

ISSN 2413-7383

Problems of Artificial Intelligence

РА

Международный
рецензируемый
научно-теоретический
журнал

International
peer-reviewed
scientific
journal

**ПРОБЛЕМЫ
ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА**

2025'4(39)

ISSN 2413-7383

Министерство науки и высшего образования
Российской Федерации

Федеральное государственное бюджетное
научное учреждение «Институт проблем
искусственного интеллекта»

**ПРОБЛЕМЫ
ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА**

2025 № 4 (39)

Международный
рецензируемый
научно-теоретический
журнал

ДОНЕЦК 2025

ISSN 2413-7383

Ministry of science and higher education
of the Russian Federation

The Federal State Budgetary
Scientific Institution «Institute of Problems
of Artificial Intelligence»

**PROBLEMS
OF ARTIFICIAL
INTELLIGENCE**

2025 № 4 (39)

International
Peer-Reviewed
Scientific
Journal

DONETSK 2025

Редакционная коллегия
Международного рецензируемого
научно-теоретического журнала
«Проблемы искусственного интеллекта» /
International Peer-Reviewed Scientific Journal
«Problems of Artificial Intelligence»

Главный редактор
Владислав Юрьевич Шелепов, д.ф.-м.н., профессор
(Донецк)

Члены редколлегии

С. С. Анцыферов, д.т.н., профессор (Москва)
Л. К. Бабенко, д.т.н., профессор (Таганрог)
О. О. Варламов, д.т.н. (Москва)
О. В. Даринцев, д.т.н., профессор (Уфа)
Т. В. Ермоленко, к.т.н., доцент (Донецк)
С. Г. Капустян, д.т.н. (Таганрог)
К. Л. Метлов, д.ф.-м.н. (Донецк)
А. В. Мищенко, к.ф.-м.н. (Роканкур, Франция)
А. В. Ниценко, к.т.н. (Донецк)
В. Н. Павлыш, д.т.н., профессор (Донецк)
А. Л. Ронжин, д.т.н., профессор (Санкт-Петербург)
К. Е. Румянцев, д.т.н., профессор (Таганрог)
И. С. Сальников, к.т.н. (Донецк)
А. С. Сигов, академик РАН, профессор, д.ф.-м.н. (Москва)
Ю. С. Харин, академик НАНБ, профессор, д.ф.-м.н. (Минск)
А. А. Харламов, д.т.н., профессор (Москва)
А. А. Шарипбаев, д.т.н., д.ф.-м.н. профессор (Алмата)
В. Ю. Шелепов, д.ф.-м.н., профессор (Донецк)

Ответственный редактор

Светлана Борисовна Иванова
И.о.директора федерального государственного бюджетного
научного учреждения «Институт проблем искусственного
интеллекта» (ФГБНУ «ИПИИ»)

Ответственный секретарь

Валентина Николаевна Пигуз, н.с.
заведующий отделом компьютерных информационных
технологий ФГБНУ «ИПИИ»

Editorial Board
of the International Peer-Reviewed Scientific Journal
«Problemy iskusstvennogo intellekta» /
«Problems of Artificial Intelligence»

Chief editor

Vladislav Yurievich Shelepov, Dr. Phys.-Math., professor
(Donetsk)

Members of the Editorial Board

S. S. Antsyferov, Dr. Eng., Professor (Moscow)
L. K. Babenko, Dr. Eng., Professor (Taganrog)
O. O. Varlamov, Dr. Eng. (Moscow)
O.V. Darintsev, Dr. Eng., Professor (Ufa)
T. V. Ermolenko, PhD. Eng. (Donetsk)
S.G. Kapustyan, Dr. Eng. (Taganrog)
K. L. Metlov, Dr. D. (Donetsk)
A. V. Mishchenko, Ph.D. (Rocancourt, France)
A. V. Nitsenko, PhD. Eng (Donetsk)
V. N. Pavlysh, Dr. Eng., Professor (Donetsk)
A.L. Ronzhin Dr. Eng., Professor (St. Petersburg)
K. E. Rummyantsev, Dr. Eng., Professor (Taganrog)
I. S. Salnikov, PhD. Eng (Donetsk)
A. S. Sigov, Academician of the RAS, Professor (Moscow)
Y.S. Kharin, Academician of NASB, Professor (Minsk)
A. A. Kharlamov, Dr. Eng., Professor (Moscow)
A. A. Sharipbaev, Dr. Eng., Professor (Nur-Sultan)
V. Y. Shelepov, Dr. Phys.-Math., professor (Donetsk)

Managing Editor

Svetlana Borisovna Ivanova
Director of the Federal State Budgetary Scientific Institution
«Institute of Problems of Artificial Intelligence»
(FSBSI «IPAI»)

Executive Secretary

Valentina Nikolaevna the Department of Computer and
Information Technologies of the FSBSI «IPAI»

Свидетельство

о регистрации средства массовой информации
Серия ААА № 000114 от 14.03.2017, ISSN 2413-7383
Международный рецензируемый
научно-теоретический журнал
«Проблемы искусственного интеллекта» /
«Problems of Artificial Intelligence» /
внесен в перечень
ВАК РФ от 25 сентября 2024 г.

В журнале публикуются
основные научные результаты диссертаций
на соискание ученой степени кандидата наук,
на соискание ученой степени доктора наук
по специальностям:

- 1.2.1. Искусственный интеллект и машинное обучение
- 1.2.2. Математическое моделирование, численные
методы и комплексы программ
- 1.2.3. Теоретическая информатика, кибернетика
- 2.3.1. Системный анализ, управление и обработка
информации, статистика
- 2.3.3. Автоматизация и управление технологическими
процессами и производствами
- 2.3.5. Математическое и программное обеспечение
вычислительных систем, комплексов
и компьютерных сетей
- 2.3.6. Методы и системы защиты информации,
информационная безопасность
- 2.3.8. Информатика и информационные процессы
- 2.5.4. Роботы, мехатроника и робототехнические системы

Рекомендовано к печати
Ученым советом ФГБНУ «ИПИИ»
Протокол № 8 от 21 ноября 2025

Международный рецензируемый научно-теоретический журнал «Проблемы искусственного интеллекта» /
«Problems of Artificial Intelligence», ISSN 2413-7383 зарегистрирован в:
научометрической базе **РИНЦ SCIENCE INDEX** (Лиц. договор № 306-05/2016 от 25 мая 2016 г.);
БД РЖ ВИНТИ РАН (Договор о сотрудничестве от 27 марта 2017 г.);
ICI Journals Master List Index Copernicus International (ICI) (17.09.2020 г.);
ЭБ Cyberleninka (Договор о сотрудничестве от 13 марта 2020 г.);
РНБ ФГБУ «Российская национальная библиотека», г. Санкт-Петербург
(Лицензионный Договор от 29 августа 2023 г.)

Mass Media Registration Certificate

Series AAA № 000114 dated by 14.03.2017, ISSN 2413-7383
International Peer-Reviewed Scientific Journal «Problemy
iskusstvennogo intellekta» /
«Problems of Artificial Intelligence»
is included into the VAK DPR list of peer-reviewed
scientific periodicals
Order of the MES RF of September 25, 2024

The journal approves basic scientific results
of Ph.D. and Doctoral dissertations:
by specialties:

- 1.2.1. Artificial intelligence and machine learning
- 1.2.2. Mathematical modeling, numerical methods
and software packages
- 1.2.3. Theoretical computer science, cybernetics
- 2.3.1. System analysis, information management and
processing, statistics
- 2.3.3. Automation and control of technological
processes and productions
- 2.3.5. Mathematical and software support of computing
systems, complexes and computer networks
- 2.3.6. Methods and systems of information protection,
information security
- 2.3.8. Computer Science and information processes
- 2.5.4. Robots, mechatronics and robotic systems

Recommended for publication
by the Academic Senate of the FSBSI «IPAI»
Protocol No 8 dated by 21 November 2025.

Раздел 1
Section 1

Искусственный интеллект
и машинное обучение

*Artificial Intelligence
and Machine Learning*

A. I. Paramonov, K. A. Pavluchenko
Belarusian State University of Informatics and Radioelectronics
6 P. Brovki Street, Minsk, 220013, Belarus

MULTI-ASPECT ANALYSIS OF EMOTIONAL CONTEXT FOR TEXT DOCUMENT CATEGORIZATION

А. И. Парамонов, К. А. Павлюченко
Белорусский государственный университет информатики и радиоэлектроники
220013, Минск, ул. П. Бровки, 6

КАТЕГОРИЗАЦИЯ ТЕКСТОВЫХ ДОКУМЕНТОВ НА ОСНОВЕ МНОГОАСПЕКТНОГО АНАЛИЗА ЭМОЦИОНАЛЬНОГО КОНТЕКСТА

The problem of texts intellectual analysis to identify their emotional context is considered. Presents the results of a review of modern approaches to the identifying emotions in texts problem, the key limitations and issues in processing complex linguistic constructions are outlined. A hypothesis is advanced: to increase the quality of digital texts categorization by emotional tone, it is necessary to account for semantic and pragmatic characteristics of their representation (sarcastic statements, emoji, contextual dependencies and other linguistic markers). A method for multi-aspect text analysis for document categorization based on identification of their emotional context is proposed. A multi-modular architecture of the method is presented, the principal component of which is a classifier built on a transformer-based language model. A description of the developed software tools using contemporary approaches is provided. The main results of computational experiments confirming the hypothesis and demonstrating the proposed approach effectiveness are discussed. Further steps for research are outlined.

Keywords: Natural Language Processing, Sentiment Analysis, Emotional Context, Document Categorization, Transformer-based Models, RoBERTa.

Рассматривается проблема интеллектуального анализа текстов с целью выявления их эмоционального контекста. Представлены итоги обзора современных подходов к задаче выявления эмоций в текстах, обозначены ключевые ограничения и проблемы при обработке сложных лингвистических конструкций. Выдвигается гипотеза: для повышения качества категоризации цифровых текстов по эмоциональному фону необходимо учитывать семантические и прагматические особенности их представления (саркастические высказывания, эмодзи, контекстуальные зависимости и другие лингвистические маркеры). Предложен метод многоаспектного анализа текста для категоризации документов на основе выявления их эмоционального контекста. Представлена многомодульная архитектура метода, основным компонентом которой является классификатор, построенный на языковой модели на основе трансформатора. Представлено описание разработанного программного средства с использованием современных подходов. Обсуждаются основные результаты вычислительных экспериментов, подтвердившие гипотезу и продемонстрировавшие эффективность предложенного подхода. Обозначены дальнейшие шаги для исследований.

Ключевые слова: обработка текста, сентимент-анализ, эмоциональный контекст, категоризация документов, трансформерные языковые модели, RoBERTa.

Introduction

Automatic text documents classification represents one of the most pressing and challenging tasks in modern computational linguistics. And when we consider contextual categorization (for example, emotional categorization), the task complication increases of the essence. In the context of exponential growth in textual information volumes across social networks, feedback systems, and digital communication platforms, there is an increasing demand for accurate and efficient methods of analyzing the emotional context of messages. Modern users of social networks and other digital communication tools actively share their thoughts, feelings, and experiences in online environments. Digital platforms daily serve as venues for exchanging textual messages and documents among millions of people. Automated analysis of textual emotional background enables businesses to understand audience sentiment, track trends, and respond to customer inquiries more quickly and effectively. An important aspect of the considered problem is that texts posted online can serve as indicators of mental health, providing opportunities to identify warning signals, depression, stress, and other issues.

Traditional approaches are often limited to simple categorization of positive and negative sentiments and consider only direct mood markers [1]. Traditional statistical methods, such as naive Bayes classifiers and support vector machines, demonstrate limited effectiveness when processing complex cases. These approaches are primarily based on analyzing the frequency of individual words or their combinations, which does not adequately account for contextual features and semantic dependencies [2]. The importance of considering emotional state in personality analysis and information processing has been the subject of numerous studies. For example, the paper [3] explores the problem of emotional expression and the multifaceted nature of its manifestation.

Natural language (NL) is complex and rich in nuances such as sarcasm, hidden meanings, metaphors, and subjective evaluations. Sarcasm and irony create contradictions between literal and implied meaning of statements, significantly complicating the interpretation of emotional content [4]. Contextual dependencies require analysis not only of individual lexical units but also their interaction within broader semantic structures. The emergence of transformer architectures such as BERT and RoBERTa has opened new possibilities for solving Natural Language Processing (NLP) tasks due to their ability to model long-term dependencies and contextual connections [5]. Contemporary research demonstrates high effectiveness of deep learning methods for emotion recognition in texts. However, even these advanced models experience difficulties when working with Russian-language texts containing sarcasm, mixed emotions, or non-standard methods of emotional expression. Analysis of the current state of the field shows the necessity of developing a comprehensive approach that combines the advantages of deep learning with specialized methods for processing complex linguistic constructions. Experience in applying various ensemble and multi-aspect text analysis methods for identifying implicit characteristics in solving various applied tasks has shown the effectiveness of such a comprehensive approach [6].

The key idea constituting the theoretical foundation of the work is the development and application of a multi-modular approach to emotional context analysis. Substantial improvement in emotion classification efficiency is proposed to be achieved through combining such key methodological components as sarcasm processing, emoji analysis, and identification of contextual features [7]. The research assumes critical importance of integrating specialized components for processing complex linguistic phenomena characteristic of modern digital communication. An important role in analyzing texts from internet sources

today is played by accounting for emoji that accompany the text itself. The emoji analyzer represents an important component of emotion analysis that allows considering the semantic load of graphic symbols and their interaction with textual content [8]. The theoretical rationale behind the proposed approach arises from the insight that contemporary texts are characterized by multilayered emotional expression and employ a wide variety of linguistic means, including nonverbal elements. Conventional approaches to emotion analysis, which rely either on straightforward lexical classification or basic neural network architectures, prove insufficient for effectively capturing the complexity and richness of modern emotional expression in texts. Integration of global analysis covering the entire textual document as a unified semantic structure for identifying dominant emotional orientation with local analysis focusing on detailed investigation of emotional coloring of individual text sentences or other smaller fragments is proposed.

The main hypothesis underlying this work is the possibility of significantly improving the accuracy of determining emotional coloring in text through combined application of specialized components for processing various text features.

1 Text Analysis Method using Transformer-based Model

A method for multi-aspect analysis of NL texts is proposed to determine the spectrum of emotions in the analyzed texts. The method is based on a multi-module architecture that includes the following major components: an emotion classifier, an emoji analyzer, a sarcasm detector, a contextual analyzer, and a compositional mechanism. The interaction between components is organized according to the principle of sequential modification of the base emotion vector, considering additional factors revealed at each stage of analysis (Figure 1). The key component of the proposed method is the emotion classifier. The classifier involves using a transformer-based model that implements the modern RoBERTa encoder architecture [9]. The encoder architecture was specially adapted and additionally trained for solving emotion recognition tasks in texts. Transformer models demonstrate effectiveness in multilingual emotion analysis tasks [10].

Text input to the RoBERTa classifier undergoes preprocessing. After cleaning and normalization, tokenization is performed using the RoBERTa Tokenizer tool, ensuring preservation of text structure and its contextual connections. Using the «Byte Pair Encoding» text segmentation algorithm allows the model to effectively process rare or unknown words by breaking them into more understandable and frequently occurring parts, helping the model work with a sufficiently wide range of texts.

The model is trained to recognize nine emotional states: eight basic ones (wariness, delight, admiration, horror, astonishment, grief, disgust, and anger) as well as a neutral undefined state. For training the developed model the modern AdamW optimization algorithm is applied, representing an improved version of the classic Adam optimizer with correct implementation of weight decay. This algorithm demonstrates high characteristics in transformer architecture training tasks due to its ability to effectively manage the gradient descent process when working with high-dimensional parameter spaces. AdamW ensures accelerated model convergence to optimal parameter values through adaptive learning rate management for each parameter individually, allowing the model to more effectively adapt to complex patterns in data [11].

The emoji analyzer component (step 2 in Figure 1) represents a classifier of emoji symbols identified in the text. Each emoji symbol is matched with a predetermined dictionary where each symbol is assigned a specific emotional category. Thus, coefficients

for correcting the basic emotion vector are formed. When calculating the contribution of emoji to the document's emotional context, not only quantitative indicators are considered, but also their application methods (combinations, placement throughout the text, etc.).

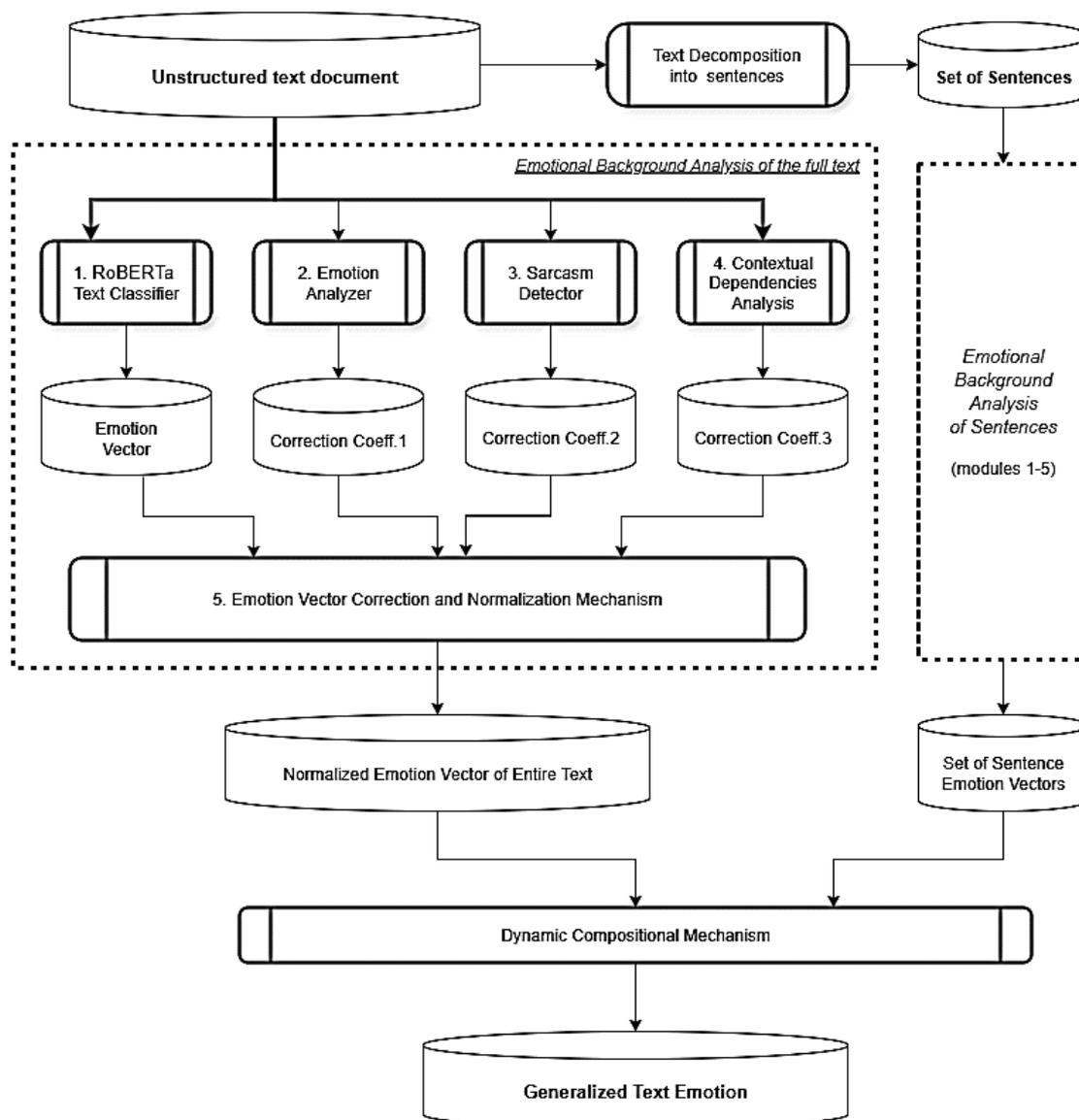


Figure 1. General structural scheme of the proposed method.

The sarcasm detector is based on recognizing characteristic linguistic patterns typical of sarcastic statements [12]. At this stage, a dictionary of markers and typical features is used, including special intensification markers. To identify sarcasm features, sentence structure analysis and syntactic analysis of punctuation mark combinations are performed. All identified sarcasm features are combined using weighted evaluation, where the weight of each feature is determined based on its statistical significance for sarcasm detection. With high sarcasm probability, emotion vector inversion or its substantial modification is possible. The contextual analyzer serves to identify logical and emotional connections between sentences – how they develop, strengthen, or change the emotional line of the text.

Contextual analysis plays a key role in modern NLP systems. The obtained set of coefficients is applied to the emotion vector for correction.

The final component of the developed architecture is the compositional mechanism, representing an intelligent system for synthesizing results. This mechanism functions as a central coordination node, ensuring mathematically justified combination of results from all specialized system components into a unified holistic and consistent assessment of the emotional context of the analyzed text. The compositional mechanism implements multilevel logic for weighting and integrating heterogeneous factors identified at previous analysis stages, including basic emotional vectors from the transformer model, corrections from the sarcasm detector, modifications from the emoji analyzer, and contextual adjustments from the specialized contextual analyzer. A key feature of this mechanism is its ability for adaptive management of weight coefficients depending on the characteristics of the specific analyzed text and the confidence level of each system component. The overall emotion score of a text is calculated as a weighted sum of all emotion vectors and their weighting coefficients. The weighting coefficients are determined based on the confidence of the classifiers. A distinctive feature of the compositional mechanism is the implementation of a dynamic weight coefficient correction system based on confidence metrics of individual classification components. This software functions according to the principle of adaptive trust, where components demonstrating high confidence in their predictions (determined through analysis of output vector probability distributions) automatically receive increased weight coefficients in the process of forming the final assessment.

After completing all stages of correction and composition of results, a procedure for normalizing the final emotion vector is performed. This operation is carried out using L1-normalization. The normalization procedure incorporates additional measures ensuring mathematical correctness, specifically by verifying the absence of negative values, controlling numerical stability throughout operations, and addressing specific edge cases like zero or infinite vector component values. The final normalized emotion vector represents the ultimate product of the entire analytical system, embodying a comprehensive multi-factor assessment of the emotional context of the analyzed textual document. This vector integrates the results of basic transformer classification, specialized processing of sarcasm and irony, semantic analysis of emoji and graphic elements, contextual analysis of linguistic features, and considers all identified factors of emotional expressiveness and communicative characteristics of the source text.

2 Software Tools for Emotional Background Detection

The proposed models and methods are implemented in the form of a specialized software tools. The developed software represents a comprehensive system for automated analysis of textual data emotional context, built on principles of modular architecture and modern approaches to NLP. The software is implemented using the Python programming language and integrates advanced machine learning libraries and NLP tools to ensure high precision in emotion classification.

The technical foundation of the software is based on the PyTorch framework, which enables efficient work with transformer architectures and support for graphics processing unit (GPU) computations to accelerate training and inference processes. Integration with the Transformers library from Hugging Face provides access to pre-trained RoBERTa models and tools for their fine-tuning for specific emotion classification tasks. For loading and executing the transformer model, performing emotion classification in text, text tokenization,

manipulation of emotion vectors, and carrying out mathematical operations, libraries such as PyTorch (torch), Transformers, nltk, NumPy and scikit-learn (sklearn) were employed. For working with emojis, specifically for extraction and identification of emojis in text, the Emoji library (emoji) was used. Methods from the tqdm and Matplotlib libraries (tqdm, matplotlib) were also used for visualization of the training process and presentation of analysis results. The software tool is organized as a set of interconnected modules, each responsible for a specific aspect of textual data processing and analysis.

The data preprocessing module performs normalization of input text, including cleaning from irrelevant symbols, correction of common typos, unification of emoji and special symbol representation. Tokenization is performed using RoBERTa Tokenizer, which ensures correct text splitting while maintaining semantic integrity.

The main classification module represents a system component that implements the interaction logic with a modified RoBERTa transformer architecture that has undergone specialized additional training for multi-class emotion classification tasks within eight target categories of the emotional spectrum. RoBERTa-type models, considering their size and complexity, require significant computational resources, especially when training on large datasets. It was considered when designing the software architecture. The classification module includes numerous functional capabilities: automated mechanisms for loading and initializing pre-trained model weight parameters, a system for performing inference procedures on new input textual data, generation of multidimensional probabilistic vectors of emotional states with accompanying quantitative metrics of reliability and confidence of classification decisions. The module provides dynamic computational resource management, automatic determination of optimal execution device (CPU or GPU), and implements exception handling mechanisms to ensure system stability under various operating conditions.

The emoji analysis module implements semantic matching of graphic symbols with textual content based on the CLIP model. The component extracts emoji from text, calculates their semantic proximity to textual content through cosine distance in vector space, and generates corrective coefficients for the basic emotional vector.

To identifying sarcastic statements, a combination of linguistic rules and statistical patterns is used (the "sarcasm detector" module). To context analysis dictionaries of linguistic markers and rules for their combination are used to generate contextual corrections.

The software tools provide several interaction interfaces for various usage scenarios. The following operating modes are supported: real-time mode for analyzing individual messages with minimal delay; batch mode for processing large document collections with performance optimization; and continuous monitoring mode for analyzing streaming data from social networks or other open external sources.

The software architecture ensures a high degree of configurability through a system of configuration files in JSON and YAML formats. Users can configure weight coefficients of various components, threshold values for sarcasm detection, emotional marker dictionaries, and other system parameters without the need to modify source code. The software tool configuration allows organizing assembly for efficient operation on both personal computers and high-performance server systems. GPU acceleration support through CUDA provides significant acceleration of analysis processes when working with large data volumes. The software includes mechanisms for caching intermediate results and memory optimization for processing long texts and large data batches. The architecture supports horizontal scaling through distributing computational load among multiple worker processes or servers, allowing the application to be adapted to various performance requirements and volumes of processed data.

3 Computer Experiment and Discussion

To evaluate the effectiveness of the proposed method and the software developed based on it, a computer experiment was conducted. Within the framework of the experiment, the task was set to assess the proposed approach in comparison with existing classification methods, models, and services.

For conducting the experiment, a specialized dataset «ru_go_emotions» was used, containing annotated Russian-language texts with emotional categories [13]. Additionally, a test sample was formed from social network comments that includes examples of sarcasm, metaphorical expressions, and ambivalent statements. Such a dataset is indicative for evaluating the method's capabilities to handle ambiguous cases. The experiment was conducted on a sample that included 1000 randomly selected texts from the «ru_go_emotions» dataset. For each text from the set, emotional context assessment was performed using two software solutions:

- A service based on the modern generative language model GPT;
- Software based on the proposed multi-aspect emotional context analysis method.

To ensure correctness of result comparison, adaptation of the dataset classification scheme to the "Flower of Emotions" model variation using a palette of eight emotions, which was used in this work, was performed. Evaluation was carried out by comparing predictions of each solution with expert annotations of the dataset.

Detailed analysis of prediction distribution revealed fundamental differences in architectural approaches and methodological principles of the two systems for solving multi-class emotion classification tasks, demonstrating critical limitations of universal generative models and advantages of specialized emotional analysis systems. GPT demonstrated serious problems with classification balancing, manifested in extremely uneven distribution of predictions across emotion categories (Table 1).

Table 1 – Number of detected documents (for each of selected emotion)

Emotion	Expert assessment	GPT Service	Software of the Proposed Method
Wary	147	443	51
Delight	168	383	62
Admiration	291	56	70
Horror	117	0	287
Astonishment	72	5	298
Grief	68	17	124
Disgust	90	50	68
Anger	47	46	33

At the same time, the software proposed in this work demonstrated a fundamentally different classification paradigm based on deep understanding of emotional structure and specialized balancing algorithms. A significant achievement of the developed approach is the complete elimination of classification "dead zones" characteristic of GPT. The phenomenon of absolute ignoring of the horror emotion (zero detections from the dataset) by GPT represents a fundamental architectural problem of generative model's incapable of adequate categorization of intense negative emotional states.

The series of experiments allows concluding that the software complex presented in this work demonstrated superiority over the modern generative GPT model. On the test sample of thousand texts, the authors' method achieved a classification accuracy of over 56 percent, which is eighteen percentage points higher than GPT's performance (with GPT

achieving thirty-eight percent). Various non-trivial cases were examined: sarcastic use of emoji and negation of reliability, mixed emotions and technical terminology, mixed emotions in long text, contradictory emotions in short sentences, and others (Table 2).

Coefficient of variation – a dimensionless indicator, calculated as the ratio of standard deviation to the arithmetic mean of the number of predictions across 8 emotion classes. Standard deviation – measured in the number of predictions, shows the absolute spread from the mean value of 125 predictions per class. Herfindahl concentration index – a dimensionless indicator from 0 to 1, calculated as the sum of squares of each class's proportions, shows the degree of prediction concentration. Maximum deviation from uniformity – represents the largest deviation of any class proportion from the ideal uniform distribution of 12.5% for 8 classes (measured in percentages).

Table 2. Detailed analysis of four key statistical metrics.

Metric	GPT-based Approach	Proposed Approach	Enhancement
Coefficient of variation	0.89	0.31	x 2.8
Standard deviation	128.4	37.2	x 3.5
Herfindahl concentration index	0.26	0.13	x 2
Maximum deviation from uniformity	22.8%	6.4%	x 3.5

The conducted experiments validated the efficacy of the proposed method in handling diverse configurations of complex texts across multiple datasets. For example, when analyzing text with heavy emotional context. In this case, the GPT model showed uncertain classification with low probabilities, while multi-aspect analysis correctly determined the final emotion value as the negative emotion of sadness with confidence of seventy-eight hundredths. "Sarcasm Detector" module showed remarkable effectiveness in handling complex cases. Specific test scenarios included the use of quotation marks around keywords, ironic juxtapositions, and inconsistencies between positively phrased sentences and negatively connoted emojis.

The results indicate substantial progress in solving the task of automatic emotion classification in texts. The achieved accuracy improvement of approximately thirty percent compared to commercial solutions demonstrates the effectiveness of a specialized approach for specific NLP tasks. Comparison with commercial solutions reveals advantages the research approach's superior accuracy in processing complex linguistic constructions. Commercial products often use simplified algorithms to ensure high processing speed, which negatively affects the quality of analysis of non-standard cases.

The developed software showed high effectiveness in working with hidden emotional contexts, the ability to distinguish neutral declarative tone from emotionally colored. The tool demonstrates the capability to accurately determine the overall emotion of a text and its variations, even in cases where individual sentences yield conflicting analyses. The ability to correctly interpret conflicts between textual content and graphic elements is critically important for analyzing messages in Social Net and messengers, where such contradictions occur regularly. Experimental validation confirms the feasibility of the proposed approach to adaptive adjustment of initial predictions based on contextual factors. Particular attention was paid to analyzing stability with rare emotion classes and performance indicators in scenarios where emotions overlap or are distributed among different parts of the text.

It's important! Limitations of the method are related to the computational complexity of transformer architectures and the need for significant resources for training.

Conclusion

Application areas of the developed method include social media monitoring systems for public opinion analysis and identification of potential crisis situations, e-commerce platforms for automatic customer review analysis and service quality improvement, psychological support systems for early detection of some mental disorders through text messages analysis, and content moderation tools for automatic identification of aggressive or destructive messages.

The scientific results of this research work include the development of a specialized sarcasm detector for Russian-language texts, integration of an emoji analyzer based on the CLIP model for semantic matching of graphic symbols and textual content. The hypothesis about the effectiveness of using multi-aspect text analysis to improve the quality of document categorization was experimentally confirmed. The developed architecture can serve as a foundation for creating more complex NLP systems with in-depth analysis of human emotions and intentions. The obtained results open perspectives for further development of emotion analysis methods toward supporting multilingualism, integrating additional modalities, and expanding the spectrum of recognizable emotional states.

References

1. Gelbukh, A. (ed.) (2013). *Computational Linguistics and Intelligent Text Processing*. Proceedings of the 14th International Conference, CICLing 2013, Samos, Greece, March 24–30, 2013, Part I. Springer-Verlag Berlin Heidelberg.
2. Dvoynikova, A.A., Karpov, A.A. (2020). Analytical Review of Approaches to Sentiment Recognition in Russian-Language Text Data. *Information Control Systems* 4, pp. 20–30. (In Russian)
3. Salnikov, I. S., & Salnikov, R. I. (2024). Problemy i voprosy emotsional'noi ekspressii i patternializatsii emotsii v issledovaniyakh Kerrolla Izarda i ego predshestvennikov, sovremennikov i posledovatelei [Problems and issues of emotional expression and emotional patterning in the research of Carroll Izard and his predecessors, contemporaries, and followers]. *Problemy Iskusstvennogo Intellekta*, (4), pp. 193–209. (In Russian)
4. Rodriguez, M., Thompson, K., Lee, S. (2022). Sarcasm Detection in Social Media: Challenges and Opportunities. *Journal of Natural Language Processing* 15(3), pp. 45–62.
5. Gardazi, N. M., et al. (2025). BERT applications in NLP: A review. *Artificial Intelligence Review*.
6. Trukhanovich, I., Paramonov, A. (2023). Multispecies Ensemble Architecture for Texts Authorship Classification. In: *Proceedings of the 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, IEEE, pp. 142–148.
7. Anderson, R., Wilson, P., Garcia, C. (2022). Contextual Analysis in Natural Language Processing: Recent Advances and Applications. *ACM Computing Surveys* 55(8), pp. 1–34.
8. Rogers, A., Romanov, A., Rumshisky, A., Volkova, S., Gronas, M. & Gribov, A. (2018). RuSentiment: An Enriched Sentiment Analysis Dataset for Social Media in Russian. In E. M. Bender, L. Derczynski & P. Isabelle (eds.), *COLING* (p./pp. 755-763), : Association for Computational Linguistics. ISBN: 978-1-948087-50-6.
9. Liu, Y., Ott, M., Goyal, N. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*.
10. Kim, Y., Park, J., Choi, D. (2023). Transformer-Based Models for Multilingual Emotion Analysis. *Computational Linguistics* 49(2), pp. 287–305.
11. Loshchilov, I., Hutter, F. (2019). Decoupled Weight Decay Regularization. *arXiv preprint arXiv:1711.05101*.
12. Chen, J., Wang, L., Zhang, H. (2023). Deep Learning Approaches for Emotion Recognition in Text: A Comprehensive Survey. *Computer Speech & Language* 78, pp. 101–118.
13. Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., Ravi, S. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 4040–4054.

Список литературы

1. Gelbukh A. (ed.). *Computational Linguistics and Intelligent Text Processing*. Proceedings of the 14th International Conference, CICLing 2013, Samos, Greece, March 24–30, 2013, Part I. Berlin Heidelberg: Springer-Verlag, 2013.
2. Двойникова, А. А. Аналитический обзор подходов к распознаванию тональности в русскоязычных текстовых данных / А. А. Двойникова, А. А. Карпов // *Системы управления и информационные технологии*. –2020. № 4. С. 20–30.
3. Сальников, И. С. Проблемы и вопросы эмоциональной экспрессии и паттерниализации эмоций в исследованиях Кэрролла Изарда и его предшественников, современников и последователей / И. С. Сальников, Р. И. Сальников // *Проблемы искусственного интеллекта*. 2024. № 4. С. 193–209.

4. Rodriguez M., Thompson K., Lee S. Sarcasm Detection in Social Media: Challenges and Opportunities // Journal of Natural Language Processing. 2022. Vol. 15, No. 3. P. 45–62.
5. Gardazi N. M. et al. BERT applications in NLP: A review // Artificial Intelligence Review. 2025.
6. Trukhanovich I., Paramonov A. Multispecies Ensemble Architecture for Texts Authorship Classification // Proceedings of the 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). IEEE, 2023. P. 142–148.
7. Anderson R., Wilson P., Garcia C. Contextual Analysis in Natural Language Processing: Recent Advances and Applications // ACM Computing Surveys. 2022. Vol. 55, No. 8. P. 1–34.
8. Rogers A., Romanov A., Rumshisky A., Volkova S., Gronas M., Gribov A. RuSentiment: An Enriched Sentiment Analysis Dataset for Social Media in Russian // COLING / ed. by E. M. Bender, L. Derczynski, P. Isabelle. Association for Computational Linguistics, 2018. P. 755–763.
9. Liu Y., Ott M., Goyal N. RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692, 2019.
10. Kim Y., Park J., Choi D. Transformer-Based Models for Multilingual Emotion Analysis // Computational Linguistics. 2023. Vol. 49, No. 2. P. 287–305.
11. Loshchilov I., Hutter F. Decoupled Weight Decay Regularization. arXiv preprint arXiv:1711.05101, 2019.
12. Chen J., Wang L., Zhang H. Deep Learning Approaches for Emotion Recognition in Text: A Comprehensive Survey // Computer Speech & Language. 2023. Vol. 78. P. 101–118.
13. Demszky D., Movshovitz-Attias D., Ko J., Cowen A., Nemade G., Ravi S. GoEmotions: A Dataset of Fine-Grained Emotions // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. P. 4040–4054.

RESUME

A. I. Paramonov, K. A. Pavluchenko

Multi-Aspect Analysis of Emotional Context for Text Document Categorization

Traditional sentiment analysis methods are limited to simple positive/ negative categorization and fail to adequately process complex linguistic constructions. Statistical approaches like naive Bayes and support vector machines demonstrate limited effectiveness due to their reliance on word frequency analysis without accounting for contextual features and semantic dependencies. Modern texts contain nuanced elements such as sarcasm, mixed emotions, emoji, and contextual dependencies that significantly complicate emotional content interpretation.

A multi-aspect text analysis method is proposed, based on a multi-modular architecture integrating specialized components: an emotion classifier built on the RoBERTa transformer-based model trained to recognize nine emotional states, an emoji analyzer using the CLIP model for semantic matching of graphic symbols with textual content, a sarcasm detector based on linguistic patterns and statistical markers, a contextual analyzer for identifying logical and emotional connections between sentences, and a compositional mechanism for intelligent synthesis of results through weighted combination of all components. The software implementation uses Python with PyTorch, Transformers library, and AdamW optimization algorithm. Evaluation was conducted on the ru_go_emotions dataset containing 1000 annotated Russian-language texts.

The proposed method achieved classification accuracy of 56%, exceeding the GPT-based commercial solution by 18 percentage points (GPT: 38%). Statistical analysis revealed significant improvements: coefficient of variation decreased 2.8-fold, standard deviation improved 3.5-fold, and Herfindahl concentration index improved 2-fold. The method demonstrated complete elimination of classification "dead zones" characteristic of generative models, successfully handling complex cases including sarcastic emoji usage, mixed emotions in long texts, and contradictory emotions in short sentences.

The computational experiments validated the effectiveness of combining neural network models with postprocessing algorithms for emotional context analysis. The method successfully accounts for subtle aspects of natural language including contextual mentions, sarcasm, and ambiguous cases. The developed architecture provides a foundation for creating advanced natural language understanding systems with in-depth emotion and intention analysis, with applications in social media monitoring, e-commerce review analysis, psychological support systems, and content moderation tools.

РЕЗЮМЕ

А. И. Парамонов, К. А. Павлюченко

Категоризация текстовых документов

на основе многоаспектного анализа эмоционального контекста

В данной статье разработан метод многоаспектного анализа текстов на естественном языке для определения спектра эмоций в анализируемых документах на основе многомодульной архитектуры. Предложена архитектура, включающая классификатор эмоций на базе трансформерной модели RoBERTa, анализатор эмпатии с использованием модели CLIP для семантического сопоставления графических символов с текстовым содержанием, детектор сарказма на основе распознавания характерных лингвистических паттернов, контекстуальный анализатор для выявления логических и эмоциональных связей между предложениями и композиционный механизм для интеллектуального синтеза результатов всех компонентов системы. Разработано программное средство с использованием языка Python, библиотек PyTorch и Transformers, алгоритма оптимизации AdamW. Проведены вычислительные эксперименты на датасете *ru_go_emotions*, содержащем 1000 аннотированных русскоязычных текстов. Результаты экспериментов показали точность классификации 56%, что превышает показатели коммерческого решения на основе GPT на 18 процентных пункта. Коэффициент вариации улучшен в почти 3 раза, стандартное отклонение – в 3,5 раза, а индекс концентрации Херфиндаля – в 2 раза. Метод продемонстрировал полное устранение «мёртвых зон» классификации и эффективную обработку сложных случаев, включая саркастическое использование эмпатии, смешанные эмоции в длинных текстах и противоречивые эмоции в коротких высказываниях. Разработанные модели и методы позволяют решить задачу повышения качества категоризации цифровых текстов по эмоциональному фону за счёт учёта семантических и прагматических особенностей их представления и могут быть использованы в системах мониторинга социальных сетей, анализа отзывов в электронной коммерции, психологической поддержки и модерации контента.

Парамонов Антон Иванович – к.т.н., доцент, заведующий кафедрой информационных систем и технологий, учреждение образования «Белорусский государственный университет информатики и радиоэлектроники», г. Минск, Республика Беларусь.

Область научных интересов: Интеллектуальные системы. Компьютерная лингвистика и АОТ. Когнитивная психология. эл. почта a.paramonov@bsuir.by, адрес: 220005, г. Минск, ул. Козлова, 28, телефон: +375293451172.

Павлюченко Кирилл Алексеевич – магистр, ассистент кафедры информационных систем и технологий, учреждение образования «Белорусский государственный университет информатики и радиоэлектроники», г. Минск, Республика Беларусь.

Область научных интересов: Эмоциональный интеллект. Машинное обучение. эл. почта gu_mendorill@mail.ru, адрес: 220005, г. Минск, ул. Козлова, 28, телефон: +375173776760

Статья поступила в редакцию 22.08.2025.

А.Г. Гурин, А.В. Григорьев
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Донецкий национальный технический университет»
283001, Донецкая Народная Республика, г. Донецк, ул. Артёма, 58

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ПРОДУКЦИОННЫХ СИСТЕМ ДЛЯ РАБОТЫ С БОЛЬШИМИ ДАННЫМИ. КОМБИНАЦИЯ ДВУХ ПОДХОДОВ

A.G. Gurin, A.V. Grigoriev
Federal State Budgetary Educational Institution of Higher Education "Donetsk National Technical
University" 283001, Donetsk People's Republic, Donetsk, st. Artyoma, 58.

COMPARATIVE ANALYSIS OF METHODS FOR ENHANCING THE EFFICIENCY OF PRODUCTION SYSTEMS FOR BIG DATA PROCESSING

Проведен анализ логического вывода в производственных системах для задач обработки больших данных. Несмотря на эффективность классических производственных систем, недостаточно изученной остается комбинация двух подходов: использование И-ИЛИ-деревьев для наглядного представления логики и применение тензорных представлений для повышения производительности. Исследование демонстрирует взаимодополняемость данных подходов, что обосновывает возможность создания адаптивного мета-алгоритма, динамически выбирающего метод в зависимости от характеристик решаемой задачи. Для образовательных систем подобный гибридный подход особенно перспективен, так как позволяет сочетать высокую производительность при обработке данных с полной интерпретируемостью принимаемых решений.

Ключевые слова: производственные системы, И-ИЛИ-дерево, тензорные представления, онтологии, образовательные системы.

An analysis of logical inference in production systems for big data processing tasks has been conducted. Despite the effectiveness of classical production systems, the combination of two approaches remains understudied: the use of AND-OR trees for intuitive logic representation and the application of tensor representations for performance enhancement. The research demonstrates the complementarity of these approaches, justifying the feasibility of creating an adaptive meta-algorithm that dynamically selects the method based on the characteristics of the problem being solved. For educational systems, such a hybrid approach is particularly promising, as it allows for combining high data processing performance with complete interpretability of the decisions made.

Key words: production systems, AND-OR-trees, tensor representations, ontologies, education system.

Введение

Предлагаемая работа направлена на развитие ряда положений предшествующих авторских работ [1-3], посвященных применению теоретико-множественных операций над порождающими грамматиками в продукционных системах автоматизированного проектирования программного обеспечения (САПР ПО). Спецификой данного подхода есть сужение порождающих грамматик до уровня И-ИЛИ-дерева, как инструмента, реализующего прецедентный подход в задачах принятия решений.

Ключевое отличие данной работы заключается в исследовании путей расширения математического аппарата теоретико-множественных операций над порождающими грамматиками, предложенных в указанных работах, новыми математическими сущностями, такими как тензора и полугруппы. В данной работе исследуется предлагаемый гибридный подход к организации логического вывода, рассматриваемого как комбинация тензорного подхода и И-ИЛИ-дерева. Обсуждаются специфика и достоинства такого подхода. Продукционные системы, основанные на правилах вида «если-то», формализуют причинно-следственные связи, что обуславливает их аналогию с формальными грамматиками, в частности, порождающими. Порождающая грамматика представляет собой классический формализм, задающий правила порождения строк или структур. В этом контексте продукционные системы могут рассматриваться как их обобщение, где правила применяются не только к строкам, но и к состояниям систем в различных предметных областях. Аспекты использования продукционных систем в рамках порождающих грамматик детально рассмотрены в [4-7].

Порождающая грамматика является не только теоретической конструкцией, но и практическим инструментом, обеспечивающим гибкость, автоматизацию и надежность при преобразовании пользовательских формул в исполняемые алгоритмы в системах сценарного моделирования [8]. Широкое применение продукционный подход находит в искусственном интеллекте, экспертных системах и автоматизации логического вывода. Рост сложности прикладных задач порождает потребность в интеграции продукционных систем с другими математическими методами, такими, как И-ИЛИ-дерева и тензорные вычисления. Подобный синтез потенциально позволяет повысить эффективность обработки многомерных данных и сложных логических структур [9], а также открывает возможность использования физической семантики для создания сложных интеллектуальных систем. Элементарная семантика уже является важным аспектом различных интеллектуальных систем [10], [11].

Практическая потребность в развитии рассматриваемых методов особенно очевидна в контексте современных образовательных и государственных информационных систем. Развитие технической инфраструктуры таких систем является критически важным фактором их эффективности. В этой связи в различных предметных областях предпринимаются попытки интеграции сложных математических и логических структур, включая онтологии и продукционные модели [10], [12-15]. Одновременно с этим, актуальной задачей становится применение онтологий при разработке информационных систем, в частности, в связи с внедрением концепции цифровой аналитической платформы для предоставления статистических данных, где ключевым требованием выступает использование онтологии данных, заложенной в модель государственных данных [16], [17]. Таким образом, возникает задача поиска эффективных методов для работы с комплексными онтологиями, объединенными с продукционными правилами.

Целью данной работы является проведение сравнительного анализа методов интеграции продукционного подхода с И-ИЛИ-деревьями и тензорными представлениями для оценки их формальных основ, применимости и ограничений при работе с комплексными онтологиями, объединенными с продукционными правилами.

Постановка задачи включает решение следующих ключевых аспектов:

- формализация математических моделей продукционных систем;
- формализация аппарата И-ИЛИ-деревьев для структурирования логического вывода;
- формализация тензорного представления для высокопроизводительных вычислений;
- проведение сравнительного анализа методов;
- исследование перспективных направлений развития.

Ключевым объединяющим аспектом рассматриваемых методов является их способность формализовать причинно-следственные связи. И-ИЛИ-деревья обеспечивают визуальную интерпретацию сложных условий через аналогию с формальными грамматиками, а тензорные представления реализуют те же правила в виде операций линейной алгебры, заимствуя принципы ковариантности из физических моделей. Исследование тензорного подхода в контексте продукционных систем направлено на создание структурного и математически строгого аппарата для работы со сложными онтологиями, выходящего за рамки классических методов.

Базовые принципы и математические модели продукционных систем

Продукционный подход формализуется в виде системы, основанной на правилах, рабочей памяти и механизме логического вывода. В контексте исследования методов интеграции с И-ИЛИ-деревьями и тензорными представлениями рассмотрим формальную структуру продукционных систем.

Продукционная система определяется тройкой [18]:

$$PS = \langle R, D, C \rangle, \quad (1)$$

где:

- $R = \{r_1, r_2, \dots, r_n\}$ — множество продукционных правил;
- D — рабочая память (база данных) - множество фактов или утверждений, например: $D = \{A, B, C, \dots\}$;
- $C: R \times D \rightarrow R$ — функция разрешения конфликтов, определяющая порядок активации правил. Основные стратегии включают приоритет по специфичности антецедента, приоритет по порядку добавления (FIFO) и приоритет по весу правил.

Каждое правило r_i имеет форму:

$$P \rightarrow Q \text{ (ЕСЛИ } P, \text{ ТО } Q), \quad (2)$$

где P — антецедент (условие), Q — консеквент (действие). Антецедент P может представлять собой булеву комбинацию условий:

$$P = A \wedge B \vee C. \quad (3)$$

В системах с неопределенностью добавляются веса или коэффициенты уверенности:

$$r_i: P \rightarrow Q \text{ (CF} = 0.8). \quad (4)$$

В логике предикатов правило может включать переменные:

$$\forall x (P(x) \rightarrow Q(x)), \quad (5)$$

например, $\forall x (\text{Студент}(x) \wedge \text{СдалЭкзамен}(x) \rightarrow \text{ДопущенКЗащите}(x))$, $x = \text{Иван}$.

Процесс вывода и его виды:

– сопоставление: для каждого правила r_i проверяется, выполняется ли его антецедент P в текущей рабочей памяти D . Формально: $D \models P$.

– активация: при истинности P , консеквент формируется в массив Q и добавляется в D или модифицирует его:

$$D_{\text{нов}} = D_{\text{стр}} \cup Q; \quad (6)$$

– цикл "Выбор-Действие": процесс повторяется, пока применимые правила не будут исчерпаны.

Процесс вывода демонстрируется на примере системы диагностики: при наличии в рабочей памяти фактов {ДвигательНеЗаводится, СветФарТусклый} активируется правило r_1 : ЕСЛИ ДвигательНеЗаводится \wedge СветФарТусклый, ТО РазряженАккумулятор, что приводит к модификации рабочей памяти и последующей активации правила r_2 . Больше примеров для понимания работы можно найти в трудах Shortliffe [19].

Продукционные системы допускают различные математические интерпретации:

– в теории автоматов: система представляется как конечный автомат $\langle S, \Sigma, \delta, s_0 \rangle$, где S — множество состояний (содержимое D), Σ — алфавит операций (правила R), δ — функция переходов (применение правил);

– в теории графов — правила и факты образуют ориентированный граф, где ребра — импликации $P \rightarrow Q$.

– в логике Хорна: каждое правила эквивалентно дизъюнктам Хорна [20]:

$$\neg P \vee Q. \quad (7)$$

К основным ограничениям относятся:

– возможность циклов при взаимозависимых правилах ($A \rightarrow B, B \rightarrow A$), решением которого является введение счетчиков активаций или мета-правил для прерывания циклов [21], [22];

– немонотонность вывода, так как добавление новых фактов может инвалидировать предыдущие выводы, решением которого является механизмы отката или нечеткая логика [22], [23].

Данная формализация обеспечивает строгий анализ продукционных систем, проектирование эффективных механизмов вывода и их интеграцию с современными методами ИИ, включая алгоритмы верификации на основе методов проверки моделей [23], [24].

Визуализация и структурирование логического вывода с помощью И-ИЛИ-деревьев

И-ИЛИ-деревья представляют формальный аппарат для структурированного представления пространства поиска решений в задачах искусственного интеллекта. В контексте интеграции с продукционными системами они обеспечивают наглядное представление сложных логических условий.

Формально И-ИЛИ-дерево определяется как ориентированный граф [21], [26]:

$$T = (V, E, \tau), \quad (8)$$

где:

– $V = V_{\text{И}} \cup V_{\text{ИЛИ}} \cup V_{\text{лист}}$ — множество вершин,

– $E \subseteq V \times V$ — множество дуг,

– $\tau: V \rightarrow \{ \text{И}, \text{ИЛИ}, \text{лист} \}$ — функция типа вершины.

Вариантом И-ИЛИ-дерева называется поддерево, полученное удалением всех дуг, кроме одной, для каждой ИЛИ-вершины. Мощностью дерева называется количество его вариантов [25].

И-ИЛИ-деревья находятся в соответствии с контекстно-свободными грамматиками, где И-узлы соответствуют конкатенации ($A \rightarrow BC$), а ИЛИ-узлы — альтернативным продукциям ($A \rightarrow B \mid C$). Это соответствие позволяет применять их для синтеза программ по техническим заданиям, автоматического порождения онтологий и верификации.

Интеграция И-ИЛИ-деревьев с продукционной системой $PS = \langle R, D, C \rangle$ осуществляется через представление антецедентов правил $r_i \in R$ в виде деревьев T_i . Логика вычисления истинности вершины v определяется рекурсивно [22]:

- если $\tau(v) = \text{И}$: значение истинно тогда и только тогда, когда истинны все потомки;
- если $\tau(v) = \text{ИЛИ}$: значение истинно тогда и только тогда, когда истинен хотя бы один потомок.

На примере правила $r: A \wedge (B \vee C) \rightarrow Q$ соответствующее И-ИЛИ-дерево представлено на рисунке 1 [22]. При активации нескольких правил выбор осуществляется стратегией конфликтного разрешения C , определенной в продукционной системе.

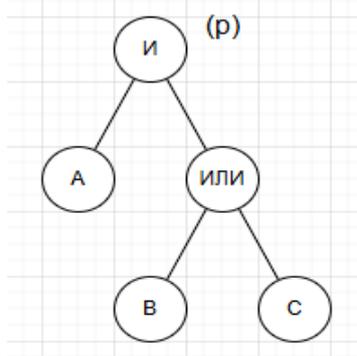


Рисунок 1 – Правило в виде дерева

Аналогичный подход к использованию формальных грамматик для преобразования графовых структур представлен в работе [27], где продукционные правила применяются для трансформации деревьев зависимостей. В исследованиях А. Григорьева [1-3] формальные грамматики используются для манипулирования И-ИЛИ-деревьями в онтологиях. К преимуществам интеграции И-ИЛИ-деревьев с продукционными системами относятся:

- структурированное представление сложных логических условий;
- эффективный инкрементальный вывод;
- естественная интеграция с логикой Хорна.

Основные ограничения метода [21]:

- экспоненциальный рост количества вариантов при ИЛИ-узлах;
- сложность верификации циклов;
- ограниченная выразительность.

Интеграция И-ИЛИ-деревьев с продукционным подходом обеспечивает высокую интерпретируемость логики принятия решений за счет структурного соответствия формальным грамматикам. Это позволяет применять методы синтаксического анализа (LL/LR), генерировать объяснения решений в виде выводимых цепочек и упрощать синтез правил через автоматизированное преобразование грамматик.

Тензорный формализм для высокопроизводительного логического вывода

Продукционная система допускает представление в виде ориентированного графа, где узлы соответствуют фактам, а дуги — правилам. Матрица смежности такого графа, имеющая размерность $m \times n$, представляет собой тензор ранга 2 и служит основой для перехода к тензорному исчислению.

Использование тензорного формализма позволяет унифицировать представление онтологий за счет объединения связанных величин и обеспечивает инвариантность уравнений относительно предметной области. Принципы физической ковариантности, используемые в тензорном формализме [28], обеспечивают инвариантность правил относительно перекодировки фактов, что согласуется с подходом к работе с физической семантикой [29].

Для формализации процесса логического вывода рабочая память D представляется как бинарный вектор $d \in \{0,1\}^n$ [21], [30], [31]:

$$d_j = \begin{cases} 1, & \text{если } f_j \in D \\ 0, & \text{иначе} \end{cases} \quad (9)$$

Однако простое умножение матрицы смежности на бинарный вектор не учитывает условия, требующие нескольких фактов одновременно. Для решения этой проблемы применяется трехмерное тензорное представление [32], где выделяются три оси:

- факты - $J = \{j_1, \dots, j_k\}$ индексируются в векторном пространстве;
- правила - $I = \{i_1, \dots, i_n\}$ индексируются в векторном пространстве;
- выводы - $K = \{k_1, \dots, k_n\}$ индексируются в векторном пространстве.

Соответствующий тензор из трех осей в соответствии с рисунком 2, определяется как:

$$T[i, j, k] = \begin{cases} 1, & \text{если } r_i : f_j \rightarrow f_k \\ 0, & \text{иначе} \end{cases} \quad (10)$$

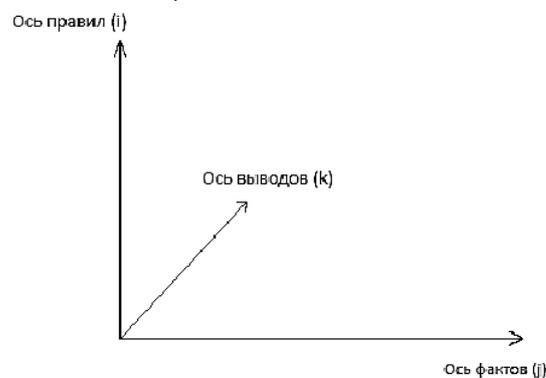


Рисунок 2 – Схема направление векторов тензора из примера

Процесс логического вывода формализуется через поиск применимых правил, что эквивалентно проверке условия $P_i \subseteq D$, что является классическим приемом в компьютерных науках [33-36]. Скалярное произведение векторов дает матрицу активаций $M \in \{0,1\}^{k \times n}$ [37]:

$$M[i, j] = d_j \cdot T[i, j, \ell] \quad (11)$$

Условие активации правила r_i определяется как:

$$\text{акв}(r_i) = \begin{cases} 1, & \text{если } \sum_{j=1}^n M[i, j] = \|P_i\| \\ 0, & \text{иначе} \end{cases} \quad (12)$$

где $\|P_i\|$ -количество фактов в антецеденте P_i .

Перспективным направлением развития является использование тензорной алгебры [38] для расширения возможностей предложенного подхода. К проблеме сложности визуализации тензорных представлений подошли авторы работы [39] и предлагают специализированные методы, такие как тензорные глифы.

Важной теоретической проблемой остается проверка достижимости вывода в тензорной парадигме. Как показано в [40], для продукционных систем, моделируемых матричными преобразованиями, проблема достижимости может оказаться алгоритмически неразрешимой уже для систем с матрицами 4×4 .

К преимуществам тензорного подхода относятся:

- вычислительная эффективность за счет сведения к операциям линейной алгебры;
- возможность параллельной обработки на GPU/TPU;
- единое представление сложных систем правил;
- поддержка вероятностного вывода через расширение моделью.

Основные ограничения метода:

- экспоненциальный рост требований к памяти;
- сложность интерпретации без дополнительных методов визуализации;
- необходимость аппроксимации при большом количестве правил.

Тензорное кодирование преобразует символьные правила в операции линейной алгебры, создавая математически строгую модель, эффективно реализуемую на современном аппаратном обеспечении. Данный подход обеспечивает масштабируемость за счет параллельных вычислений и методов тензорного разложения.

Выводы

В работе выполнен сравнительный анализ методов интеграции продукционных систем со структурами И-ИЛИ-деревьев и тензорными представлениями. Анализ проводился с позиций оценки формальных основ, применимости и ограничений указанных подходов применительно к задачам обработки онтологий и работы с большими данными в образовательных системах.

Классические продукционные системы, обладая высокой интерпретируемостью и относительной простотой реализации, демонстрируют недостаточную производительность при обработке крупномасштабных массивов данных. Интеграция с И-ИЛИ-деревьями способствует структурированию сложных логических условий и обеспечивает прозрачность механизма логического вывода, включая поддержку инкрементальных вычислений. Однако данный подход сопряжен с проблемами комбинаторного роста вариантов и повышенной сложности верификации при циклических зависимостях.

Альтернативный подход на основе тензорных представлений позволяет перейти от символьных правил к операциям линейной алгебры, что обеспечивает существенное повышение производительности вычислений за счет параллелизации и низкоуровневой оптимизации. Тензорные методы обеспечивают компактное представление онтологических структур и обладают высокой масштабируемостью, однако требуют значительных вычислительных ресурсов и затрудняют интерпретацию результатов.

Сравнительный анализ выявил комплементарный характер исследуемых подходов: тензорные представления демонстрируют более высокую производительность, в то время, как И-ИЛИ-деревья обладают преимуществом в интерпретируемости и гибкости настройки.

В качестве перспективного направления выделяется разработка адаптивного мета-алгоритма, способного динамически выбирать между использованием древовидных структур, тензорных вычислений или их гибридной комбинации. Критериями выбора должны служить объем обрабатываемых данных, требования к интерпретируемости модели и доступные вычислительные ресурсы.

Дальнейшие исследования целесообразно сосредоточить на исследовании алгебраических свойств последовательностей преобразований в продукционных системах в рамках теории полугрупп и создании методов автоматического преобразования между различными формами представления правил.

Таким образом, синтез продукционного подхода с анализируемыми математическими формализмами позволяет создавать системы, сочетающие интерпретируемость логического вывода с вычислительной эффективностью, что особенно актуально для образовательной сферы с ее требованием обслуживания значительных потоков пользователей при сохранении прозрачности принимаемых решений.

Список литературы

1. Григорьев, А. В. Алгоритм выполнения теоретико-множественных операций над грамматиками в среде специализированной оболочки для создания интеллектуальных САПР // Научные труды Донецкого государственного технического университета. Серия: Проблемы моделирования и автоматизации проектирования динамических систем (МАП-2002). 2002. № 52. С. 83–93.
2. Григорьев, А. В. Упорядочивание обликов в семиотической модели САПР // Искусственный интеллект. 2005. № 4. С. 465–477.
3. Григорьев, А. В. Специфика выполнения теоретико-множественных операций над контекстно-свободными грамматиками в условиях различных форм дополнительных семантических правил в семиотической модели интеллектуальных САПР // Научные труды Донецкого национального технического университета. Серия «Проблемы моделирования и автоматизации проектирования динамических систем» (МАП — 2006). 2006. Вып. 5 (116). С. 91–104.
4. Khan M. F. Syntactic Transformation in Large Language Models (LLMs): A Transformational Generative Grammar (TGG) Perspective // Dicon Journal of Languages. 2025. Vol. 1, Issue 1. P. 24–43. DOI: <https://doi.org/10.64169/djl.23>.
5. Portelance E., Jasbi M. On the compatibility of generative AI and generative linguistics // Nature Computational Science. 2025. Vol. 5, No. 9. DOI: 10.1038/s43588-025-00861-2.
6. Pullum G. K. Creation Myths of Generative Grammar and the Mathematics of Syntactic Structures // The Mathematics of Language: collection of works / ed. by C. Ebert, G. Jager, J. Michaelis. – Berlin; Heidelberg: Springer, 2010. P. 238–254. (Lecture Notes in Computer Science; vol. 6149). DOI: 10.1007/978-3-642-14322-9_18. ISBN 978-3-642-14321-2.
7. Атакищев, А. О. Продукционные сети как модели генерации и анализа вариантов принимаемых управленческих решений / А. О. Атакищев, С. А. Макаренков, И. М. Круглов, Д. О. Соколов // Т-Comm. 2008. № 4. С. 23–24.
8. Савушкин С. А., Бородин В. А., Цыганов В. В. Информационно-логическая компонента сценарного моделирования // Информационные технологии в науке, образовании и управлении. 2021. № 1. С. 70–76. DOI: 10.47501/ITNOU.2021.1.70-76.
9. Гурин, А. Г. Обзор методов представления онтологий с физической семантикой / А. Г. Гурин, А. В. Григорьев // Современные информационные технологии в образовании и научных исследованиях: материалы VIII Всерос. науч.-техн. конф. (СИТОНИ-2023). Донецк, 2023. С. 87–92.
10. Анцыферов, С. С. Методология развития интеллектуальных систем / С. С. Анцыферов, А. С. Сигов, К. Н. Фазилова // Проблемы искусственного интеллекта. 2022. № 2 (25). С. 42–47.
11. Харламов, А. А. Перспективы развития семантических представлений, основывающиеся на тенденциях нейроинформатики // Проблемы искусственного интеллекта. 2020. № 3 (18). – С. 3–17.
12. Андриевская, Н. К. Обобщенная модифицированная модель представления текстовых информационных ресурсов // Информатика и кибернетика. 2020. № 4 (22). С. 21–30.
13. Николаев, А. А. Международный опыт и перспективы использования искусственного интеллекта в образовании / А. А. Николаев, М. Ю. Кузнецов, В. А. Николаев // Управление образованием: теория и практика. 2024. Т. 14, № 5-1. DOI 10.25726/e8567-8724-8003-k. – EDN ZNISEA.

14. Чудина, Е. Ю. Решение задачи оптимизации уровня учебной успеваемости и самоактуализации обучающихся на основании двухфакторной регрессии с применением метода идеальной точки // Проблемы искусственного интеллекта. 2023. № 1 (28). С. 60–68.
15. Андриевская, Н. К. Онтологический подход в системах обработки данных научных и научно-образовательных организаций // Проблемы искусственного интеллекта. 2020. № 1 (16). С. 23–36.
16. Гаркуша, Д. А. Функциональные особенности реализованных онтологических платформ // Проблемы искусственного интеллекта. 2023. № 4 (31). С. 4–11.
17. Дорохина Г.В. Формализованные онтологии и задачи построения компьютерной информационной технологии цифрового сбора, обработки и анализа данных. Искусственный интеллект: теоретические аспекты, практическое применение: материалы Донецкого международного научного круглого стола. Донецк: ГУ ИПИИ, 2020. С. 229–234.
18. Waterman, D. A. Pattern-Directed Inference Systems / D. A. Waterman, F. Hayes-Roth. Academic Press, 1978. 680 p. ISBN 978-0-12-737550-0.
19. Shortliffe, E. H. Computer-Based Medical Consultations: MYCIN. Elsevier, 1976. 264 p. ISBN 978-0-444-00183-0.
20. Brachman, R. J. Knowledge Representation and Reasoning / R. J. Brachman, H. J. Levesque. Morgan Kaufmann, 2004. 381 p. ISBN 978-1558609327.
21. Russell, S. Artificial Intelligence: A Modern Approach / S. Russell, P. Norvig. 4th ed. Pearson, 2020. 1136 p. ISBN 978-0134610993.
22. Люгер, Дж. Ф. Искусственный интеллект: стратегии и методы решения сложных проблем / Дж. Ф. Люгер; пер. с англ. М.: Вильямс, 2003. 864 с.
23. Baier, C. Principles of Model Checking / C. Baier, J.-P. Katoen. MIT Press, 2008. 975 p. ISBN 978-0262026499.
24. Model checking [Электронный ресурс] // Wikipedia. URL: https://en.wikipedia.org/wiki/Model_checking (дата обращения: 25.09.2024).
25. Метод генерации тестовых заданий на основе деревьев И/ИЛИ и его программная реализация [Электронный ресурс] // Habr. URL: <https://habr.com/ru/articles/198068/> (дата обращения: 25.09.2024).
26. Nilsson, N. J. Principles of Artificial Intelligence. Springer, 1982. 476 p. ISBN 978-3-540-11340-1.
27. Полетаев А.Ю., Парамонов И.В., Бойчук Е.И. Алгоритм построения дерева синтаксических единиц русскоязычного предложения по дереву синтаксических связей // Информатика и автоматизация. 2023. Т. 22. № 6. С. 1323-1353. DOI: 10.15622/ia.22.6.3.
28. Румянцев, В. В. Моделирование физических объектов / В. В. Румянцев, С. А. Федоров, А. Е. Рыбалка // Проблемы искусственного интеллекта. 2023. № 1 (28). С. 50–59.
29. Андриевская, Н. К. Гибридная интеллектуальная мера оценки семантической близости / Н. К. Андриевская // Проблемы искусственного интеллекта. 2021. № 1 (20). С. 4–17.
30. Newell, A. Human Problem Solving. The state of the theory in 1970 / A. Newell, H. A. Simon // American Psychologist. 1971. Vol. 26, no. 2. P. 145–159.
31. Laird, J. E. SOAR: An Architecture for General Intelligence / J. E. Laird, A. Newell, P. S. Rosenbloom // Artificial Intelligence. 1987. Vol. 33, no. 1. P. 1–64.
32. Rocktäschel, T. End-to-End Differentiable Proving / T. Rocktäschel, S. Riedel // Advances in Neural Information Processing Systems 30 (NIPS 2017). 2017. P. 3788–3800.
33. Salton, G. Introduction to Modern Information Retrieval / G. Salton, M. J. McGill. McGraw-Hill, 1983. 448 p. – ISBN 0070544840.
34. Strang, G. Linear Algebra and Its Applications. 4th ed. Thomson Learning, 2006. 487 p. ISBN 978-0030105678.
35. Maida, A. S. Intensional Concepts in Propositional Semantic Networks / A. S. Maida, S. C. Shapiro // Cognitive Science. – 1982. – Vol. 6, no. 4. – P. 291–330.
36. Agrawal, R. Fast algorithms for mining association rules in large databases / R. Agrawal, R. Srikant // Proceedings of the 20th International Conference on Very Large Data Bases (VLDB). 1994. P. 487–499.
37. Kolda, T.G. Tensor Decompositions and Applications / T.G. Kolda, B.W. Bader // SIAM Review. 2009. – Vol. 51, no. 3. P. 455–500.
38. Дзедобоев, Д. И. Тензорная алгебра / Д. И. Дзедобоев // Проблемы искусственного интеллекта. 2025. № 3 (38). С. 11–18.
39. Воробьева Г. Р., Воробьев А. В., Орлов Г. О. Концепция обработки, анализа и визуализации геофизических данных на основе элементов тензорного исчисления // Информатика и автоматизация. 2024. Т. 23. № 2 С. 572–604. DOI: 10.15622/ia.23.2.10.
40. Рыбалов А. Н. О генерической сложности проблемы вхождения для полугрупп целочисленных матриц // Прикладная дискретная математика. 2022. № 55. С. 95–101. DOI: 10.17223/20710410/55/7.

References

1. Grigoriev, A. V. (2002). An Algorithm for Performing Set-Theoretic Operations on Grammars in a Specialized Shell for Creating Intelligent CAD Systems. *Scientific Works of Donetsk State Technical University. Series: Problems of Modeling and Automation of Dynamic Systems Design (MAP-2002)* (52), 83–93.
2. Grigoriev, A. V. (2005). Ordering of Appearances in the Semiotic Model of CAD. *Artificial Intelligence*, (4), 465–477.
3. Grigoriev, A. V. (2006). Specifics of Performing Set-Theoretic Operations on Context-Free Grammars Under Different Forms of Additional Semantic Rules in the Semiotic Model of Intelligent CAD. *Scientific Works of Donetsk National Technical University. Series "Problems of Modeling and Automation of Dynamic Systems Design" (MAP-2006)*, 5(116), 91–104.
4. Khan, M. F. (2025). Syntactic Transformation in Large Language Models (LLMs): A Transformational Generative Grammar (TGG) Perspective. *Dibon Journal of Languages*, 1(1), 24–43. <https://doi.org/10.64169/djl.23>
5. Portelance, E., & Jasbi, M. (2025). On the compatibility of generative AI and generative linguistics. *Nature Computational Science*, 5(9). <https://doi.org/10.1038/s43588-025-00861-2>
6. Pullum, G. K. (2010). Creation Myths of Generative Grammar and the Mathematics of Syntactic Structures. In C. Ebert, G. Jager, & J. Michaelis (Eds.), *The Mathematics of Language* (pp. 238–254). Springer. (Lecture Notes in Computer Science, Vol. 6149). https://doi.org/10.1007/978-3-642-14322-9_18
7. Atakishchev, A. O., Makarenkov, S. A., Kruglov, I. M., & Sokolov, D. O. (2008). Production Networks as Models for Generating and Analyzing Alternatives for Management Decisions. *T-Comm*, (4), 23–24.
8. Savushkin, S. A., Borodin, V. A., & Tsyganov, V. V. (2021). Information-Logical Component of Scenario Modeling. *Information Technologies in Science, Education and Management*, (1), 70–76. <https://doi.org/10.47501/ITNOU.2021.1.70-76>
9. Gurin, A. G., & Grigoriev, A. V. (2023). A Review of Methods for Representing Ontologies with Physical Semantics. In *Modern Information Technologies in Education and Scientific Research: Materials of the VIII All-Russian Scientific and Technical Conference (SITONI-2023)* (pp. 87–92). Donetsk.
10. Antsiferov, S. S., Sigov, A. S., & Fazilova, K. N. (2022). Methodology for the Development of Intelligent Systems. *Problems of Artificial Intelligence*, 2(25), 42–47.
11. Kharlamov, A. A. (2020). Prospects for the Development of Semantic Representations Based on Neuroinformatics Trends. *Problems of Artificial Intelligence*, 3(18), 3–17.
12. Andrievskaya, N. K. (2020). Generalized Modified Model for Representing Textual Information Resources. *Informatics and Cybernetics*, 4(22), 21–30.
13. Nikolaev, A. A., Kuznetsov, M. Y., & Nikolaev, V. A. (2024). International Experience and Prospects for the Use of Artificial Intelligence in Education. *Education Management: Theory and Practice*, 14(5-1). <https://doi.org/10.25726/e8567-8724-8003-k>
14. Chudina, E. Y. (2023). Solving the Problem of Optimizing the Level of Academic Achievement and Self-Actualization of Students Based on Two-Factor Regression Using the Ideal Point Method. *Problems of Artificial Intelligence*, 1(28), 60–68.
15. Andrievskaya, N. K. (2020). Ontological Approach in Data Processing Systems of Scientific and Scientific-Educational Organizations. *Problems of Artificial Intelligence*, 1(16), 23–36.
16. Garkusha, D. A. (2023). Functional Features of Implemented Ontological Platforms. *Problems of Artificial Intelligence*, 4(31), 4–11.
17. Dorokhina, G. V. (2020). Formalized Ontologies and the Tasks of Building a Computer Information Technology for Digital Collection, Processing and Data Analysis. In *Artificial Intelligence: Theoretical Aspects, Practical Application: Materials of the Donetsk International Scientific Round Table* (pp. 229–234). State Institution IPII.
18. Waterman, D. A., & Hayes-Roth, F. (Eds.). (1978). *Pattern-Directed Inference Systems*. Academic Press.
19. Shortliffe, E. H. (1976). *Computer-Based Medical Consultations: MYCIN*. Elsevier.
20. Brachman, R. J., & Levesque, H. J. (2004). *Knowledge Representation and Reasoning*. Morgan Kaufmann.
21. Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
22. Luger, G. F. (2003). *Artificial Intelligence: Strategies and Methods for Solving Complex Problems* (Trans. from English). Williams.
23. Baier, C., & Katoen, J.-P. (2008). *Principles of Model Checking*. MIT Press.
24. Model checking. (2024, September 25). In Wikipedia. Retrieved September 25, 2024, from https://en.wikipedia.org/wiki/Model_checking

25. A Method for Generating Test Tasks Based on AND/OR Trees and Its Software Implementation. (n.d.). Habr. Retrieved September 25, 2024, from <https://habr.com/ru/articles/198068/>
26. Nilsson, N. J. (1982). *Principles of Artificial Intelligence*. Springer.
27. Poletaev, A. Y., Paramonov, I. V., & Boichuk, E. I. (2023). An Algorithm for Constructing a Tree of Syntactic Units of a Russian-Language Sentence from a Dependency Tree. *Informatics and Automation*, 22(6), 1323–1353. <https://doi.org/10.15622/ia.22.6.3>
28. Romyantsev, V. V., Fedorov, S. A., & Rybalka, A. E. (2023). Modeling of Physical Objects. *Problems of Artificial Intelligence*, 1(28), 50–59.
29. Andrievskaya, N. K. (2021). A Hybrid Intelligent Measure for Assessing Semantic Similarity. *Problems of Artificial Intelligence*, 1(20), 4–17.
30. Newell, A., & Simon, H. A. (1971). Human Problem Solving. The state of the theory in 1970. *American Psychologist*, 26(2), 145–159.
31. Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An Architecture for General Intelligence. *Artificial Intelligence*, 33(1), 1–64.
32. Rocktäschel, T., & Riedel, S. (2017). End-to-End Differentiable Proving. *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 3788–3800.
33. Salton, G., & McGill, M. J. (1983). *Introduction to Modern Information Retrieval*. McGraw-Hill.
34. Strang, G. (2006). *Linear Algebra and Its Applications* (4th ed.). Thomson Learning.
35. Maida, A. S., & Shapiro, S. C. (1982). Intensional Concepts in Propositional Semantic Networks. *Cognitive Science*, 6(4), 291–330.
36. Agrawal, R., & Srikant, R. (1994). Fast Algorithms for Mining Association Rules in Large Databases. *Proceedings of the 20th International Conference on Very Large Data Bases (VLDB)*, 487–499.
37. Kolda, T. G., & Bader, B. W. (2009). Tensor Decompositions and Applications. *SIAM Review*, 51(3), 455–500.
38. Dzeboev, D. I. (2025). Tensorial Algebra. *Problems of Artificial Intelligence*, 3(38), 11–18.
39. Vorobeva, G. R., Vorobev, A. V., & Orlov, G. O. (2024). A Concept for Processing, Analysis and Visualization of Geophysical Data Based on Elements of Tensor Calculus. *Informatics and Automation*, 23(2), 572–604. <https://doi.org/10.15622/ia.23.2.10>
40. Rybalov, A. N. (2022). On the Generic Complexity of the Membership Problem for Semigroups of Integer Matrices. *Applied Discrete Mathematics*, (55), 95–101. <https://doi.org/10.17223/20710410/55/7>

RESUME

Gurin A.G., Grigoriev A.V.

Comparative analysis of methods for enhancing the efficiency of production systems for big data processing

A comparative analysis of two directions for modernizing production systems has been conducted: integration with AND-OR trees and the application of tensor representations in the context of educational systems. The study evaluates the formal foundations, applicability, and limitations of these approaches for processing ontologies and working with big data.

Integration with AND-OR trees provides structural representation of complex logical conditions and a high degree of model interpretability, supporting efficient incremental inference. However, this approach is susceptible to combinatorial explosion and challenges with cyclic dependencies. The tensor method demonstrates enhanced performance through parallelism and compact data representation, contributing to system scalability. Its disadvantages include high memory requirements and complexity in interpreting results.

The analysis confirms the complementary nature of the approaches: tensor representations are effective for processing large data volumes, while AND-OR trees excel in interpretability and flexibility.

A promising direction is the development of an adaptive meta-algorithm capable of dynamically selecting or combining approaches based on current context (data volume, explainability requirements, available resources). An additional research direction involves integrating fuzzy logic and probabilistic inference methods for processing large rule sets.

РЕЗЮМЕ

Гурин А.Г., Григорьев А.В.

Сравнительный анализ методов повышения эффективности продукционных систем для работы с большими данными

Проведен сравнительный анализ двух направлений модернизации продукционных систем: интеграции с И-ИЛИ-деревьями и применения тензорных представлений в контексте образовательных систем. Исследование оценивает формальные основы, применимость и ограничения данных подходов для задач обработки онтологий и работы с большими данными.

Интеграция с И-ИЛИ-деревьями обеспечивает структурное представление сложных логических условий и высокий уровень интерпретируемости модели, поддерживая эффективный инкрементальный вывод. Однако данный подход подвержен комбинаторному взрыву и сложностям при циклических зависимостях. Тензорный метод демонстрирует повышенную производительность за счет параллелизма и компактного представления данных, что способствует масштабируемости системы. К недостаткам относятся высокие требования к памяти и сложность интерпретации результатов.

Анализ подтверждает комплементарный характер подходов: тензорные представления эффективны для обработки больших объемов данных, тогда как И-ИЛИ-деревья превосходят в интерпретируемости и гибкости.

Перспективным направлением является разработка адаптивного мета-алгоритма, способного динамически выбирать или комбинировать подходы на основе текущего контекста (объема данных, требований к объяснимости, доступных ресурсов). Дополнительное направление исследований включает интеграцию методов нечеткой логики и вероятностного вывода для обработки крупных наборов правил.

Гурин А.Г. – соискатель, аспирант, ФГБОУ ВО ДонНТУ, кафедра «Программная инженерия им. Л.П. Фельдмана», 283001, Донецк, ул. Артема, 58, тел +7(949) 364 1724, gurin.sasha.1996@gmail.com.

Область научных интересов: моделирование интеллектуальных систем, онтологии, тензорные исчисления, полугруппы, нейронные сети.

Григорьев А.В. – к.т.н., доцент, ФГБОУ ВО ДонНТУ, кафедра «Программная инженерия им. Л.П. Фельдмана», 283001, Донецк, ул. Артема, 58, тел +7(949) 334 9185, grigorievalvl@gmail.com.

Область научных интересов: интеллектуальные системы автоматизации проектирования, формальные грамматики, онтологии, нейронные сети.

Статья поступила в редакцию 09.09.2025.

УДК 512.554.1

Д. И. Дзебоев

Национальный исследовательский университет «Высшая школа экономики»
109028, Москва, Покровский бульвар, 11

ПОИСК ЕДИНИЦ В ТЕНЗОРИАЛЬНОЙ АЛГЕБРЕ

D. I. Dzeboev

National Research University "Higher School of Economics"
109028, 11 Pokrovsky Boulevard, Moscow

SEARCH OF UNITS IN TENSORIAL ALGEBRA

В работе рассматривается задача аналитического поиска единичных элементов в тензориальной алгебре — обобщённой алгебраической системе билинейного умножения векторов, заданной тензором алгебры. Показано, что действие любого тензориального числа в алгебре может быть представлено в виде линейного оператора, а поиск единицы сводится к решению системы линейных уравнений, при которой свёртка тензора алгебры с неизвестным вектором образует единичную матрицу. Получены общие формулы для левой и правой единицы в терминах псевдообратных матриц Мура–Пенроуза и показано, что при существовании хотя бы одной правой и одной левой единицы они совпадают, образуя единственную двустороннюю (коммутативную) единицу, делающую алгебру унитарной.

Ключевые слова: тензориальная алгебра, тензор алгебры, единица алгебры, правая и левая единица, псевдообратная матрица, линейный оператор, унитарная структура

The paper considers the analytical problem of finding unit elements in tensorial algebra — a generalized algebraic system of bilinear vector multiplication defined by an algebra tensor. It is shown that the action of any tensorial number in the algebra can be represented as a linear operator, and that the search for the unit element reduces to solving a system of linear equations where the contraction of the algebra tensor with an unknown vector yields the identity matrix. General formulas for the left and right units are derived in terms of Moore–Penrose pseudoinverses. It is proven that if at least one right and one left unit exist, they coincide, forming a unique two-sided (commutative) unit that renders the algebra unital.

Key words: tensorial algebra, algebra tensor, algebraic unit, right and left unit, Moore–Penrose pseudoinverse, linear operator, unital structure

Введение

Тензориальная алгебра является одним из направлений современного развития обобщённых алгебраических структур, в которых операция умножения задаётся не аналитически, а конструктивно — через тензор алгебры, определяющий результат произведения любых двух базисных элементов [1]. Такой подход обеспечивает возможность построения алгебр произвольной размерности, не ограничивающихся классическими гиперкомплексными системами (вещественными, комплексными числами, кватернионами и т.д.), и допускает как аналитическое, так и обучаемое определение операции умножения. Это делает тензориальную алгебру универсальным инструментом для моделирования, символической математики и нейросетевых архитектур, основанных на алгебраических преобразованиях.

Ключевой особенностью тензориальной алгебры является сохранение размерности пространства при выполнении операции умножения, что обеспечивает её замкнутость и позволяет рассматривать пространство алгебры как кольцо. Для дальнейшего анализа таких систем необходимо определить внутренние свойства — наличие единичного элемента, возможность построения обратных и, следовательно, реализацию операции деления. В отличие от традиционных числовых алгебр, где существование единицы очевидно, в тензориальной алгебре наличие и единственность единичного элемента зависят от структуры тензора умножения и потому требуют строгого аналитического описания.

В работе [1] было введено основное определение тензориальной алгебры и показано, что её операция может быть реализована в виде программного класса, выполняющего эйнштейнову свёртку по тензору структуры. В патенте на программу для ЭВМ № 2025667782 [2] эта алгебра была зарегистрирована как универсальный инструмент для вычисления обобщённого произведения тензоров произвольного ранга.

Дальнейшее развитие получила идея представления действия тензориального числа в алгебре как линейного оператора [3]. Такое рассмотрение позволило перейти к линейной форме, где каждое тензориальное число сопоставляется матрице действия, а исследование его свойств сводится к задачам линейной алгебры. На основе этого подхода в патенте № 2025681990 [4] была предложена методика поиска правых и левых единиц в тензориальной алгебре, сводящая задачу к решению систем линейных уравнений относительно вектора, при котором матрица действия принимает вид единичной.

В первой части данной работы формулируется теорема о действии тензориального числа как линейного оператора, во второй — теорема о поиске левой и правой единицы и их аналитическом выражении через псевдообратные матрицы Мура–Пенроуза, а в заключении доказывается теорема о единственности и коммутативности единицы при существовании хотя бы одной пары правой и левой нейтрали.

1 Действие тензориального числа как линейного оператора

Для того чтобы определить, что именно будет являться единичным элементом в тензориальной алгебре, необходимо установить связь между тензориальным числом и его действием на остальные элементы алгебры. Это действие удобно рассматривать как линейный оператор в базисном пространстве. Такой подход позволяет перейти от тензориальной формы записи к привычным операциям линейной алгебры — матричному умножению и исследованию свойств операторов (ранг, определитель, спектр).

Пусть алгебра \mathcal{A} задана тензором структуры третьего ранга $T_{ijk} \in \mathbb{R}^{n \times n \times n}$. Для любых $a, b \in \mathbb{R}^n$ операция умножения определяется формулой:

$$c = a \boxtimes b, \quad c_k = \sum_{i,j} a_i b_j T_{ijk}$$

Теорема 1.

Любое тензориальное число $A \in \mathbb{R}^n$ в алгебре $\mathcal{A}(T)$ может быть представлено в виде матрицы линейного оператора, действующего на произвольный элемент алгебры посредством матричного умножения.

При этом матрицы левого и правого действия задаются свёрткой вектора A с тензором структуры:

$$M^L(A)_{jk} = \sum_i A_i T_{ijk}$$

$$M^R(A)_{ik} = \sum_j A_j T_{ijk}$$

Тогда для любого $x \in \mathbb{R}^n$ выполняется:

$$A \boxtimes x = M^L(A) x$$

$$x \boxtimes A = M^R(A) x$$

Доказательство.

Из определения операции тензориального умножения следует:

$$(A \boxtimes x)_k = \sum_{i,j} A_i x_j T_{ijk}$$

Зафиксируем вектор A . Тогда каждая компонента A_i определяет линейную комбинацию матриц $T_i = (T_{ijk})_{jk}$, и, следовательно:

$$M^L(A) = \sum_i A_i T_i, \quad \text{где } (T_i)_{jk} = T_{ijk}$$

Отсюда:

$$(A \boxtimes x)_k = \sum_j M^L(A)_{jk} x_j = (M^L(A)x)_k$$

Аналогично, если фиксировать индекс j , то:

$$M^R(A) = \sum_j A_j T^j, \quad \text{где } (T^j)_{ik} = T_{ijk}$$

и, следовательно:

$$(x \boxtimes A)_k = (M^R(A)x)_k$$

Таким образом, каждое тензориальное число может рассматриваться как линейный оператор, однозначно определяющий своё действие на остальные элементы алгебры.

2 Поиск единиц в тензориальной алгебре

Единичным элементом тензориальной алгебры называется такой вектор ($I \in \mathbb{R}^n$), который сохраняет любой элемент алгебры при умножении с соответствующей стороны. Для правой единицы это условие имеет вид:

$$\forall A \in \mathbb{R}^n: \quad \text{einsum}("i, j, ijk \rightarrow k", A, I, T) = A$$

для левой —

$$\forall A \in \mathbb{R}^n: \quad \text{einsum}("i, j, ijk \rightarrow k", I, A, T) = A,$$

где T — тензор алгебры.

Теорема 2.

Вектор I является единичным элементом тензориальной алгебры тогда и только тогда, когда соответствующая ему матрица действия тождественно равна единичной матрице:

$$M_R(I) = \text{einsum}("j, ijk \rightarrow ik", I, T) = E_n$$

$$M_L(I) = \text{einsum}("i, ijk \rightarrow jk", I, T) = E_n$$

При этом $M_R(I)$ и $M_L(I)$ являются матрицами правого и левого действия тензориального числа I на пространство алгебры.

Доказательство.

Согласно Теореме 1, действие любого тензориального числа A на элемент алгебры можно рассматривать как линейный оператор, определяемый свёрткой его координат с тензором алгебры:

$$M_R(A) = \text{einsum}("j, ijk \rightarrow ik", A, T)$$

$$M_L(A) = \text{einsum}("i, ijk \rightarrow jk", A, T)$$

Для того чтобы I было правой единицей, необходимо выполнение условия

$$A = \text{einsum}("i, j, ijk \rightarrow k", A, I, T) = \text{einsum}("i, ik \rightarrow k", A, M_R(I)) = A,$$

что возможно тогда и только тогда, когда $(M_R(I) = E_n$.

Аналогично, для левой единицы:

$$A = \text{einsum}("i, j, ijk \rightarrow k", I, A, T) = \text{einsum}("j, jk \rightarrow k", A, M_L(I)) = A$$

что выполняется при $M_L(I) = E_n$.

Следовательно, поиск единицы сводится к решению системы линейных уравнений относительно вектора I , при котором свёртка тензора алгебры с I по первому (для левой) или второму (для правой) индексу порождает единичную матрицу:

$$\text{einsum}("i, ijk \rightarrow jk", I, T) = E_n \text{ (левая единица)}$$

$$\text{einsum}("j, ijk \rightarrow ik", I, T) = E_n \text{ (правая единица)}$$

Совместность этих систем определяет существование единиц, а их совпадение — наличие единственной коммутативной единицы, делающей алгебру унитарной.

Теорема 3. Формула единицы

Формулы единицы тензориальной алгебры равны:

$$I_L = \text{vec}_{(j,k)}[T_{ijk}]^+ \text{vec}(E_n) + (I_n - \text{vec}_{(j,k)}[T_{ijk}]^+ \text{vec}_{(j,k)}[T_{ijk}]) w_L, \quad w_L \in \mathbb{R}^n$$

$$I_R = \text{vec}_{(i,k)}[T_{ijk}]^+ \text{vec}(E_n) + (I_n - \text{vec}_{(i,k)}[T_{ijk}]^+ \text{vec}_{(i,k)}[T_{ijk}]) w_R, \quad w_R \in \mathbb{R}^n$$

Доказательство.

Пусть тензориальная алгебра задана тензором алгебры T_{ijk} , и пусть

$$A_L = \text{vec}_{(j,k)}[T_{ijk}] \in \mathbb{R}^{n^2 \times n}$$

$$A_R = \text{vec}_{(i,k)}[T_{ijk}] \in \mathbb{R}^{n^2 \times n}$$

$$b = \text{vec}(E_n) \in \mathbb{R}^{n^2}$$

Тогда множество всех левых (соответственно правых) единиц есть множество всех решений системы

$$A_L I_L = b \quad (\text{соответственно } A_R I_R = b)$$

Условия существования (Кронекера–Капелли)

Система $AI = b$ совместна тогда и только тогда, когда

$$\text{rank}(A) = \text{rank}([A \mid b]) \Leftrightarrow AA^+b = b$$

где $A \in \{A_L, A_R\}$ и A^+ — псевдообратная Мура–Пенроуза.

Общее решение при существовании.

Если система $AI = b$ совместна, то все решения имеют вид [5]

$$I = A^+b + (I_n - A^+A)w, \quad w \in \mathbb{R}^n$$

где A^+b — фиксированное частное (минимальной нормы) решение, а свободное слагаемое $(I_n - A^+A)w$ пробегает ядро A .

В частности,

$$\begin{aligned} I_L &= A_L^+b + (I_n - A_L^+A_L)w_L \\ I_R &= A_R^+b + (I_n - A_R^+A_R)w_R \\ w_L, w_R &\in \mathbb{R}^n \end{aligned}$$

Поэтому, возвращаясь к исходным единицам:

$$I_L = \text{vec}_{(j,k)}[T_{ijk}]^+ \text{vec}(E_n) + (I_n - \text{vec}_{(j,k)}[T_{ijk}]^+ \text{vec}_{(j,k)}[T_{ijk}])w_L, \quad w_L \in \mathbb{R}^n$$

$$I_R = \text{vec}_{(i,k)}[T_{ijk}]^+ \text{vec}(E_n) + (I_n - \text{vec}_{(i,k)}[T_{ijk}]^+ \text{vec}_{(i,k)}[T_{ijk}])w_R, \quad w_R \in \mathbb{R}^n$$

Размерность и единственность.

Размерность аффинного множества решений равна $\dim \ker(A) = n - \text{rank}(A)$

Соответственно:

$$\text{rank}(A) = n \Leftrightarrow \ker(A) = \{0\} \Leftrightarrow I = A^+b \text{ единственно.}$$

Применяя к A_L и A_R , получаем критерии единственности левой/правой единицы.

Теорема 4. О единственности и коммутативности единицы при наличии хотя бы одной правой и хотя бы одной левой.

Пусть тензориальная алгебра с билинейным умножением \boxtimes задана тензором алгебры T . Предположим, что существует хотя бы одна правая единица I_R и хотя бы одна левая единица I_L , то есть

$$\forall A: A \boxtimes I_R = A$$

$$\forall A: I_L \boxtimes A = A$$

Тогда $I_R = I_L$. Обозначая этот элемент через I , получаем единственную двустороннюю единицу (унитарность алгебры), которая автоматически коммутирует с любым элементом:

$$\forall A: I \boxtimes A = A = A \boxtimes I$$

Доказательство.

Подставим $A = I_R$ в равенство для левой единицы и $A = I_L$ в равенство для правой:

$$I_L \boxtimes I_R = I_R$$

$$I_L \boxtimes I_L = I_L$$

Отсюда немедленно следует $I_L = I_R$. Обозначим $I = I_L = I_R$. Тогда для любого A

$$I \boxtimes A = A \quad \text{и} \quad A \boxtimes I = A,$$

то есть I — двусторонняя единица и, в частности, «коммутативная» (находится в центре алгебры в том смысле, что $I \boxtimes A = A \boxtimes I$ для всех A)

Единственность: если I'_L — ещё одна левая единица или I'_R — ещё одна правая, то

$$\begin{aligned} I'_L &= I'_L \otimes I = I \\ I &= I \otimes I'_R = I'_R \end{aligned}$$

Следовательно двусторонняя единица единственна.

3 Программная реализация

Алгоритмы, описывающие действие тензориального числа как линейного оператора и поиск правых и левых единиц, реализованы в программном модуле TensorialAlgebra, написанном на языке Python с использованием библиотеки PyTorch. Такой выбор обусловлен возможностью работы с тензорами произвольной размерности, наличием встроенных средств линейной алгебры и функцией автоматического дифференцирования, что делает модуль применимым как в аналитических задачах, так и в обучаемых нейросетевых моделях тензориальной алгебры.

Класс TensorialAlgebra представляет собой параметризованный объект, содержащий тензор структуры, который определяет правила умножения элементов алгебры.

Метод forward(x, y) реализует основную операцию тензориального умножения по эйнштейновой свёртке, что соответствует операции $x \otimes y$ в аналитической форме.

Метод number_as_operator(A, side) вычисляет действие фиксированного тензориального числа A в виде линейного оператора. В зависимости от стороны действия формируются матрицы, которые отображают левое и правое умножение соответственно.

Для нахождения единиц реализован метод search_unit(side, w=None), который строит соответствующую систему линейных уравнений в матричной форме.

Корректность реализации подтверждена тестами на примерах комплексных чисел и кватернионов, где полученные единицы совпадают с известными аналитическими результатами ($I=(1,0)$ и $I=(1,0,0,0)$ соответственно). Кроме того, модуль допускает работу с произвольными случайными тензорами, что позволяет проводить вычислительные эксперименты и исследовать условия существования и единственности единиц в неассоциативных и обучаемых алгебрах.

4 Применение в практических системах

Единичный элемент в тензориальной алгебре играет ключевую роль при построении вычислительных и физических моделей, в которых требуется сохранение инвариантности по отношению к операции умножения. Наличие единственной коммутативной единицы обеспечивает корректность операций нормировки, обратимости и масштабирования в рамках алгебраической структуры, что делает тензориальную алгебру не только теоретически замкнутой системой, но и практическим инструментом для описания процессов в различных областях науки и техники.

Одним из направлений применения результатов данного исследования является моделирование колебательных систем и резонансных взаимодействий. В работе В. Н. Беловодского и С. Л. Букина [6] рассмотрено поведение материальной частицы в вибрационной транспортирующей машине нелинейного типа, где движение частицы определяется системой уравнений с билинейными характеристиками. Подобные задачи могут быть естественно описаны средствами тензориальной алгебры, где операция умножения отражает взаимодействие колебательных составляющих, а единичный

элемент задаёт состояние динамического равновесия системы. Таким образом, введение единственной тензориальной единицы позволяет корректно формализовать переходные режимы между суб- и супергармоническими колебаниями и описать устойчивые состояния как тензориальные нейтралы.

Другим направлением является применение тензориальных структур в задачах искусственного интеллекта и технологического моделирования. В статье В. С. Солода, В. М. Зуева и С. Б. Ивановой [7] показана возможность использования нейросетевых моделей для оптимизации металлургических процессов. Подобные системы могут быть дополнены тензориальными алгебрами с обучаемыми структурами, где единица играет роль эталонного состояния технологического режима — базовой точки стабилизации в параметрическом пространстве. Это обеспечивает возможность реализовать обучаемые модели, сохраняющие внутреннюю согласованность и обратимость при изменении производственных параметров.

Кроме того, физические интерпретации единицы тензориальной алгебры находят отклик в моделировании упорядоченных систем, описываемых топологическими и симметричными свойствами. Так, в работе С. А. Федорова, Ю. А. Безуса и А. Е. Рыбалки [8] исследованы поляритонные возбуждения в гексагональных наноструктурах. Единица тензориальной алгебры в таких моделях может рассматриваться как инвариантный элемент, задающий симметричное базисное состояние поля, относительно которого описываются все нарушения и возмущения. Это позволяет использовать тензориальные операторы для построения симметрично-инвариантных гамильтонианов и анализа энергетических спектров сложных кристаллических решёток.

Таким образом, понятие единицы в тензориальной алгебре выходит за рамки абстрактной алгебраической конструкции и становится фундаментальным элементом прикладных моделей — от механики и физики колебаний до нейросетевого моделирования и квантово-топологических систем. Наличие единственной коммутативной единицы гарантирует устойчивость описания, возможность обратных преобразований и корректную интерпретацию физических или вычислительных процессов в рамках единого алгебраического пространства.

Заключение

Выполненное исследование позволило установить строгие аналитические условия существования и единственности/бесконечности единичного элемента в тензориальной алгебре, а также их формулы. Показано, что действие тензориального числа в алгебре может рассматриваться как линейный оператор, а поиск правой и левой единицы сводится к решению систем линейных уравнений, в которых свёртка тензора алгебры с неизвестным вектором образует единичную матрицу.

Получены общие формулы для левой и правой единицы через псевдообратные матрицы Мура–Пенроуза, что позволяет аналитически описывать все случаи — отсутствие решения, единственность или существование бесконечного множества единиц. Доказано, что при наличии хотя бы одной правой и одной левой единицы они совпадают, образуя единственную коммутативную (двустороннюю) единицу, делающую алгебру унитарной.

Результаты имеют не только теоретическое значение, но и практическое применение: программная реализация в модуле `TensorialAlgebra` позволяет автоматически определять существование и структуру единицы для произвольного тензора алгебры, включая обучаемые модели. Это делает возможным интеграцию тензориальных алгебр в системы символьных и нейросетевых вычислений, а также применение в задачах физического моделирования и анализа симметрий.

Список литературы

1. Дзебоев Д. И. Тензорная алгебра. *Проблемы искусственного интеллекта*. 2025. № 3(38). Раздел: Искусственный интеллект и машинное обучение. С. 11–18.
2. Патент на программу для ЭВМ № 2025667782. Тензорная алгебра / Дзебоев Д. И. Роспатент, 2025.
3. Патент на программу для ЭВМ № 2025681930 – Действие тензорного числа в алгебре как линейного оператора / Дзебоев Д. И. Роспатент, 2025.
4. Патент на программу для ЭВМ № 2025681990. Поиск правых и левых единиц в тензорной алгебре / Дзебоев Д. И. Роспатент, 2025.
5. Ben-Israel A., Greville T. N. E. *Generalized Inverses: Theory and Applications*. New York: Springer, 2003.
6. Беловодский В. Н., Букин С. Л. О перемещениях материальной частицы на горизонтальной транспортирующей машине нелинейного типа, совершающей суб- и супергармонические колебания. *Проблемы искусственного интеллекта*. 2025. № 3(38). С. 15–24.
7. Солод В. С., Зуев В. М., Иванова С. Б. Оценка возможностей использования нейросетевых моделей для оптимизации технологии охлаждения арматурной стали. *Проблемы искусственного интеллекта*. 2025. № 3(38). С. 25–33.
8. Федоров С. А., Безус Ю. А., Рыбалка А. Е. Поляритонные возбуждения в топологически упорядоченном неидеальном гексагональном массиве микропор с примитивной решёткой. *Проблемы искусственного интеллекта*. 2025. № 3(38). С. 34–42.

References

1. Dzeboev D. I. Tensorial Algebra. // *Problems of Artificial Intelligence*. – 2025. – No. 3(38). – Section: Artificial Intelligence and Machine Learning.
2. Computer Program Patent No. 2025667782 – Tensorial Algebra / Dzeboev D. I. – Rospatent, 2025.
3. Computer Program Patent No. 2025681930 – Action of a Tensorial Number in Algebra as a Linear Operator / Dzeboev D. I. – Rospatent, 2025.
4. Computer Program Patent No. 2025681990 – Search of Right and Left Units in Tensorial Algebra / Dzeboev D. I. Rospatent, 2025.
5. Penrose R. *The Road to Reality: A Complete Guide to the Laws of the Universe*. London: Vintage, 2004. 1136 p.
6. Belovodsky V. N., Bukin S. L. On the Displacement of a Material Particle on a Horizontal Nonlinear-Type Transporting Machine Performing Sub- and Superharmonic Oscillations. // *Problems of Artificial Intelligence*. – 2025. – No. 3(38). – pp. 15–24.
7. Solod V. S., Zuev V. M., Ivanova S. B. Assessment of the Possibilities of Using Neural Network Models to Optimize the Technology of Reinforcing Steel Cooling. // *Problems of Artificial Intelligence*. – 2025. – No. 3(38). pp. 25–33.
8. Fedorov S. A., Bezus Yu. A., Rybalka A. E. Polaritonic Excitations in a Topologically Ordered Non-Ideal Hexagonal Array of Micropores with a Primitive Lattice. // *Problems of Artificial Intelligence*. – 2025. – No. 3(38). pp. 34–42.

RESUME

D. I. Dzeboev

Search of units in Tensorial Algebra

Background: Tensorial Algebra, which defines multiplication through an algebra tensor, generalizes hypercomplex systems while preserving dimensionality under multiplication. For such structures to be well-defined and reversible, the existence of a unit element must be analytically established rather than assumed.

Materials and methods: The action of a tensorial number is represented as a linear operator. The search for left and right units is reduced to solving a system of linear equations where the contraction of the algebra tensor with an unknown vector yields the identity matrix. Moore–Penrose pseudoinverse matrices are used to obtain analytical expressions for all possible cases — nonexistence, uniqueness, or infinite multiplicity of units.

Results: Explicit formulas for left and right units are derived, and their equivalence is proven: if at least one right and one left unit exist, they coincide, forming a unique commutative identity that renders the algebra unital. A computational implementation of the method confirms correctness on classical algebras such as complex numbers and quaternions.

Conclusion: The existence of a unique commutative unit is a necessary condition for unital tensorial algebras. The proposed analytical and computational framework enables automated detection and verification of unit elements for arbitrary algebra tensors, supporting applications in artificial intelligence, neural modeling, and physical systems with tensorial symmetry.

РЕЗЮМЕ

Д. И. Дзобоев

Поиск единиц в тензориальной алгебре

Тензориальная алгебра, задающая умножение элементов с помощью тензора алгебры, является обобщением гиперкомплексных систем и обеспечивает сохранение размерности пространства при умножении. Для корректного построения алгебраической структуры требуется наличие единичного элемента, что в общем случае не гарантируется свойствами тензора.

Рассмотрено представление действия тензориального числа как линейного оператора и доказано, что поиск единицы сводится к решению системы линейных уравнений относительно вектора, при котором свёртка тензора алгебры с этим вектором порождает единичную матрицу. Для нахождения единиц использованы развёртки тензора и псевдообратные матрицы Мура–Пенроуза, позволяющие описывать все типы решений.

Результаты: Получены аналитические формулы для левой и правой единицы, доказаны критерии их существования и показано, что при наличии хотя бы одной правой и одной левой единицы они совпадают, образуя единственную коммутативную нейтраль. Приведена программная реализация алгоритмов поиска единиц и проверена их корректность на примерах известных алгебр.

Наличие единственной коммутативной единицы является необходимым условием унитарности тензориальной алгебры. Разработанный аппарат позволяет выполнять автоматический поиск и анализ структуры единицы в алгебрах произвольной размерности, что открывает возможности применения в задачах искусственного интеллекта, нейросетевого моделирования и физико-математических симметрий.

Дзобоев Д. И. – стажёр-исследователь, НИУ ВШЭ, Факультет компьютерных наук, Лаборатория моделирования и управления сложными системами, 109028, Москва, Покровский бульвар, 11, тел +7(965)199-29-44, dzeboev.daniil@gmail.com. *Область научных интересов:* Тензориальная алгебра, гиперкомплексные числа, неассоциативные структуры, искусственный интеллект, нейронные сети. Orcid 0009-0008-4004-8750

Статья поступила в редакцию 09.10.2025.

В. О. Елисеев¹, В. И. Бондаренко², А. Ю. Максимова¹

¹Федеральное государственное бюджетное научное учреждение

«Институт прикладной механики и математики»

283048, Донецкая Народная Республика, г. Донецк, ул. Розы Люксембург, 74

²Федеральное государственное бюджетное образовательное учреждение

высшего образования «Донецкий государственный университет»

283001, Донецкая Народная Республика, г. Донецк, ул. Университетская, 24

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОГО АССИСТЕНТА ДЛЯ ОБРАЗОВАТЕЛЬНОЙ ОРГАНИЗАЦИИ*

V. O. Eliseev¹, V. I. Bondarenko², A. Y. Maksimova¹

¹Federal State Budgetary Scientific Institution "Institute of Applied Mathematics and Mechanics"

283048, Donetsk People's Republic, Donetsk, Rosa Luxemburg st, 74

²Federal State Budgetary Educational Institution of Higher Education "Donetsk State University"

283001, Donetsk People's Republic, Donetsk, University st, 24

DEVELOPMENT OF AN INTELLIGENT ASSISTANT FOR AN EDUCATIONAL INSTITUTION

В статье предложен метод построения интеллектуальных ассистентов для образовательных организаций на основе больших языковых моделей (LLM). Это позволит пользователям получать справочные ответы об организации в течение минуты вместо часов/суток. Метод был реализован для построения ассистента Донецкого Государственного университета (ДонГУ). Выполнен сбор и предобработка корпоративных данных, обоснован выбор способа их интеграции в LLM. Итоговая система представляет собой классический Retrieval Augmented Generation (RAG) с дообученной LLaMA-3.1-8b (генератор) и Giga-Embeddings-instruct (ретривер). Для оценки качества работы модели генератора использовались метрики совпадения на уровне токенов и оценка семантической близости ответов ассистента к целевым ответам для ряда заготовленных сценариев. Исходный код разработанных скриптов и финальная версия ассистента для ДонГУ находятся в открытом доступе, предложенный подход применим к адаптации для других образовательных организаций.

Ключевые слова: большая языковая модель, обработка естественного языка, вопросно-ответный поиск, генерация, дополненная поиском, LLAMA, тонкая настройка.

The article proposes the method for development intelligent assistants for educational institutions based on Large Language Models (LLMs). This enables users to receive informational responses about an institution within a minute, as opposed to the hours or days typically required. The method was implemented to create an assistant for Donetsk State University (DonSU). The study involved selecting a method for integrating corporate data into an LLM, followed by the direct collection and preprocessing of this data. The final system is a standard Retrieval-Augmented Generation (RAG) pipeline, which utilizes a fine-tuned LLaMA-3.1-8b model as the generator and Giga-Embeddings-instruct as the retriever. To evaluate the performance of the generator model, we used token-level similarity metrics and assessed the semantic similarity between the assistant's responses and the target answers for a set of predefined scenarios. The source code for all implemented steps and the final version of the DonSU assistant are open-source what makes the proposed approach applicable for other educational institutions.

Key words: large language model, natural language processing, question answering, Retrieval-Augmented Generation, LLAMA, fine-tuning.

* Работа выполнена при финансовой поддержке Минобрнауки России в рамках научной темы "Разработка и совершенствование интеллектуальных методов классификации и прогнозирования для задач распознавания образов и моделирования информационных процессов" FREM-2024-0001 (Регистрационный номер 1023111000141-9-1.2.1)

Введение

Развитие современных методов NLP и появление больших языковых моделей (LLM) предоставляет возможность автоматизации работы агентов поддержки путем построения интеллектуальных ассистентов, способных отвечать на вопросы пользователей в режиме реального времени. Внедрение ассистентов в вузы сократит время ожидания ответа для заинтересованных пользователей и освободит время сотрудников. Существует явная необходимость в разработке методологии и прикладных инструментов для создания подобных цифровых помощников.

Существующие LLM не могут быть использованы напрямую в образовательных организациях, таких как Донецкий Государственный университет (ДонГУ), поскольку данные большинства организаций не были использованы для их обучения и по умолчанию модели не способны давать специфические ответы на вопросы, касающиеся их деятельности. Для решения данной проблемы необходимо определить источники корпоративных данных и методы их интеграции с LLM. Также нужно разработать механизм актуализации информации. Настройка такого взаимодействия позволит создать прикладной сервис в виде чат-бота, где можно задавать вопросы и получать ответы, основанные на специфике вуза.

В данной статье предложена методика создания интеллектуальных ассистентов для образовательных организаций и описаны проделанные шаги – от выбора LLM и способа интеграции корпоративных данных до реализации прикладного сервиса.

Цель исследования – разработать методику построения ассистентов для учебных заведений и реализовать ассистента для ДонГУ. Методика должна содержать универсальные шаги для вузов РФ и учитывать ограниченность ресурсов (в ДонГУ использовали 2 ПК с NVIDIA RTX 4080 SUPER и RTX A4000).

Для достижения поставленной цели необходимо решить следующие *задачи*:

- 1) выбрать LLM для генерации ответов;
- 2) разработать интеграцию корпоративных данных;
- 3) выполнить предобработку и фильтрацию данных;
- 4) реализовать сервис для взаимодействия с ассистентом

Выбор LLM и механизма интеграции данных

Большинство общедоступных LLM не обладает знаниями о внутренних процессах конкретных организаций (исключением могут быть общеизвестные корпорации, однако и в этом случае знания будут носить лишь общий характер). В связи с этим, для построения интеллектуального ассистента, работающего на основе LLM, необходимо найти способ внедрить корпоративную информацию в модель.

Существует 3 способа интеграции корпоративных данных в LLM: обучить собственную модель с нуля, как это было в случае BloombergGPT [1]; провести тонкую настройку уже предобученной LLM [2], либо же использовать механизм генерации, дополненной поиском (Retrieval Augmented Generation, RAG) [3].

Обучение собственной модели с нуля не подходит для решаемой задачи, поскольку данный процесс является крайне ресурсоемким (необходимо иметь в распоряжении кластеры GPU) и требует наличия текстовых корпусов большого объема (сотни миллиардов токенов) и высокого качества.

Тонкая настройка предобученной LLM снижает требования к вычислительным ресурсам и размеру и качеству текстового корпуса. Поскольку модель не учится моделировать язык с нуля, ей достаточно «показать» текстовый корпус с корпоративными

данными конкретной организации, и она будет способна отвечать на связанные с ней вопросы. В данном случае достаточно иметь корпус объемом от сотен тысяч до миллионов токенов, а процесс осуществим при ограниченных ресурсах, при использовании метода адаптации матрицами низкого ранга (LoRA) [4]. Так как процессы в образовательных организациях подвержены частым изменениям, важно поддерживать знания модели в актуальном состоянии. Для их актуализации необходимо часто переобучать модель, что ресурсозатратно, поэтому дообучение в чистом виде не подходит для решаемой задачи.

Наименее ресурсоемким способом внедрить корпоративные знания в LLM является построение Retrieval Augmented Generation (RAG) систем. Данный подход не требует дообучения LLM и позволяет использовать как локальные модели из открытого доступа, так и коммерческие по API. Корпоративные данные сначала векторизуются с помощью модели-ретривера – предобученной encoder-only модели, после чего помещаются в векторную базу. Во время работы системы ретривер получает пользовательский запрос, векторизует его, после чего ищет в векторной базе N наиболее семантически близких документов, в которых потенциально может находиться ответ. В качестве метрики близости [5] обычно используется косинусное сходство или L2-расстояние. Найденные документы передаются вместе с запросом пользователя модели-генератору, которая должна сформировать ответ на запрос на основе собственных данных и найденной ретривером информации. Недостатком такой системы является зависимость от качества модели-ретривера [6], так как в случае ошибок в его работе модель-генератор не получит необходимой корпоративной информации и не сможет сгенерировать качественный ответ.

В данной работе мы объединили дообучение LLM и RAG, используя дообученную специальным образом LLM в качестве генератора. Однако мы не просто дообучили модель отвечать на вопросы, связанные с деятельностью ДонГУ, передав в нее соответствующие корпоративные данные в виде инструкционного датасета [7], а натренировали ее извлекать релевантные ответы из предложенных ретривером документов в соответствии с пользовательским запросом, как того требует RAG. Для этого был разработан контекстно-инструкционный набор данных [8], состоящий из потенциального пользовательского запроса, трех наиболее релевантных относительно запроса фрагментов документов из векторной базы и целевого ответа модели. Процедура сбора данных для данной работы будет описана в следующем разделе. Подобный подход был успешно применен при дообучении моделей-генераторов для систем RAG, решающих общие задачи используя мультидоменные знания [9], [10]. В нашем случае набор данных был собран под задачу вопросно-ответного поиска по внутренним процессам ДонГУ.

В качестве дообучаемой модели была выбрана квантизованная LLaMA-3.1-8b, что призвано провести процесс дообучения с помощью LoRA на ограниченных ресурсах и сохранить приемлемый уровень качества ответов.

Сбор и подготовка контекстно-инструкционного набора данных

Для дообучения генератора собирается обучающий корпус и данные для заполнения базы знаний. Источником выбран официальный сайт ДонГУ, так как его содержимое структурировано, машиночитаемо и обновляемо, что регламентировано нормативами [11]. Это делает предложенный способ применимым для других вузов.

Сбор и подготовка данных включают:

- 1) выгрузку текстов с сайта,
- 2) фильтрацию нерелевантных материалов,
- 3) построение датасета и векторной базы,
- 4) добавление релевантных фрагментов как context.

Данные собираются скриптами из открытого доступа [12]. Текстовое содержимое страниц доступно через уникальный в рамках сайта CSS-селектор, что позволяет использовать скрипты для других вузов. В результате обхода получено 2912 файлов для обучения модели и наполнения базы, и 3546 документов для инференса.

Для фильтрации нерелевантного содержимого использовались три LLM (Gemma2-9b-it [13], LLaMA-3.1-70B-Instruct, LLaMA-3.3-70B-Instruct), которые оценивали файлы по пятибалльной шкале. Распределение оценок показано на рисунке 1.

Распределения оценок файлов

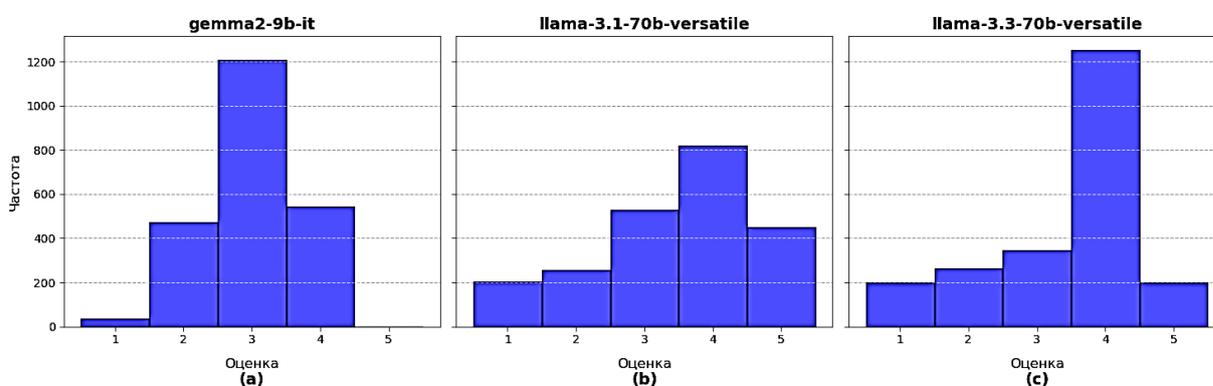


Рисунок 1 – Распределения оценок файлов, полученные от выбранных LLM

Полученные распределения оценок смещены вправо, а мода совпадает с предыдущим максимальным значением. Анализ распределений показал, что модели LLaMA более эффективно анализируют текстовые данные и лучше подходят для синтетической генерации инструкционных пар.

Файлы со средней оценкой 2 балла и ниже были исключены, что свидетельствует о нерелевантности данных текстов. После фильтрации из 2912 файлов осталось 2337, этого объема достаточно для формирования качественного инструкционного датасета. Исходный код модуля фильтрации и использованный промпт доступны в открытом доступе [14].

Для построения классического инструкционного датасета на основе отобранных текстовых файлов было решено также использовать модель LLaMA-3.3-70B-Instruct как лучшую в процессе исключения нерелевантных текстов. Был создан промпт, включающий в себя название файла, его содержимое и метку последнего обновления. и запрос модели сгенерировать максимально возможное число инструкций по заданной информации о странице, либо, если поданный текст не является релевантным – отказаться от генерации, выдав сообщение об ошибке.

После первичного запуска процесса генерации инструкций и анализа распределения количеств инструкций по файлам было выявлено, что для ряда информативных файлов сгенерировано малое количество инструкций. Повторный промптинг моделей для файлов, получивших 3 и менее инструкции, с пониженными значениями температуры (с 1.0 до 0.7) и top-p (с 1.0 до 0.8) позволил получить более детерминированные ответы модели [15]. В результате получено 806 новых инструкций. Итого

датасет включил в себя 25633 инструкции, что достаточно для дообучения модели, работающей с единственным доменом. Исходный код модуля генерации инструкционного датасета также находится в открытом доступе [16]. Для расширения построенного инструкционного датасета полем с контекстом была построена векторная база на основе релевантных текстовых файлов. В качестве ретривера выбрана модель Giga-Embeddings-instruct [17], разработанная командой Сбера. Несмотря на относительно небольшой размер (2.5 миллиарда параметров), контекст модели 4096 токенов, а размерность получаемых эмбеддингов 2048, что делает их более репрезентативными и позволяет лучше описать семантику входной последовательности [18]. Выбранная модель была обучена на русском языке, что позволяет описать большой объем текста меньшим числом токенов.

Документы были фрагментированы по 500 токенов с перекрытием 50 по абзацам, что позволило уместить их в контекстное окно модели LLaMA-3.1-8b и избежать разрывов контекста внутри предложений. После удаления дубликатов и фрагментации сформировано 9935 уникальных фрагментов для векторной базы.

Контекстно-инструкционный набор данных для дообучения LLM в системе RAG был построен следующим образом: для каждого input инструкционного датасета выбраны три наиболее релевантных фрагмента из векторной базы на основе косинусного сходства, их содержимое объединено в поле context результирующего датасета. Исходный код модулей проекта доступен в открытом доступе [16].

Дообучение и тестирование модели-генератора

Для обучения seq2seq моделей требуется объединение текстовых полей датасета, что реализовано с помощью форматированного промпта. Такой подход показал себя лучше всего для систем RAG [19], так как имитирует процесс инференса.

Важную роль играет способ разметки входных и выходных данных. Исследованы два варианта: первый — input содержит форматированный промпт с контекстом и запросом, output — целевой ответ; второй — input и output совпадают, но в output все токены до метки ответа («[ANSWER]») замаскированы значением -100. Это позволяет функции потерь учитывать только токены ответа, что фокусирует модель на генерации финальной реплики [20].

В таблице 1 приведены описательные характеристики полного датасета, обучающей и тестовой подвыборок в токенизированном виде (разделение 90/10).

Таблица 1 – Описательные статистики токенизированного датасета

Статистика (в токенах, округлено до целых)	Полный датасет	Train	Test
количество	40 156 778	36 137 150	4 019 628
минимальный пример	396	396	436
максимальный пример	3282	3282	2288
средняя длина примера	1567	1567	1568
медианная длина примера	1673	1673	1674

Модель-генератор была обучена более чем на 36 миллионах токенов. На рисунке 2 представлено распределение длин экземпляров в полном датасете.

Распределение смещено вправо, есть выбросы, размером около 3000 токенов, но основная часть датасета состоит из примеров длиной в 1500-2000 токенов. Так как в нашем случае ресурсы ограничены, было решено не использовать мини-пакеты напрямую, во избежание необходимости паддинга и выравнивания датасета по длине.

Дообучение проходило на машине с видеоадаптером NVIDIA GeForce RTX 4080 SUPER (16 Гб видеопамяти), дообучаемая модель была квантизована в формате NF4 [21]. Квантизация проходила в 2 этапа: сначала веса переводились в 8-битный формат, а после в 4-битный, снижая ошибку квантизации без увеличения размера модели [21].

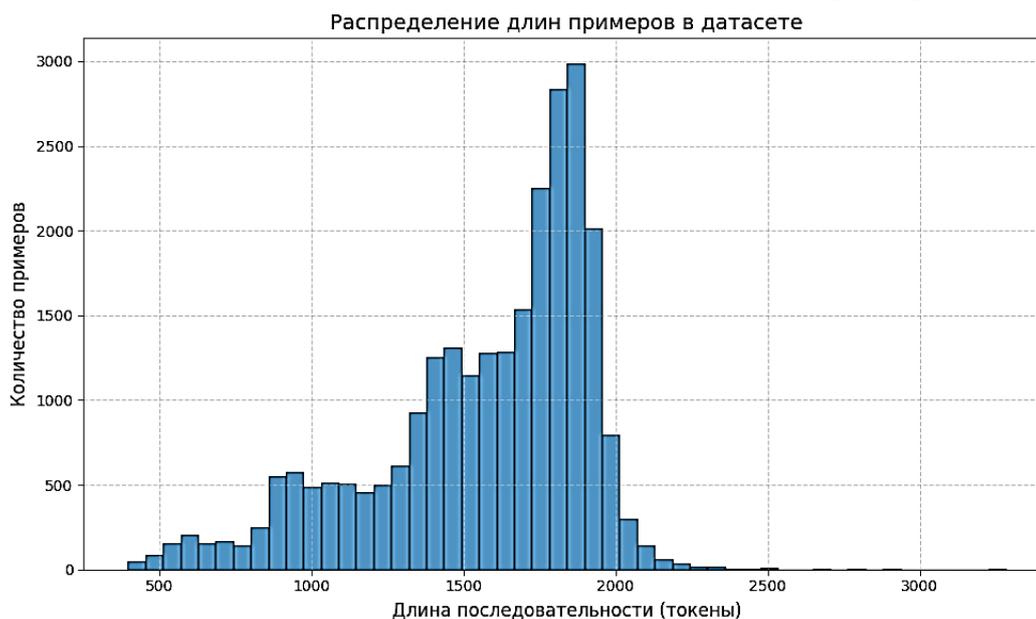


Рисунок 2 – Распределение длин экземпляров в датасете (в токенах)

В качестве метода дообучения модели выбрана LoRA [6] с параметрами:

- 1) $r = 8$ (ранг матриц A и B, аппроксимирующих матрицу градиентов, чем ниже, тем меньше параметров необходимо обучить);
- 2) $\alpha = 32$ (коэффициент масштабирования, показывающий, насколько новые знания повлияют на имеющиеся знания – полученная LoRA-вставка усиливается в α/r раз перед слиянием с моделью);
- 4) $\text{dropout} = 0.05$;
- 5) дообучаемые модули: q_proj , k_proj , v_proj , o_proj – модули блоков внимания и $gate_proj$, up_proj , $down_proj$ – модули MLP.

Использование квантизации и механизма LoRA позволит существенно сократить ресурсные затраты на процесс обучения. Так, после квантизации модель LLaMA-3.1-8b имеет 1 050 939 392 обучаемых параметра. При включении LoRA-адаптеров при указанных выше параметрах модель содержит только 20 971 520 обучаемых параметров (поскольку в данном случае исходные параметры модели замораживаются), тем самым мы уменьшили число обучаемых параметров более чем в 50 раз относительно квантизованной версии и в 400 относительно исходной.

В качестве главных параметров обучения были выбраны следующие:

- 1) $\text{learning_rate} = 2e-4$;
- 2) $\text{weight_decay} = 1e-3$;

3) *warmup_ratio* = 0.1 (10% от общего числа шагов обучения будут использоваться для постепенного увеличения *learning_rate* от нуля до $2e-4$);

4) *num_epoch* = 25;

5) *gradient_accumulation_steps* = 8 (градиентов матриц А и В обновляются каждые 8 шагов, имитируя *batch_size* равный 8, без увеличения расхода видеопамяти);

6) *batch_size* = 1 (фактически на используемой GPU в один момент будет обрабатываться только 1 экземпляр данных).

Процесс обучения для каждого способа передачи данных в модель на одной GPU NVIDIA GeForce RTX 4080 SUPER занял около 200 часов. На рисунке 3 показаны графики изменения функции потерь (кросс-энтропия) и средней точности предсказания токенов. Поскольку для разных способов графики практически идентичны, приведем только второй.

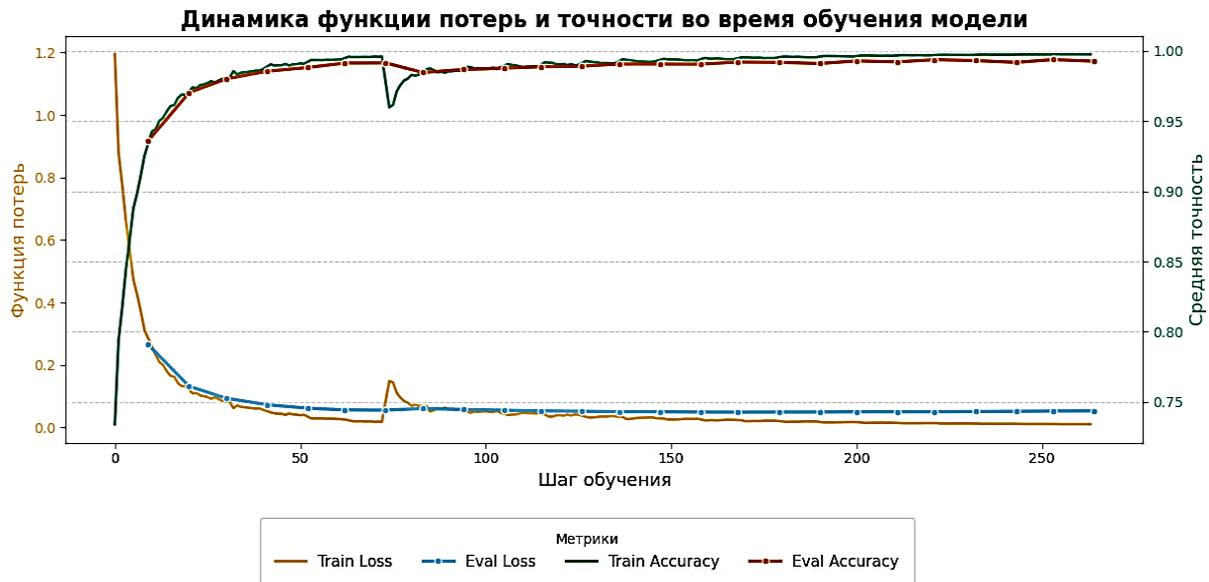


Рисунок 3 – График изменения функции потерь и средней точности предсказания токенов

На финальной, 25-й эпохе средняя точность модели на тестовой выборке при первом способе дообучения составила 99.35%, а при втором – 99.26%. Наилучшую точность на тестовой выборке при первом способе обучения модель показала после 24-й эпохи (99.45%). При втором способе обучения лучшую точность модель также показала после 24-й эпохи, и она составила 99.38%. Значение функции потерь на финальной эпохе при первом способе обучения составило примерно 0.04, лучший результат наблюдался на 16-й, составив примерно 0.042. При втором способе дообучения значение функции потерь на финальной эпохе составило примерно 0.05, лучший результат, равный 0.049, модель показала на 16-й. Также исходя из графиков можно заметить отсутствие значительного разрыва между качеством работы модели на тренировочных и на тестовых данных, что дает возможность предположить отсутствие факта переобучения модели.

Перед тестированием дообученной модели требуется интеграция LoRA-адаптеров в исходные веса. Если модель была квантизована, а адаптеры — *bfloat16*, то для слияния сначала проводится деквантизация либо загрузка в формате *bfloat16* и объединение весов, что требует большой видеопамяти и работает не с любым оборудованием. Для проверки была реализована интеграция обоими способами: при ограниченной

памяти модель выгружалась на CPU, делилась на части до 5 Гб и объединялась пофрагментно на GPU, позволяя интегрировать адаптеры без дополнительной квантизации.

Для сравнения качества моделей использовали 30% примеров из контекстно-инструкционного датасета. Ответы на вопросы сначала генерировала базовая, затем дообученные версии LLaMA-3.1-8b. Для оценки применяли BLEU [22], ROUGE [23] и косинусное сходство (Giga-Embeddings-instruct), затем усредняли значения по тестовой подвыборке (7690 примеров). В таблице 2 приведены результаты для бейзлайна и дообученных моделей. Суффиксы «sep» — для моделей с отдельной подачей context в inputs и labels; «un» — для моделей с полным промптом с маскировкой; приставка «nod» — для моделей слияния весов без квантизации.

Таблица 2 – Усредненные метрики сходства эталонных ответов и ответов, сгенерированных классической LLaMA-3.1-8b и ее дообученными версиями

Метрика	baseline-llama	tuned-uni-llama	nod-tuned-uni-llama	tuned-sep-llama	nod-tuned-sep-llama
BLEU	0.177851	0.301470	0.311701	0.114690	0.130832
ROUGE-1	0.152560	0.138591	0.133058	0.112703	0.117614
ROUGE-2	0.065618	0.074913	0.072057	0.043937	0.046375
ROUGE-L	0.151232	0.137470	0.132025	0.110703	0.115748
cosine-similarity	0.716595	0.755788	0.761709	0.675485	0.685384

Обучение моделей с отдельной подачей промпта и ответа дало худший результат, чем базовая LLaMA-3.1-8b, то есть этот метод обучения неэффективен. Зато обучение с одинаковыми inputs и labels (с маскированием части запроса у labels) дало двукратный прирост BLEU, небольшой прирост ROUGE-2, незначительное отставание по ROUGE-1, ROUGE-L, а также преимущество по семантической близости генераций к целевым ответам. Модель уверенно генерирует биграммы, но немного уступает по отдельным словам и предложениям при ROUGE-сравнении, зато превосходит по смысловой близости.

Преимущество моделей с интегрированными без квантизации весами минимально и проявляется на третьих-четвертых знаках после запятой, то есть деквантизация весов после квантизации негативно не влияет при отсутствии возможностей для слияния весов в исходном виде. Все дообученные варианты моделей доступны на HuggingFace [24].

Реализация сервиса интеллектуального ассистента

После дообучения модели можно строить прикладной сервис, включающий RAG пайплайн и пользовательский интерфейс взаимодействия с системой. На рисунке 5 представлена структура реализованной системы интеллектуального ассистента.

Векторная база знаний предварительно наполнялась релевантными документами, использованными для инструкционного датасета. В базе документы делились на фрагменты по 2000 токенов с перекрытием 200 (Giga-Embeddings-instruct). Для обновления базы создан инструмент, не требующий изменений в системе.

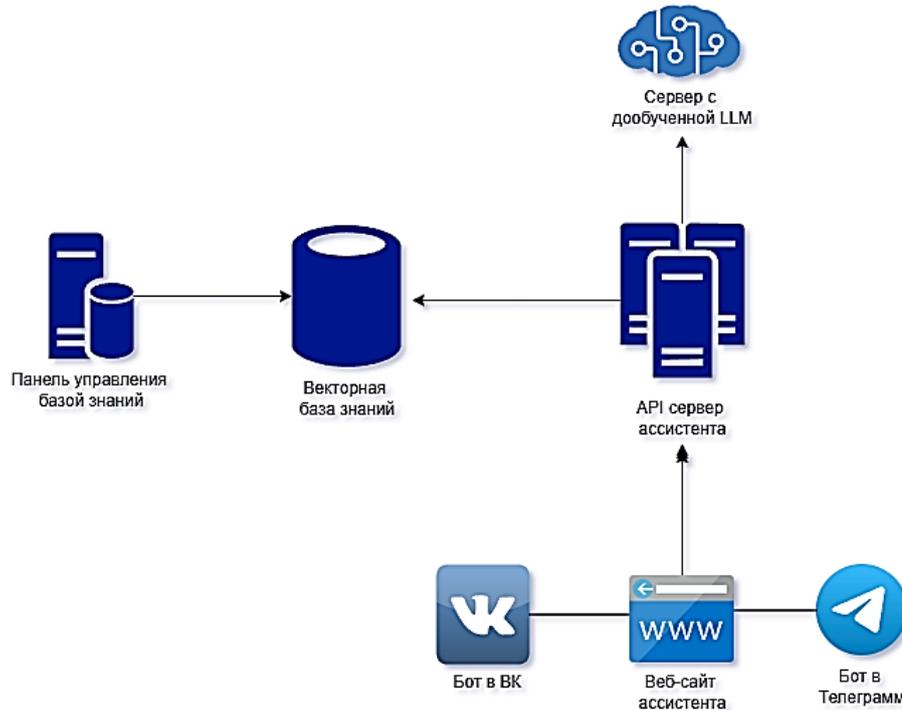


Рисунок 5 – Структура системы интеллектуального ассистента

Дообученная модель размещена на сервере с RTX 4080 SUPER, работая через LMStudio (OpenAI API совместимость). Использована 4-битная квантизация Q4_K_M (K-means), которая сокращает размер при сохранении производительности [25], и формат GGUF, поддерживающий запуск на CPU [26].

API ассистента реализован на FastAPI и обслуживает клиентские приложения, хранит историю чатов (20 последних сообщений), автоматически перефразирует запросы с учётом истории для лучшего понимания и разрешения местоименной анафоры, извлекает 3 ближайших документа из базы и формирует промпт для LLM.

В качестве клиентов доступны веб-чат, боты в VK и Telegram, а также API для сторонних сервисов. Весь исходный код находится в открытом доступе [27].

Выводы

Разработана методика создания ассистентов на основе LLM для образовательных организаций, реализован ассистент для ДонГУ. Методика использует дообученную на контекстно-инструкционном датасете LLM как генератор для RAG, что преодолевает ограничения раздельного применения дообучения и RAG.

Для ассистента ДонГУ была выбрана модель LLaMa-3.1-8b, дообученная QLoRA на датасете, собранном с сайта университета. Модель показала результаты выше базовой на 13% (BLEU) и 4.5% (векторная близость). После сравнения лучшая версия интегрирована в систему RAG с обновляемой базой знаний и реализован сервис с клиентскими приложениями.

Весь исходный код находится в открытом доступе, что упрощает использование методики для других учебных заведений.

Список литературы

1. Wu S. et al. Bloombergpt: A large language model for finance //arXiv preprint arXiv:2303.17564. – 2023.
2. Suchin Gururangan, Ana Marasovic, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A Smith. Don't stop pretraining: Adapt language models to domains and tasks. arXiv preprint arXiv:2004.10964, 2020.
3. Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Kuttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in Neural Information Processing Systems, 33:9459–9474, 2020.
4. Hu, E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L. & Chen, W. Lora: Low-rank adaptation of large language models, International Conference On Learning Representations. (2021).
5. Андриевская, Н. К. Гибридная интеллектуальная мера оценки семантической близости / Н. К. Андриевская // Проблемы искусственного интеллекта. – 2021. – № 1(20). – С. 4-17. – EDN ZDZK GK.
6. Scott Barnett, Stefanus Kurniawan, Srikanth Thudumu, Zach Brannelly, and Mohamed Abdelrazek. Seven failure points when engineering a retrieval augmented generation system. In Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI, pp. 194–199, 2024.
7. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A. & Others Training language models to follow instructions with human feedback, Advances In Neural Information Processing Systems. 35 pp. 27730-27744 (2022).
8. Eliseev, V.; Maksimova, A.; Bondarenko, V. Transforming raw corporate texts into instruction dataset for fine-tuning generator within a RAG system. System Informatics 2025, 27, 77-92.
9. Zhang T. et al. Raft: Adapting language model to domain specific rag //First Conference on Language Modeling. 2024.
10. Liu W. et al. RAG-Instruct: Boosting LLMs with Diverse Retrieval-Augmented Instructions //arXiv preprint arXiv:2501.00353. 2024.
11. Методические рекомендации «Методические рекомендации представления информации об образовательной организации высшего образования в открытых источниках с учетом соблюдения требований законодательства в сфере образования» (вступают в силу с 01.09.2024) // Федеральная служба по надзору в сфере образования и науки. – 2024
12. Eliseev Vadim. Website Crawler [Электронный ресурс]. URL: <https://github.com/EliseevVadim/WebsiteCrawler>.
13. Team G. et al. Gemma 2: Improving open language models at a practical size //arXiv preprint arXiv:2408.00118. 2024.
14. Eliseev Vadim. Texts Estimator [Электронный ресурс]. URL: <https://github.com/EliseevVadim/TextsEstimator>.
15. Lee M. H. J. Examining the Robustness of Homogeneity Bias to Hyperparameter Adjustments in GPT-4 //arXiv preprint arXiv:2501.02211. – 2025.
16. Eliseev Vadim. Instructions Generator [Электронный ресурс]. URL: <https://github.com/EliseevVadim/InstructionsGenerator>.
17. Giga-Embeddings-instruct | HuggingFace [Электронный ресурс]. URL: <https://huggingface.co/ai-sage/Giga-Embeddings-instruct>.
18. Kaplan J. et al. Scaling laws for neural language models //arXiv preprint arXiv:2001.08361. – 2020.
19. Rodrigues J., Branco A. Meta-prompting Optimized Retrieval-Augmented Generation //EPIA Conference on Artificial Intelligence. – Cham: Springer Nature Switzerland, 2024. С. 203-214.
20. Huerta-Enochian M., Ko S. Y. Instruction Fine-Tuning: Does Prompt Loss Matter? //arXiv preprint arXiv:2401.13586. – 2024.
21. Dettmers T. et al. Qlora: Efficient finetuning of quantized llms //Advances in neural information processing systems. 2023. Т. 36. С. 10088-10115.
22. Reiter E. A structured review of the validity of BLEU //Computational Linguistics. 2018. Т. 44. №. 3. С. 393-401.
23. Lin C. Y. Rouge: A package for automatic evaluation of summaries //Text summarization branches out. 2004. С. 74-81.
24. Eliseev Vadim. DonSU LLaMa Family [Электронный ресурс]. URL: <https://huggingface.co/collections/insidious316/donsu-llama-family>.
25. Zhao E. et al. Quantitative Analysis of Performance Drop in DeepSeek Model Quantization //arXiv preprint arXiv:2505.02390. – 2025.
26. Kavathekar I. et al. Small Models, Big Tasks: An Exploratory Empirical Study on Small Language Models for Function Calling //arXiv preprint arXiv:2504.19277. 2025.
27. Eliseev Vadim. DonSU Intelligent Assistant [Электронный ресурс]. URL: <https://github.com/EliseevVadim/DonSU-Intelligent-Assistant>.

References

1. Wu S. et al. Bloomberggpt: A large language model for finance //arXiv preprint arXiv:2303.17564. – 2023.
2. Suchin Gururangan, Ana Marasovic, Swabha Swayamdipta, Kyle Lo, Iz Beltagy, Doug Downey, and Noah A Smith. Don't stop pretraining: Adapt language models to domains and tasks. arXiv preprint arXiv:2004.10964, 2020.
3. Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Kuttler, Mike Lewis, Wen-tau Yih, Tim Rocktaschel, et al. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474, 2020.
4. Hu, E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L. & Chen, W. Lora: Low-rank adaptation of large language models, *International Conference On Learning Representations*. (2021).
5. Andrievskaya N. K. Hybrid Intelligent Measure for Assessing Semantic Similarity // *Problems of Artificial Intelligence*. – 2021. – No. 1(20). – Pp. 4-17.
6. Scott Barnett, Stefanus Kurniawan, Srikanth Thudumu, Zach Brannelly, and Mohamed Abdelrazek. Seven failure points when engineering a retrieval augmented generation system. In *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI*, pp. 194–199, 2024.
7. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A. & Others Training language models to follow instructions with human feedback, *Advances In Neural Information Processing Systems*. 35 pp. 27730-27744 (2022).
8. Eliseev, V.; Maksimova, A.; Bondarenko, V. Transforming raw corporate texts into instruction dataset for fine-tuning generator within a RAG system. *System Informatics 2025*, 27, 77-92.
9. Zhang T. et al. Raft: Adapting language model to domain specific rag // *First Conference on Language Modeling*. – 2024.
10. Liu W. et al. RAG-Instruct: Boosting LLMs with Diverse Retrieval-Augmented Instructions //arXiv preprint arXiv:2501.00353. – 2024.
11. Methodical Guidelines: Methodical Guidelines for the Presentation of Information about a Higher-Education Institution in Open Sources, Taking into Account Compliance with Education-Legislation Requirements (effective 01.09.2024) // *Federal Service for Supervision in Education and Science*, 2024.
12. Eliseev Vadim. Website Crawler [Электронный ресурс]. – URL: <https://github.com/EliseevVadim/WebsiteCrawler>.
13. Team G. et al. Gemma 2: Improving open language models at a practical size //arXiv preprint arXiv:2408.00118. – 2024.
14. Eliseev Vadim. Texts Estimator [Electronic resource]. – URL: <https://github.com/EliseevVadim/TextsEstimator>.
15. Lee M. H. J. Examining the Robustness of Homogeneity Bias to Hyperparameter Adjustments in GPT-4 //arXiv preprint arXiv:2501.02211. – 2025.
16. Eliseev Vadim. Instructions Generator [Electronic resource]. – URL: <https://github.com/EliseevVadim/InstructionsGenerator>.
17. Giga-Embeddings-instruct | HuggingFace [Электронный ресурс]. – URL: <https://huggingface.co/ai-sage/Giga-Embeddings-instruct>.
18. Kaplan J. et al. Scaling laws for neural language models //arXiv preprint arXiv:2001.08361. – 2020.
19. Rodrigues J., Branco A. Meta-prompting Optimized Retrieval-Augmented Generation // *EPIA Conference on Artificial Intelligence*. – Cham: Springer Nature Switzerland, 2024. – C. 203-214.
20. Huerta-Enochian M., Ko S. Y. Instruction Fine-Tuning: Does Prompt Loss Matter? //arXiv preprint arXiv:2401.13586. – 2024.
21. Dettmers T. et al. Qlora: Efficient finetuning of quantized llms // *Advances in neural information processing systems*. – 2023. – T. 36. – C. 10088-10115.
22. Reiter E. A structured review of the validity of BLEU // *Computational Linguistics*. 2018. T. 44. №. 3. – C. 393-401.
23. Lin C. Y. Rouge: A package for automatic evaluation of summaries // *Text summarization branches out*. – 2004. – C. 74-81.
24. Eliseev Vadim. DonSU LLaMa Family [Electronic resource]. – URL: <https://huggingface.co/collections/insidious316/donsu-llama-family>.
25. Zhao E. et al. Quantitative Analysis of Performance Drop in DeepSeek Model Quantization //arXiv preprint arXiv:2505.02390. – 2025.
26. Kavathekar I. et al. Small Models, Big Tasks: An Exploratory Empirical Study on Small Language Models for Function Calling //arXiv preprint arXiv:2504.19277. – 2025.
27. Eliseev Vadim. DonSU Intelligent Assistant [Electronic resource]. URL: <https://github.com/EliseevVadim/DonSU-Intelligent-Assistant>.

RESUME

V. O. Eliseev, V. I. Bondarenko, A. Y. Maksimova

Development of an Intelligent Assistant for an Educational Institution

Background: timely provision of reference information is crucial in both education and business. Previously, this problem was addressed through human support staff, what caused big delays. Advances in Natural Language Processing (NLP) and the emergence of Large Language Models (LLM) enabled the automation of this process by developing intelligent assistants. The main goal of this study is creating the method that enables building of such assistants for educational institutions and which is applicable for majority of Russian universities with minimal effort. Also, the intelligent assistant of Donetsk State University (DonSU) was built using this method.

Materials and methods: various methods of integrating of corporate data into LLMs were analyzed among with their disadvantages. The combined approach of LLMs fine-tuning and building of Retrieval Augmented Generation (RAG) systems was proposed. It involves fine-tuning an LLM using the context based instructional dataset that consists of input-context-output triples and using it as a generator in RAG system. As part of applying the method for DonSU, a dataset of the proposed format was built based on raw textual data from the university's website. The LLaMA-3.1-8B model was then fine-tuned on this dataset to serve as a generator. The Giga-Embeddings-Instruct model was used as the retriever for dataset construction and in the system in general.

Results: the LLaMA-3.1-8B model fine-tuned on the built dataset demonstrated a 13% improvement over the base model in BLEU score and a 4.5% increase in cosine similarity between generated and target responses across several interaction scenarios. The resulting best-performing model was integrated into the RAG-based intelligent assistant system of DonSU, along with a vector database populated with content from the university's website and an implemented mechanism for its updating. Additionally, an applied service was developed, enabling interaction with the assistant via a web interface and bots in Telegram and VK.

Conclusion: the proposed method for building intelligent assistants proved effective in the development of the DonSU intelligent assistant. The final system produces high-quality responses to user queries while requiring minimal resources during development. The source code for the scripts used to collect and prepare the context based instructional dataset, as well as the code for the DonSU intelligent assistant system, is open source, making it easy to implement the approach for other educational institutions.

РЕЗЮМЕ

В. О. Елисеев, В. И. Бондаренко, А. Ю. Максимова

*Разработка интеллектуального ассистента
для образовательной организации*

Справочная информация: своевременное предоставление справочной информации востребовано как в образовании, так и в бизнесе. Ранее задача решалась вручную сотрудниками поддержки, что приводило к большим задержкам. Развитие технологий обработки естественного языка (NLP) и появление больших языковых моделей (LLM) позволили автоматизировать этот процесс за счет создания интеллектуальных ассистентов. Целью этой работы является разработка методики, позволяющей создавать интеллектуальных ассистентов для образовательных организаций, и применимой к различным вузам РФ с минимальными затратами, а также ее реализация для построения интеллектуального ассистента Донецкого Государственного университета (ДонГУ).

Материалы и методы: были рассмотрены существующие способы интеграции корпоративных данных в LLM, затронуты их недостатки. Предложен подход, комбинирующий дообучение LLM и построение систем генерации, дополненной поиском, путем дообучения модели на контекстно-инструкционном наборе данных (тройки запрос-контекст-ответ) и ее использования в качестве генератора в системе RAG. В рамках применения метода для ДонГУ был собран набор данных предложенного формата на основе сырых текстовых данных с сайта вуза, после чего на нем была дообучена LLaMA-3.1-8b для роли генератора. В качестве ретривера для формирования набора данных выступала модель Giga-Embeddings-instruct.

Результаты: дообученная на собранном датасете LLaMA-3.1-8b показала прирост по сравнению с базовой моделью на 13% по метрике BLEU и на 4.5% по косинусному сходству генераций к целевым ответам для ряда сценариев взаимодействия. Полученная лучшая модель была интегрирована в систему RAG интеллектуального ассистента ДонГУ вместе с векторной базой данных, наполненной содержимым сайта университета и реализованным механизмом ее актуализации. Также был реализован прикладной сервис, позволяющий взаимодействовать с ассистентом с помощью веб-интерфейса и ботов в Telegram и VK.

Вывод: предложенная методика построения интеллектуальных ассистентов успешно продемонстрировала себя при разработке интеллектуального ассистента для ДонГУ, полученная система генерирует качественные ответы на запросы пользователей при использовании минимальных ресурсов при ее разработке. Исходный код скриптов по сбору и подготовке контекстно-инструкционного датасета, а также код системы интеллектуального ассистента ДонГУ находятся в открытом доступе, что позволяет легко имплементировать подход для других образовательных организаций.

Елисеев Вадим Олегович – стажер-исследователь лаборатории интеллектуальных систем Федерального государственного бюджетного научного учреждения Институт прикладной математики и механики.

Область научных интересов: искусственный интеллект, машинное обучение, нейронные сети, обработка естественного языка, генеративные и большие языковые модели.

Бондаренко Виталий Иванович – кандидат технических наук, доцент кафедры компьютерных технологий физико-технического факультета, Федерального государственного бюджетного образовательного учреждения высшего образования "Донецкого Государственного Университета".

Область научных интересов: искусственный интеллект, интеллектуальный анализ данных, машинное обучение, математическое моделирование гидро- и теплофизических процессов, разработка пользовательских интерфейсов для прикладных программ моделирования.

Максимова Александра Юрьевна – кандидат технических наук, заведующий лабораторией интеллектуальных систем, Федерального государственного бюджетного научного учреждения Институт прикладной математики и механики.

Область научных интересов: машинное обучение, анализ данных, распознавание образов, миварный логический вывод, нейросетевые модели.

Статья поступила в редакцию 15.10.2025.

УДК 004.932.2

DOI 10.24412/2413-7383-2025-4-39-49-59

О. А. Лямцев^{1,2}, И. И. Максименко^{1,2}¹ФГБНУ «Институт прикладной математики и механики»

283048, г. Донецк, ул. Розы Люксембург, 74

²ФГБОУ ВО «Донецкий государственный университет»

283001, г. Донецк, ул. Университетская, 24

ОБЗОР И ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ ТРЕХМЕРНОЙ ОЦЕНКИ ПОЗЫ ЧЕЛОВЕКА ПО ОДНОМУ ИЗОБРАЖЕНИЮ

O. A. Lyamtsev^{1,2}, I. I. Maksimenko^{1,2}¹Federal State Budgetary Scientific Institution "Institute of Applied Mathematics and Mechanics"

283048, Donetsk, Rosa Luxemburg str., 74

²Federal State Budgetary Educational Institution of Higher Education "Donetsk State University"

283001, Donetsk, Universitetskaya str., 24

OVERVIEW AND CHALLENGES OF USING DEEP LEARNING FOR THREE-DIMENSIONAL ESTIMATION OF A PERSON'S POSTURE FROM A SINGLE IMAGE

В статье рассмотрена задача трехмерной оценки позы человека по монокулярному изображению, приведено математическое описание, выделены основные подходы, проблемы в задаче и их решения, архитектуры нейронных сетей. Среди архитектур нейронных сетей наиболее популярны графовые сверточные сети, извлекающие локальную информацию из ближайших суставов и трансформеры, моделирующие глобальные зависимости через механизм внутреннего внимания. Среди ключевых проблем при решении задачи можно выделить: 1) неопределенность глубины; 2) окклюзия; 3) недостаточный объем и разнообразие обучающих данных.

Ключевые слова: трехмерная оценка позы человека, компьютерное зрение, глубокое обучение, нейронные сети, монокулярное изображение

The article considers the task of 3D human pose estimation from monocular image, provides a mathematical description, highlights the main approaches, problems in the task and their solutions, neural network architectures. Among neural network architectures, graph convolutional networks are the most popular, extracting local information from nearby joints and transformers, modeling global dependencies through the internal attention mechanism. Among the key problems in solving the task, we can distinguish: 1) depth uncertainty; 2) occlusion; 3) insufficient volume and diversity of training data.

Keywords: 3D human pose estimation, computer vision, deep learning, neural networks, monocular image

Введение

Оценка позы человека в трёхмерном пространстве (*Human Pose Estimation, HPE*) [1] – это активно развивающаяся область, чья актуальность обусловлена прогрессом в глубоком обучении и растущим спросом со стороны прикладных сфер. В отличие от двумерной оценки, эта задача нацелена на точное предсказание пространственных координат ключевых точек тела, что даёт более полное и точное описание его положения. Такая детализация критически важна для решения прикладных задач высокого уровня, таких как дополненная реальность, биомеханический анализ и спортивная аналитика.

Постановка задачи

В рамках работы проводится анализ ключевых проблем и методов глубокого обучения в задаче трёхмерного оценивания позы человека. Для достижения поставленной цели были определены следующие задачи:

1. Сформулировать математическую постановку задачи 3D HPE.
2. Провести анализ существующих аппаратных подходов 3D HPE.
3. Изучить разновидности представлений человеческого тела
4. Классифицировать подходы для оценки позы с точки зрения на модельные и безмодельные
5. Рассмотреть разновидности безмодельного подхода
6. Исследовать наиболее эффективные архитектуры нейронных сетей, применяемые для решения данной задачи
7. Выявить основные проблемы в задаче 3D HPE

Математическая постановка задачи

Пусть у нас есть входное изображение I . Необходимо предсказать 3D-координаты $J = \{j_1, j_2, \dots, j_N\}$ из N ключевых точек тела. Каждая ключевая точка j_k представлена вектором 3D-координат $j_k = (x_k, y_k, z_k)$.

$I \in R^{H \times W \times 3}$ – входное RGB изображение высотой H и шириной W . В случае видео входные данные будут представлять собой последовательность изображений I_t , где t – временной шаг.

$J = \{(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_N, y_N, z_N)\}$ – результирующий набор 3D-координат N ключевых точек, где каждая (x_k, y_k, z_k) – это 3D-координаты k -й ключевой точки.

Задача заключается в нахождении функции f , которая отображает входное изображение I на 3D-позу J : $f: I \rightarrow J$ или, в случае с видео: $f: \{I_t\} \rightarrow \{J_t\}$. Эта функция f обычно параметризуется нейронной сетью с весами θ . Таким образом, наша задача – оптимизировать θ так, чтобы предсказанные позы \hat{J} были максимально близки к истинным позам J_{gt} .

Многие подходы сначала предсказывают 2D-позу, а затем используют ее для вывода 3D-позы. В этом случае функция f может быть декомпозирована на две подфункции:

$$f_{2D}: I \rightarrow P_{2D}, \text{ где } P_{2D} \text{ – 2D-координаты ключевых точек.}$$

$$f_{3D}: P_{2D} \rightarrow J$$

Некоторые методы используют параметрические модели тела, такие как SMPL (*Skinned Multi-Person Linear model*), для представления 3D-позы. В этом случае f предсказывает параметры модели (например, параметры формы и позы), а не напрямую 3D-координаты ключевых точек.

$$f: I \rightarrow (\beta, \alpha), \beta \text{ – параметры формы, } \alpha \text{ – параметры позы.}$$

Затем 3D-поза J получается из (β, α) с помощью дифференцируемого оператора SMPL.

Аппаратные подходы

HPE определяет положение тела и суставов человека по изображениям или видео. Существуют два основных подхода: маркерный и безмаркерный [2].

Маркерная система использует костюм с датчиками, данные с которых фиксируются камерами и преобразуются в 3D-модель. Этот подход обеспечивает высокую точность, но требует дорогого оборудования, ограничивает движения и уязвим к повреждениям датчиков.

Безмаркерная технология основана на компьютерном зрении и использует обычные или специализированные камеры. Для получения 3D-позы можно применять многокамерные системы с последующей триангуляцией или RGB-D камеры, предоставляющие данные о глубине. Однако движения, захваченные с одной точки, часто зашумлены, что требует дополнительной обработки методами машинного обучения.

Альтернативные технологии, такие как RADAR и LIDAR, хорошо справляются с окклюзией и обеспечивают высокое разрешение, но сталкиваются с проблемами разрозненности и разреженности данных, а также являются дорогостоящими. Инфракрасные датчики, например Kinect, подвержены помехам от солнечного света.

Наиболее современным и сложным подходом является использование одной RGB-камеры без данных о глубине, что стало возможным благодаря нейронным сетям. Его ключевое преимущество – масштабируемость и низкая стоимость, однако он сталкивается с ограничениями в виде окклюзии и неопределенности глубины.

Учитывая вышеперечисленное, основное внимание в исследованиях уделяется камерным решениям, поскольку именно они позиционируются как наиболее перспективные и передовые, в то время как работы на основе RADAR, LIDAR или ИК-детекторов сравнительно менее продвинуты.

Представление человеческого тела

Моделирование человеческого тела является ключевым компонентом HPE [3]. Человеческое тело представляет собой сложный нежесткий объект, обладающий множеством характеристик, таких как кинематическая структура, текстура поверхности, форма и положение суставов. При этом модель не обязательно должна включать все атрибуты тела. В зависимости от сценариев применения в HPE обычно используются скелетные, контурные и объемные модели.

Скелетная модель включает положения суставов и ориентации конечностей, представляя структуру тела. Она интуитивно понятна и успешно применяется в 2D и 3D HPE, обеспечивая гибкое графическое представление, хотя и ограничена в передаче информации о текстуре и форме.

Контурная модель отражает взаимосвязи между частями тела, представляя его форму и внешний вид. В ней части тела аппроксимируются прямоугольниками, соответствующими контурам тела.

Объемная модель используется для 3D-оценки позы, предоставляя точные данные о форме и текстуре с помощью сканирования всего тела, что позволяет получать сетки тела различных форм и поз.

Генеративный и дискриминативный подход

До того, как методы, основанные на обучении, заняли основное место, генеративные модели были популярны и доминировали в 3D HPE [4]. Генеративные методы строят свои решения на основе моделей человеческого тела, тогда как дискрими-

нативные подходы ориентированы на нахождение соответствий между наблюдаемыми данными и позой человека без использования явной модели тела. Таким образом, их рабочие направления являются полностью противоположными.

Генеративные подходы к оценке позы основаны на использовании априорных знаний, заложенных в параметрических моделях человеческого тела. Восстановление позы осуществляется либо инвертированием кинематики из двумерных ключевых точек, либо поиском в пространстве конфигураций, который обычно формулируется детерминированно как задача нелинейной оптимизации или максимизации правдоподобия. Данный процесс включает два этапа: на этапе моделирования формируется функция правдоподобия, интегрирующая признаки изображения, кинематику модели, параметры камеры и анатомические ограничения; на этапе оценки определяется набор параметров позы, наилучшим образом согласующийся с наблюдениями. Хотя эти методы не требуют обучения на парных 3D-данных и обеспечивают высокую точность и физическую состоятельность результатов, их главным ограничением является зависимость от точной начальной инициализации и склонность к сходимости к локальным минимумам. Тем не менее, они демонстрируют высокую обобщающую способность к сложным позам, вариациям одежды и аксессуаров.

В противоположность этому, дискриминативные подходы (безмодельные или основанные на обучении), не используют явные параметрические модели тела и обычно реализуются с помощью машинного обучения. Вместо этого они изучают функцию отображения непосредственно из пространства изображений в пространство поз. Они не требуют инициализации и итеративной оптимизации, что делает их чрезвычайно быстрыми на этапе тестирования. После обучения процесс вывода сводится к вычислительной задаче прямого прогноза или ограниченного поиска, а не к оптимизации в многомерном параметрическом пространстве. Однако эта скорость достигается ценой зависимости от больших объёмов данных с трёхмерными аннотациями, которые зачастую труднодоступны. Кроме того, поскольку многие из этих методов моделируют суставы как независимые точки, они могут принимать физически некорректные позы, например, с нереалистичными углами вращения в суставах или плавающей длиной костей. Как следствие, дискриминативные методы демонстрируют высокую производительность в пределах распределения обучающих данных, но могут быть менее надёжными при столкновении с новыми, нетипичными позами. Но к настоящему моменту генеративные подходы чаще применяются в составе гибридных методов

В попытке преодолеть ограничения каждого из подходов были разработаны гибридные методы, которые комбинируют сильные стороны обеих парадигм. Типичная гибридная схема использует быстрый дискриминативный метод для получения грубой начальной оценки позы. Эта оценка служит качественной инициализацией для последующего итеративного процесса оптимизации на основе генеративной модели. В таком сценарии правдоподобие наблюдения, вычисляемое генеративной моделью, используется для проверки и уточнения гипотез, предложенных дискриминативной моделью. Другой способ интегрирует нейронные сети для прямой регрессии параметров модели, минуя этап явного обнаружения ключевых точек. Например, использование CNN, для прямого предсказания параметров модели SMPL [5], после чего эта модель проецируется на изображение, и функция потерь, оценивающая согласованность в двумерном пространстве, минимизируется для тонкой настройки параметров. Таким образом, гибридные подходы смягчают проблему низкой эффективности генеративных методов и компенсируют слабую обобщающую способность дискриминативных методов, обеспечивая баланс между скоростью, точностью и физической достоверностью итоговой позы.

Таким образом, современные исследования в области 3D НРЕ в значительной степени сфокусированы на дискриминативных подходах, чей прогресс был катализирован развитием глубокого обучения и благодаря которым также продвинулись гибридные подходы. Далее будут подробно рассмотрены ключевые разновидности дискриминативных методов.

Разновидности дискриминативного подхода

Оценка трёхмерной позы одного человека, в рамках дискриминативного подхода, может быть условно разделена на два основных подхода: прямую регрессию [6] и подъём (lifting) из 2D в 3D [7]. Кроме того, отдельного внимания заслуживает класс методов, основанных на тепловых картах [8].

Прямая регрессия предполагает непосредственное предсказание трёхмерных координат суставов по входным изображениям или видеопоследовательностям с помощью единой нейросетевой модели. Несмотря на концептуальную простоту, такой подход страдает от фундаментальной проблемы неоднозначности глубины: из одного 2D-изображения невозможно однозначно восстановить 3D-структуру сцены без дополнительных предположений или контекста. Это ограничивает точность и обобщающую способность прямых методов, особенно в условиях сложных поз или частичных окклюзий.

В отличие от прямой регрессии, подходы на основе подъёма (2D-to-3D lifting) разделяют задачу на два этапа. На первом этапе с использованием хорошо зарекомендовавших себя 2D-детекторов поз – таких как OpenPose [9], DeepCut [10] или BlazePose [11] – из входного изображения извлекается двумерный скелет человека. На втором этапе специализированная модель преобразует полученную 2D-позу в её трёхмерный аналог. Такой декомпозиционный подход позволяет эффективно использовать достижения в области 2D-оценки поз и значительно смягчает проблему неоднозначности глубины за счёт явного моделирования геометрических и кинематических ограничений человеческого тела. Эмпирически показано, что методы подъёма обеспечивают существенно более высокую точность, чем прямые регрессионные модели. Более того, использование временных последовательностей 2D-скелетов в качестве входа для лифтинга позволяет учитывать динамику движения и дополнительно улучшает качество 3D-реконструкции по сравнению с обработкой отдельных кадров.

Отдельную категорию составляют методы на основе тепловых карт. В таких подходах нейронная сеть не предсказывает координаты напрямую, а генерирует для каждой ключевой точки трёхмерную тепловую карту, обычно моделируемую как трёхмерное гауссово распределение с центром в истинном положении сустава. Окончательная 3D-поза восстанавливается на этапе постобработки путём поиска локальных максимумов на этих картах. Такой подход обеспечивает более устойчивую оценку позы за счёт мягкого представления пространственной неопределённости и часто используется в гибридных архитектурах, сочетающих преимущества регрессии и тепловых карт.

Таким образом, несмотря на теоретическую привлекательность прямых методов, современные решения в области 3D-оценки позы человека преимущественно опираются на двухэтапные стратегии с использованием 2D-детекторов и последующего подъёма, а также на представления в виде тепловых карт для повышения точности.

Архитектуры

Графовые сверточные сети (*Graph Convolutional Networks, GCN*) [12,] [13] продемонстрировали высокую эффективность в задачах 3D-НРЕ, особенно в двухэтапных подходах, где на первом этапе из изображения извлекаются двумерные ключевые точки, а на втором – по ним восстанавливается трёхмерная поза. В этом контексте GCN позволяют явно моделировать пространственные зависимости между суставами, используя графовую структуру скелета для распространения информации между соседними узлами. Благодаря этому модель способна учитывать биомеханические ограничения, такие как фиксированная длина конечностей, и эффективно разрешать неоднозначности, вызванные окклюзиями или отсутствием глубины в 2D-наблюдениях. Однако стандартные GCN обладают принципиальным ограничением: их агрегация информации локальна и, как правило, ограничивается непосредственными соседями в графе. Это затрудняет моделирование долгосрочных взаимодействий между анатомически удалёнными суставами (например, между локтем и коленом), что может снижать точность восстановления сложных или нестандартных поз.

В то же время трансформеры, изначально разработанные для задач обработки естественного языка, показали выдающиеся результаты в компьютерном зрении благодаря своей способности моделировать глобальные зависимости через механизм внутреннего внимания (*self-attention*) [14], [15]. В контексте 3D-НРЕ трансформеры позволяют каждому суставу напрямую взаимодействовать со всеми остальными, независимо от их топологической близости в скелете. Это особенно ценно при наличии частичных наблюдений: даже если часть суставов скрыта, модель может корректно предсказать их 3D-положение, опираясь на глобальный контекст и изученные анатомические корреляции. Механизм внимания динамически взвешивает вклад каждого сустава в предсказание других, обеспечивая гибкое и семантически осмысленное представление пространственных отношений.

Однако трансформерные подходы к моделированию скелетных данных имеют существенные ограничения. Во-первых, при представлении скелета в виде плоской последовательности они игнорируют врождённую графовую топологию тела, что заставляет модель неявно восстанавливать анатомические связи из данных и снижает её обобщающую способность. Во-вторых, глобальный характер механизма внимания часто приводит к недостаточному учёту локальных временных закономерностей – таких как плавность траекторий суставов или краткосрочные кинематические зависимости, – что критично для точного восстановления динамики движений

Основные проблемы в задаче 3D НРЕ

Неопределенность глубины – самая серьёзная проблема. Одно двумерное изображение не имеет прямых признаков глубины, а это означает, что несколько различных 3D-поз могут проецироваться на один и тот же двумерный силуэт.

Эмпирически доказано, что добавление канала глубины к RGB-изображению повышает точность модели. Карту глубины можно получить из монокулярного RGB-изображения с помощью специализированных нейросетей (например, MiDaS [16] и Qwen Image [17]), создав таким образом RGB-D вход для последующего восстановления 3D-позы.

Части человеческого тела могут быть скрыты одеждой, другими объектами сцены или даже другими частями тела (самоокклюзия). Эта частичная видимость затрудняет точную оценку полной трёхмерной позы. Кроме того, большое разнообразие форм и размеров человеческого тела ещё больше усложняет задачу, поскольку модели должны быть обобщены для различных анатомических особенностей. Одним из решений проблемы окклюзии является использование набора данных, собранного

на основе многоракурсного подхода. В такой системе поза, невидимая с одного угла обзора, с высокой вероятностью будет видна с другого, что позволяет обеспечить более надежные результаты за счет межракурсного вывода. Кроме того, извлечению дополнительной информации из изображения способствуют гибридные архитектуры, сочетающие графовые сверточные сети (GCN) и трансформеры. GCN эффективно кодируют локальную информацию о суставах скелета, в то время как трансформеры компенсируют ограниченность локального поля восприятия, обеспечивая глобальное понимание позы. Современные подходы стремятся к комплексному подходу [18-20], где GCN используются для начального кодирования топологических зависимостей, а трансформеры уточняют предсказания, что повышает устойчивость к сложным сценариям.

Многие модели хорошо работают на контролируемых лабораторных наборах данных, но испытывают трудности при применении к реальным изображениям с различным фоном, освещением и ракурсами и требуют большого объема размеченных 3D-данных, что ограничивает их применимость. Для решения этой проблемы можно использовать альтернативные стратегии обучения. Один из вариантов, обучение без учителя, которое исследует внутренние закономерности в данных, используя, например, геометрическую самосогласованность или состязательное обучение для восстановления 3D-позы без 3D-разметки. Самообучение, как подвид обучения без учителя, использует специальные задачи, такие как многовидовая реконструкция или циклы проекции, для получения 3D-позы. Обучение со слабым контролем и перенос обучения позволяют эффективно использовать ограниченные или слабо размеченные данные, уменьшая зависимость от дорогостоящей полной разметки.

Дальнейшие работы

В перспективе планируется разработка системы для оцифровки трехмерной позы пользователя на основе изображения с одной RGB-камеры и последующего переноса полученной позы на модель виртуального аватара. Ключевыми требованиями к системе являются высокая скорость обработки и точность определения позы, что обусловлено необходимостью работы в реальном времени. Кроме того, система должна быть универсальной и работать исключительно с одной стандартной RGB-камерой, без необходимости использования дополнительных датчиков, нескольких камер или специализированного оборудования.

Среди рассмотренных подходов – генеративный, гибридный и дискриминативный – предпочтение отдано последнему, из-за меньшей вычислительной сложности. Для компенсации возможного снижения точности будет использоваться метод подъема 2D-позы в 3D (2D-to-3D lifting). Эмпирически доказано, что этот подход превосходит по точности прямую регрессию, а также обеспечивает модульность системы: 2D-детектор позы может быть заменен без необходимости модификации модуля реконструкции 3D-позы.

В качестве базовой архитектуры предлагается гибридное решение, сочетающее графовые сверточные сети (GCN) и трансформеры. Для обеспечения высокой производительности планируется применение ряда оптимизаций, таких как квантование весов, использование формата ONNX и выполнение вычислений на графических ускорителях (GPU). Это позволит сохранить высокую скорость работы при увеличении объема извлекаемой релевантной информации.

Для дальнейшего повышения точности оценки позы в входные данные планируется добавить карту глубины в качестве дополнительного канала изображения. Карта глубины будет извлекаться из исходного RGB-изображения с помощью отдельно обученной нейронной сети.

Выводы

В работе приведено математическое описание задачи 3D НРЕ: описаны входные (для изображения и видео), выходные данные для прямой регрессии, подъема и SMPL задачи. Было рассмотрено три подхода в задаче 3D НРЕ: 1) генеративный (с использованием модели); 2) дискриминативный (без использования модели); 3) гибридный. Можно сделать вывод, что актуальными являются гибридный и дискриминативный подход, а генеративный сейчас чаще используется в составе гибридного. Были рассмотрены разновидности дискриминативного подхода: прямая регрессия, подъем из 2D в 3D и тепловые карты, из которых наиболее точным является второй подход. Также описаны две наиболее эффективные архитектуры нейронных сетей для задачи 3D НРЕ: трансформер, графовая сеть и основанный на них гибридный подход.

Рассмотрены основные проблемы 3D НРЕ и предложены решения некоторых проблем. 1) Поскольку один силуэт может соответствовать разным 3D-конфигурациям, это создает неопределенность глубины. Решением служит использование RGB-D данных, где карта глубины генерируется нейросетями. 2) Скрытие частей тела одеждой или объектами, а также анатомическое разнообразие людей. Для борьбы с этим применяют многоракурсные системы, где поза, невидимая с одного ракурса, видна с другого. Также эффективны гибридные архитектуры, сочетающие графовые сверточные сети (GCN) для анализа локальных связей и трансформеры для глобального контекста. 3) Модели, обученные на лабораторных данных, часто плохо обобщаются на реальные условия. Чтобы снизить зависимость от дорогой 3D-разметки, используют альтернативные стратегии: обучение без учителя (геометрическая самосогласованность, составительные сети), самообучение, слабый контроль и перенос обучения.

В результате можно сделать следующие выводы. 1) Генеративный подход позволяет обеспечить анатомическую достоверность позы и не требует обучающих данных, но вычислительно затратен и требует наличие точной трехмерной модели. Дискриминативный – его полная противоположность. 2) Подъем из 2D в 3D также обеспечивает модульность системы, позволяя заменять 2D-детектор при необходимости на другой, без замены всей остальной части системы. 3) Достичь высокой скорости обработки можно и без упрощения системы или замены подхода. Для этого можно воспользоваться квантованием весов нейронной сети и ее экспортом в формат ONNX.

Список литературы

1. El Kaid A., Baïna K. A Systematic Review of Recent Deep Learning Approaches for 3D Human Pose Estimation / A. El Kaid, K. Baïna // *Journal of Imaging*. – 2023. – P. 8-15. – DOI: 10.3390/jimaging9120275
2. Киселев Ю.В., Богомолов И.А., Розалиев В.Л., Баклан В.А. Анализ подходов, методов и решений для детектирования позы человека. выбор инструмента для задачи определения эмоционального состояния человека по его позе // *Современные наукоемкие технологии*. – 2023. – №6. – С. 41-47. – DOI: <https://doi.org/10.17513/snt.39629>
3. Шергин И.А., Рыжов А.П. Проблема оценки позы человека: задачи, методы, решения. // *Интеллектуальные системы. Теория и приложения*. – 2024. – С. 69-84.
4. Gong W., Zhang X., González J., Sobral A., Bouwmans T., Tu C., Zahzah E. Human Pose Estimation from Monocular Images: A Comprehensive Survey / W. Gong, X. Zhang, J. González, A. Sobral, T. Bouwmans, C. Tu // *Sensors*. – 2016. – P. 13-18. – DOI: 10.3390/s16121966
5. Loper M., Mahmood N., Romero J., Pons-Moll G., Black M. J. SMPL: A Skinned Multi-Person Linear Model / M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, M. J. Black. – 2015. – P.1-10. – DOI: 10.1145/3596711.3596800
6. Zheng C., Wu W., Chen C., Yang T., Zhu S., Shen J., Kehtarnavaz N., Shah M. Deep Learning-Based Human Pose Estimation: A Survey – 2023. – URL: <https://arxiv.org/pdf/2012.13392> (дата обращения 15.09.2025).
7. Zhou L., Meng X., Liu Z., Wu M., Gao Z., Wang P. Human Pose-based Estimation, Tracking and Action Recognition with Deep Learning: A Survey / L. Zhou, X. Meng, Z. Liu, M. Wu, Z. Gao, P. Wang – 2023. – P. 4-10. – DOI: arXiv.2310.13039v1

8. Knap P. Human Modelling and Pose Estimation Overview / P. Knap . – 2024. – P. 1-4. – DOI: arXiv. 2406.19290v1
9. Cao Z., Hidalgo G., Simon T., Wei S., Sheikh Y. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields / Z. Cao, G. Hidalgo, T. Simon, S. Wei, Y. Sheikh. – 2017. – P. 1-5. – DOI: arXiv:1812.08008
10. Pishchulin L., Insafutdinov E., Tang S., Andres B., Andriluka M., Gehler P., Schiele B. DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation / L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. Gehler, B. Schiele. – 2015. – P.1-15. – DOI: arXiv:1511.06645
11. Bazarevsky V., Grishchenko I., Raveendran K., Zhu T., Zhang F., Grundmann M. BlazePose: On-device Real-time Body Pose tracking / V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, M. Grundmann. – 2020. – P. 1-4. – DOI: arXiv:2006.10204
12. Shahjahan M., Hamza B. Flexible graph convolutional network for 3D human pose estimation / M. Shahjahan, B. Hamza. – 2024. – P. 1-7. – DOI: arXiv:2407.19077
13. Azizi N., Possegger H., Rodolà E., Bischof H. 3D Human Pose Estimation Using Möbius Graph Convolutional Networks / N. Azizi, H. Possegger, E. Rodolà. – 2022. – P. 1-13. – DOI: arXiv:2203.10554
14. Yang Z., Loo J. PyCAT4: A Hierarchical Vision Transformer-based Framework for 3D Human Pose Estimation – 2025. – URL: <https://arxiv.org/pdf/2508.02806> (дата обращения 10.09.2025)
15. Zheng C., Zhu S., Mendieta M., Yang T., Chen C., Ding Z. 3D Human Pose Estimation with Spatial and Temporal Transformers / C. Zheng, S. Zhu, M. Mendieta, T. Yang, C. Chen, Z. Ding. – 2021. – P. 1-13. – DOI: arXiv:2103.10455v3
16. Birkl R., Wofk D., Muller M. MiDaS v3.1 – A Model Zoo for Robust Monocular Relative Depth Estimation – 2023. – URL: <https://arxiv.org/pdf/2307.14460> (дата обращения 13.09.2025).
17. Wu C., Li J., Zhou J., Lin J., Gao K., Yan K. Qwen-Image Technical Report / C. Wu, J. Li, J. Zhou, J. Lin, K. Gao, K. Yan. – 2025. – P.1-23. – DOI: arXiv:2508.02324
18. Aouaidjiaa K., Lia A., Zhanga W., Zhanga C., 3D Human Pose Estimation via Spatial Graph Order Attention and Temporal Body Aware Transformer / K. Aouaidjiaa, A. Lia, W. Zhanga, C. Zhanga. – 2025. – P. 1-16. – DOI: arXiv.2505.01003v1
19. Mehraban S., Adeli V., Taati B. MotionAGFormer: Enhancing 3D Human Pose Estimation with a Transformer-GCNFormer Network / S. Mehraban, V. Adeli, B. Taati. – 2023. – P. 1-11. – DOI: arXiv:2310.16288
20. Fu Y., Huang C., Li J., Kong H., Tian Y., Li H., Zhang Z. HDiffTG: A Lightweight Hybrid Diffusion-Transformer-GCN Architecture for 3D Human Pose Estimation / Y. Fu, C. Huang, J. Li, H. Kong, Y. Tian, H. Li, Z. Zhang. – 2025. – P. 1-8. – DOI: arXiv:2505.04276

References

1. El Kaid A., Baïna K. A Systematic Review of Recent Deep Learning Approaches for 3D Human Pose Estimation / A. El Kaid, K. Baïna // Journal of Imaging. – 2023. – P. 8-15. – DOI: 10.3390/jimaging9120275
2. Kiselev Yu.V., Bogomolov I.A., Rozaliev V.L., Baklan V.A. ANALYSIS OF APPROACHES, METHODS AND SOLUTIONS FOR DETECTING HUMAN POSTURE. CHOOSING A TOOL FOR THE TASK OF DETERMINING A PERSON'S EMOTIONAL STATE BY THEIR POSTURE // Modern high-tech technologies. – 2023. – No. 6. – pp. 41-47. – DOI: <https://doi.org/10.17513/snt.39629>
3. Shergin I.A., Ryzhov A.P. The problem of assessing human posture: tasks, methods, solutions. // Intelligent systems. Theory and applications. – 2024. – Pp. 69-84.
4. Gong W., Zhang X., González J., Sobral A., Bouwmans T., Tu C., Zahzah E. Human Pose Estimation from Monocular Images: A Comprehensive Survey / W. Gong, X. Zhang, J. González, A. Sobral, T. Bouwmans, C. Tu // Sensors . – 2016. – P. 13-18. – DOI: 10.3390/s16121966
5. Loper M., Mahmood N., Romero J., Pons-Moll G., Black M. J. SMPL: A Skinned Multi-Person Linear Model / M. Loper, N. Mahmood, J. Romero, G. Pons-Moll, M. J. Black. – 2015. – P.1-10. – DOI: 10.1145/3596711.3596800
6. Zheng C., Wu W., Chen C., Yang T., Zhu S., Shen J., Kehtarnavaz N., Shah M. Deep Learning-Based Human Pose Estimation: A Survey – 2023. – URL: <https://arxiv.org/pdf/2012.13392> (accessed 15.09.2025).
7. Zhou L., Meng X., Liu Z., Wu M., Gao Z., Wang P. Human Pose-based Estimation, Tracking and Action Recognition with Deep Learning: A Survey / L. Zhou, X. Meng, Z. Liu, M. Wu, Z. Gao, P. Wang – 2023. – P. 4-10. – DOI: arXiv.2310.13039v1
8. Knap P. Human Modelling and Pose Estimation Overview / P. Knap . – 2024. – P. 1-4. – DOI: arXiv. 2406.19290v1
9. Cao Z., Hidalgo G., Simon T., Wei S., Sheikh Y. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields / Z. Cao, G. Hidalgo, T. Simon, S. Wei, Y. Sheikh. – 2017. – P. 1-5. – DOI: arXiv:1812.08008

10. Pishchulin L., Insafutdinov E., Tang S., Andres B., Andriluka M., Gehler P., Schiele B. DeepCut: Joint Subset Partition and Labeling for Multi Person Pose Estimation / L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. Gehler, B. Schiele. – 2015. – P.1-15. – DOI: arXiv:1511.06645
11. Bazarevsky V., Grishchenko I., Raveendran K., Zhu T., Zhang F., Grundmann M. BlazePose: On-device Real-time Body Pose tracking / V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, M. Grundmann. – 2020. – P. 1-4. – DOI: arXiv:2006.10204
12. Shahjahan M., Hamza B. Flexible graph convolutional network for 3D human pose estimation / M. Shahjahan, B. Hamza. – 2024. – P. 1-7. – DOI: arXiv:2407.19077
13. Azizi N., Possegger H., Rodolà E., Bischof H. 3D Human Pose Estimation Using Möbius Graph Convolutional Networks / N. Azizi, H. Possegger, E. Rodolà. – 2022. – P. 1-13. – DOI: arXiv:2203.10554
14. Yang Z., Loo J. PyCAT4: A Hierarchical Vision Transformer-based Framework for 3D Human Pose Estimation – 2025. – URL: <https://arxiv.org/pdf/2508.02806> (accessed 09/10/2025)
15. Zheng C., Zhu S., Mendieta M., Yang T., Chen C., Ding Z. 3D Human Pose Estimation with Spatial and Temporal Transformers / C. Zheng, S. Zhu, M. Mendieta, T. Yang, C. Chen, Z. Ding. – 2021. – P. 1-13. – DOI: arXiv:2103.10455v3
16. Birkl R., Wofk D., Muller M. MiDaS v3.1 – A Model Zoo for Robust Monocular Relative Depth Estimation – 2023. – URL: <https://arxiv.org/pdf/2307.14460> (accessed 09/13/2025).
17. Wu C., Li J., Zhou J., Lin J., Gao K., Yan K. Qwen-Image Technical Report / C. Wu, J. Li, J. Zhou, J. Lin, K. Gao, K. Yan. – 2025. – P.1-23. – DOI: arXiv:2508.02324
18. Aouaidjiaa K., Lia A., Zhanga W., Zhanga C., 3D Human Pose Estimation via Spatial Graph Order Attention and Temporal Body Aware Transformer / K. Aouaidjiaa, A. Lia, W. Zhanga, C. Zhanga. – 2025. – P. 1-16. – DOI: arXiv.2505.01003v1
19. Mehraban S., Adeli V., Taati B. MotionAGFormer: Enhancing 3D Human Pose Estimation with a Transformer-GCNFormer Network / S. Mehraban, V. Adeli, B. Taati. – 2023. – P. 1-11. – DOI: arXiv:2310.16288
20. Fu Y., Huang C., Li J., Kong H., Tian Y., Li H., Zhang Z. HDiffTG: A Lightweight Hybrid Diffusion-Transformer-GCN Architecture for 3D Human Pose Estimation / Y. Fu, C. Huang, J. Li, H. Kong, Y. Tian, H. Li, Z. Zhang. – 2025. – P. 1-8. – DOI: arXiv:2505.04276

RESUME

O. A. Lyamtsev, I. I. Maksimenko

Overview and problems of using deep learning for three-dimensional estimation of human pose from a single image

The article is devoted to the problems and directions of deep learning in the problem of three-dimensional assessment of human posture. The paper provides a mathematical description of the problem of three-dimensional assessment of posture, three approaches are considered: 1) generative; 2) discriminative; 3) hybrid. The discriminative approach is considered in more detail and three more approaches are highlighted: 1) direct regression; 2) ascent from 2D to 3D; 3) heat maps.

The paper considers the key problems of three-dimensional assessment of human posture and suggests ways to solve them. The main difficulty is the uncertainty of depth in 2D images, since different 3D poses can be projected into identical silhouettes. To increase accuracy, RGB to RGB-D conversion is used with the addition of a depth map generated by neural networks. Another problem is occlusion, when body parts are hidden, as well as the anatomical diversity of people. The solution is multi-camera systems that capture the pose from different angles, as well as hybrid architectures that combine graph convolutional networks to analyze local joint connections and transformers for a global understanding of the scene. In addition, models that are effective in the laboratory often turn out to be unsuitable for real images due to the need for a large amount of labeled 3D data. An alternative is unsupervised learning strategies that explore the internal patterns of data, as well as self-learning, weak control, and learning transfer that reduce dependence on expensive markup.

The generative method creates anatomically accurate poses, but requires a lot of calculations, whereas the discriminative method is its exact opposite. The upgrade from 2D to 3D makes the system modular, making it easy to replace the 2D detector. High processing speed can be achieved without changing the architecture, using quantization of neural network weights and export to ONNX format.

РЕЗЮМЕ

О. А. Лямцев, И. И. Максименко

Обзор и проблемы использование глубокого обучения для трехмерной оценки позы человека по одному изображению

Статья посвящена проблемам и направлениям глубокого обучения в задаче трехмерной оценки позы человека. В работе дано математическое описание задачи трехмерной оценки позы, рассмотрено три подхода: 1) генеративный; 2) дискриминативный; 3) гибридный. Рассмотрен более подробно дискриминативный подход и выделены еще три подхода: 1) прямая регрессия; 2) подъем из 2D в 3D; 3) тепловые карты.

В работе рассмотрены ключевые проблемы трёхмерной оценки позы человека и предложены пути их решения. Основной сложностью является неопределенность глубины на 2D-изображениях, поскольку разные 3D-позы могут проецироваться в идентичные силуэты. Для повышения точности применяется преобразование RGB в RGB-D с добавлением карты глубины, генерируемой нейросетями. Другая проблема – окклюзия, когда части тела скрыты, а также анатомическое разнообразие людей. Решением служат многокамерные системы, фиксирующие позу с разных ракурсов, а также гибридные архитектуры, сочетающие графовые сверточные сети для анализа локальных связей суставов и трансформеры для глобального понимания сцены. Кроме того, модели, эффективные в лабораторных условиях, часто оказываются непригодными для реальных изображений из-за необходимости в большом объеме размеченных 3D-данных. Альтернативой являются стратегии обучения без учителя, исследующие внутренние закономерности данных, а также самообучение, слабый контроль и перенос обучения, снижающие зависимость от дорогостоящей разметки.

Генеративный метод создает анатомически точные позы, но требует больших вычислений, тогда как дискриминативный – его полная противоположность. Подъем из 2D в 3D делает систему модульной, позволяя легко заменять 2D-детектор. Высокую скорость обработки можно достичь без изменения архитектуры, используя квантование весов нейросети и экспорт в формат ONNX.

Лямцев Олег Алексеевич – стажер-исследователь ФГБНУ "Институт прикладной математики и механики", ул. Розы Люксембург, 74, Донецк, 283048, Россия, бакалавр кафедры компьютерных технологий ФГБОУ ВО «ДонГУ», ул. Университетская, 24, Донецк, 283001, Россия, gelo2003@yandex.ru. *Область научных интересов:* компьютерное зрение, машинное обучение, нейронные сети. Число научных публикаций – 5.

Максименко Игорь Иванович – заведующий отделом теории управляющих систем ФГБНУ "Институт прикладной математики и механики", ул. Розы Люксембург, 74, Донецк, 283048, Россия, доцент кафедры компьютерных технологий физико-технического факультета ФГБОУ ВО «ДонГУ», ул. Университетская, 24, Донецк, 283001, Россия, igor.maksimenko_1967@mail.ru. *Область научных интересов:* теория автоматов, теория графов, системы искусственного интеллекта. Число научных публикаций – более 40.

Статья поступила в редакцию 22.09.2025

УДК 004.932.2

DOI 10.24412/2413-7383-2025-4-39-60-70

Е. С. Мороз, Я. С. Пикалёв

Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта», г. Донецк
283048, г. Донецк, ул. Артема, 118 б

АНАЛИЗ Sota-АРХИТЕКТУР ДЛЯ СЕМАНТИЧЕСКОЙ АУГМЕНТАЦИИ*

Ye. S. Moroz, Y.S. Pikaliyov

Federal State Budgetary Scientific Institution «Institute of Artificial Intelligence Problems»
283048, Donetsk, Artema str, 118-b.

ANALYSIS OF SOTA-ARCHITECTURES FOR SEMANTIC AUGMENTATION

В статье рассмотрены задачи семантической сегментации в контексте прикладного компьютерного зрения и связанного круга задач: обнаружение объектов, сегментация экземпляров и панорамная сегментация. Проанализированы лёгкие и высокоскоростные архитектуры реального времени, с точки зрения баланса между точностью и задержкой на наборах Cityscapes, CamVid и UAVid. Основное внимание уделено вопросам подготовки данных: аугментации, аннотированию изображений, предварительной обработке и интерпретируемости моделей.

Ключевые слова: семантическая сегментация, компьютерное зрение, обнаружение объектов, сегментация экземпляров, панорамная сегментация, аугментация данных, перекрёстная геолокализация.

The paper considers semantic segmentation tasks in the context of applied computer vision and related problems such as object detection, instance segmentation, and panoptic segmentation. Lightweight and high-speed real-time architectures are analyzed in terms of the trade-off between accuracy and latency on the Cityscapes, CamVid, and UAVid datasets. Special attention is paid to data preparation issues, including augmentation, image annotation, preprocessing, and model interpretability.

Key words: semantic segmentation, computer vision, object detection, instance segmentation, panoptic segmentation, data augmentation, cross-view geolocation.

* Статья выполнена в Федерального государственного бюджетного научного учреждения «Институт проблем искусственного интеллекта». «Лаборатории интеллектуальных систем и анализа данных»

Введение

Семантическая сегментация в последние годы приобрела особую значимость как ключевая задача компьютерного зрения, направленная на классификацию каждого пикселя изображения [1]. Этот метод позволяет системам интерпретировать сцены, выделяя объекты различных классов и формируя карту визуальной информации. С развитием глубокого обучения и ростом объёмов данных её эффективность существенно возросла, в том числе при работе с искусственными цифровыми изображениями [2].

Цель работы – оценка подхода к применению семантической сегментации аэрофотоснимков для задач перекрёстной геолокации с использованием аугментации данных. В настоящий момент семантическая сегментация применяется в автономном вождении, медицинской диагностике, робототехнике, промышленном контроле, анализе окружающей среды, а также при навигации по аэрофотоснимкам [3]. Она используется для обнаружения ключевых объектов и уточнения их пространственного положения на снимках местности, что требует специализированных методик работы с аэрофотоснимками [4] и продуманного формирования целевых наборов изображений [5]. При этом качество разметки и устойчивость моделей существенно зависят от выбора методов аугментации, расширяющих обучающие выборки и повышающих вариативность данных [6].

Постановка задачи исследования заключается в том, чтобы по результатам анализа рассмотренных SOTA-архитектур лёгкой семантической сегментации оценить их пригодность для обработки аугментированных наборов аэрофотоснимков (UAVid, Aescapes, iSAID, xView, Drone Tracking, SynDrone), принимая во внимание, что аугментация изменяет распределение обучающих данных и усложняет их вариативность, что может влиять как на точность и устойчивость сегментации, так и на допустимые вычислительные затраты при выборе архитектуры.

Основная часть

Для лучшего понимания роли семантической сегментации рассмотрим связанные задачи компьютерного зрения, направленные на анализ содержания изображений, а также вопросы подготовки и аннотирования данных для обучения соответствующих моделей [7].

1. Обнаружение объектов (*Object Detection*) - ключевая задача, направленная на поиск, локализацию и классификацию объектов на изображении. Цель - определить положение каждого объекта с помощью ограничивающей рамки (bounding box) и отнести его к определённому классу. Для успешного решения задачи важно обеспечить корректную предварительную обработку изображений и формирование признаков, удобных для последующего обучения [8].

2. Ориентированное обнаружение (*Oriented Object Detection*) расширяет классическую детекцию, учитывая ориентацию объектов и их форму. Вместо осево-ориентированных рамок применяются рамки с произвольным углом наклона, что позволяет точнее описывать наклонённые и продолговатые объекты. Преимущества заключаются в более точной локализации объектов произвольной ориентации и формы, а также в лучшей адаптации к реальным сценам, включая видеопоследовательности и сложные динамические сюжеты [9]. Недостатком является увеличение вычислительной сложности и необходимость специальных алгоритмов для оценки ориентации.

3. Сегментация экземпляров (*Instance Segmentation*) направлена на выделение и классификацию отдельных объектов на изображении, в отличие от семантической

сегментации, которая группирует пиксели по классам без разделения разных объектов одного типа. Она позволяет отличать отдельные объекты друг от друга, что важно для робототехники, медицинской диагностики и автономных систем. Для повышения качества таких моделей применяются расширенные техники аугментации и построения наборов данных изображений [10].

4. Семантическая сегментация (*Semantic Segmentation*) решает задачу классификации каждого пикселя изображения - результатом является полноценная карта, где каждому пикселю присвоен класс из заданного множества. Этот подход обеспечивает полное понимание сцены на уровне пикселей и находит применение там, где требуется детальный анализ: диагностика по изображениям, оценка состояния объектов и интерпретация пространственных структур, что ставит задачу интерпретируемости нейросемантических моделей в прикладных областях [11].

5. Панорамная сегментация (*Panoptic Segmentation*) классифицирует каждое пиксельное значение и разделяет объекты по экземплярам, обеспечивая полное понимание сцены, включая как фоновые классы, так и конкретные объекты. Методы панорамной сегментации, как правило, используют сложные мультизадачные архитектуры, например объединение выходов семантической и экземплярной сегментации, и предъявляют высокие требования к качеству разметки и устойчивости моделей, что подтверждается исследованиями инстанс-сегментации сложных объектов на реальных изображениях [12].

Развитие приложений, требующих обработки видео и изображений в реальном времени (например, на борту беспилотника, в автомобиле или на мобильном устройстве), стимулировало создание специализированных облегчённых и высокоскоростных архитектур сегментации. Ниже рассмотрены современные модели, ориентированные на работу в реальном времени при сохранении высокой точности.

1. BiSeNet - архитектура семантической сегментации для задач компьютерного зрения (computer vision). Решение использует двуветвевую схему: пространственный путь (spatial path) с малым шагом свёрток формирует представления высокого разрешения, сохраняя геометрию и границы, а контекстный путь (context path) с быстрой понижающей дискретизацией расширяет поле восприятия. В контекстном пути задействован модуль уточнения внимания (attention refinement module) и глобальное усреднение по пространству, что стабилизирует агрегирование признаков крупного масштаба. Слияние ветвей выполняет модуль слияния признаков (feature fusion module), который адаптивно объединяет детальные и контекстные признаки и балансирует вклад каналов. Подход реализуется на стандартных лёгких CNN-бэбконах, допускает переключение разрешения входа и сохраняет работу в реальном времени. На наборе Cityscapes для конфигурации с Xception (бэбкон) выдаёт 68,4% mIoU при 105,8 кадрах в секунду на NVIDIA Titan Xp при входе 1536x768, при полном 2048x1024 скорость ниже при сопоставимой точности.

Преимуществами BiSeNet являются: сохранение мелких деталей за счёт пространственного пути, устойчивый захват контекста при помощи контекстного пути и модуль уточнения внимания, высокая скорость инференса (вывод результатов работы обученной модели, в процессе решения задачи на новых данных) на практических разрешениях. Недостатками являются: ограничение верхней точности относительно современных трансформеров, чувствительность к настройке слияния признаков в блоке ffm, снижение качества на редких и мелких классах без доменной донастройки [13].

2. PIDNet-S. Система реализует трёхветвевую схему: P-ветвь обрабатывает высокое разрешение и сохраняет мелкие структуры, I-ветвь быстро понижает разрешение и накапливает контекст, D-ветвь выделяет границы и по карте границ направляет слияние ветвей деталей и контекста в узле, направляя объединение P и I для повышения точности на границах объектов; лёгкие блоки и взвешивание каналов уменьшают смешивание классов на стыках и сохраняют высокую частоту кадров. На наборе Cityscapes точность по метрике mIoU равен 78,6% при 93,2 кадра в секунду, на CamVid - 80,1% mIoU при 153,7 кадра в секунду, что подтверждает компромисс между скоростью и точностью для применения в реальном времени [14]. Решение ориентировано на сцены с выраженными границами и тонкими объектами, где качество контуров критично для последующей постобработки и трекинга.

Преимуществами PIDNet-S являются: высокая чёткость контуров за счёт граничного внимания, конкурентный баланс скорости и точности на городских сценах, сохранение деталей при работе в реальном времени. Недостатками являются: ограничение верхней точности относительно тяжёлых трансформерных схем, чувствительность к качеству разметки границ и настройкам слияния, снижение устойчивости на редких и мелких классах без доменной донастройки.

3. DDRNet. Сеть с ветвями высокого и низкого разрешения, между которыми выполняются многократные билатеральные слияния. Ветвь низкого разрешения ускоряет накопление контекста за счёт глубокой понижающей дискретизации, а ветвь высокого разрешения сохраняет геометрию и границы объектов. Модуль пирамидального пулинга с глубокой агрегацией (deep aggregation pyramid pooling module) расширяет эффективное поле восприятия и агрегирует многошкальный контекст на картах низкого разрешения. Билатеральные слияния и DAPPM совместно обеспечивают согласование детальных и контекстных представлений при ограниченных вычислительных затратах.

На наборе Cityscapes точность равна 77,4% mIoU при 102 кадрах в секунду на RTX 2080 Ti, на наборе CamVid - 74,7% mIoU при 230 кадрах в секунду, что отражает компромисс между точностью и скоростью [15]. Преимуществами DDRNet являются: высокая чёткость границ за счёт ветви высокого разрешения, сбор контекста с разных масштабов блоком DAPPM, конкурентная скорость инференса. Недостатками являются: ограничение верхней точности относительно тяжёлых трансформерных схем, чувствительность к балансу между ветвями, необходимость доменной донастройки при переносе на новые сцены.

4. SegFormer. Модель с иерархическим энкодером и лёгким декодером – извлекает признаки на разных масштабах и учитывает дальние связи. Энкодер делит изображение на пересекающиеся куски и обрабатывает их слоями, что позволяет отказать от позиционного кодирования без потери локальной привязки. Декодер агрегирует представления из нескольких уровней и восстанавливает карту классов без тяжёлых операций апсемплинга, что упрощает инференс.

Конфигурации, такие как V0 ориентированы на малое число параметров и низкую задержку инференса при сохранении качества на городских и аэровизуальных сценах. На наборе UAVid точность mIoU равна 66,19% при 147 кадрах в секунду, что допускает выполнение на борту при ограниченных вычислительных ресурсах [16].

Преимуществами SegFormer-V0 являются: низкая задержка инференса, компактность модели, устойчивый захват контекста и мелких деталей. Недостатками являются: ограничение верхней точности относительно крупных конфигураций SegFormer, зависимость метрик от входного разрешения и аппаратной платформы, необходимость доменной донастройки для редких и мелких классов.

5. DF1-Seg - средство семантической сегментации для задач компьютерного зрения. Решение опирается на алгоритм частичного упорядоченного отсека (partial order pruning), который применяется для поиска архитектуры декодера и формирует семейство сетей, ориентированных на работу в реальном времени. Поиск декодера упрощает процесс восстановления пространственного разрешения и снижает вычислительные затраты без усложнения апсемплинга. Бэкбоны DF обеспечивают высокую пропускную способность и совместимы со стандартными схемами обучения, что позволяет масштабировать вход и управлять компромиссом между скоростью и точностью. По результатам классификационных испытаний DF1 и DF2 достигают точности выше, чем ResNet-18 и ResNet-50, при меньшей латентности на сопоставимом оборудовании [17]. В конфигурации сегментации DF1-Seg демонстрирует работу в реальном времени как на производительных графических процессорах, так и на встраиваемых системах. На наборе Cityscapes mIoU равно 74,1% при 106,4 кадра в секунду на GTX 1080 Ti при разрешении 1024x2048, что подтверждает пригодность для поккадровой обработки городских сцен. На Jetson TX2 достигается 21,8 кадра в секунду при 1280x720 при сохранении стабильности инференса и умеренного потребления памяти. Архитектура допускает доменную донастройку на целевых данных и совместима с приёмами ускорения инференса за счёт понижения точности вычислений.

Преимуществами DF1-Seg являются: выраженный баланс скорости и точности на разных классах устройств, компактный декодер из поиска POP, низкая латентность на встраиваемых платформах, предсказуемое масштабирование по разрешению. Недостатками являются: ограничение верхней точности относительно тяжёлых трансформерных схем, чувствительность к выбору бэкбона и параметров поиска, необходимость доменной донастройки при переносе на новые сцены.

6. SFNet. Решение реализует модуль выравнивания потока (flow alignment module), который оценивает семантический поток и выравнивает многоуровневые признаки перед слиянием. Такое выравнивание корректирует расхождения между масштабами и повышает качество границ без существенного вычисления. Система поддерживает разные бэкбоны, что позволяет выбирать компромисс между глубиной и скоростью под целевое оборудование. В конфигурации с ResNet-18 на наборе Cityscapes mIoU равно 80,4% при 26 кадрах в секунду на GTX 1080 Ti при предварительном предобучении на больших уличных датасетах с пиксельной разметкой (Mapillary Vistas), без предобучения - 78,9% mIoU. При использовании DF2 в качестве бэкбона достигается 77,8% mIoU при 61 кадре в секунду и 74,5% mIoU при 121 кадре в секунду в разных режимах скорости и точности. С более глубокими бэкбонами, например ResNet-101, отмечается до 81,8% mIoU, при этом итоговая латентность зависит от стека инференса и входного разрешения [18].

Преимуществами SFNet являются: точное слияние многоуровневых признаков за счёт модуля выравнивания потока, масштабируемости по бэкбонам, конкурентный баланс скорости и точности, прирост от предобучения на уличных датасетах. Недостатками являются: чувствительность к выбору бэкбона и режиму инференса, снижение показателей без предобучения и оптимизаций, необходимость доменной донастройки при переносе на новые сцены.

7. BiSeNet v2. Решение реализует двуветвевую схему без трансформеров: ветвь детализации (detail branch) с широкими каналами и малой глубиной формирует представления высокого разрешения, а семантическая ветвь (semantic branch) с узкими каналами и глубокой иерархией накапливает контекст при понижении разрешения.

В семантической ветви вычислительная нагрузка снижается за счёт быстрой стратегии уменьшения разрешения и лёгких блоков на по-канальных свёртках (depthwise), дополненных модулем контекстного встраивания (context embedding), который стабилизирует учёт глобального контекста. Двусторонний направляемый слой агрегации (bilateral guided aggregation layer) обеспечивает направляемое объединение детальных и контекстных признаков и уменьшает смешивание классов на границах. В обучении применяется схема с дополнительными головами, которые удаляются на инференсе, что повышает точность без накладных расходов в рабочем режиме.

На наборе Cityscapes при входе 2048x1024 точность равна 72,6% mIoU при 156 кадрах в секунду на NVIDIA GTX 1080 Ti, что подтверждает пригодность для обработки видео в реальном времени [19]. Архитектура ориентирована на сцены с выраженными границами и небольшой толщиной объектов, где важны сохранение деталей и устойчивый контекст, и позволяет балансировать скорость и качество через выбор разрешения и глубины бэкбона. Преимуществами ViSeNet v2 являются: высокая скорость инференса при полном входном разрешении, отдельная обработка деталей и контекста через DB и SB, отсутствие накладных расходов от обучения на стадии работы. Недостатками являются: ограничение верхней точности относительно тяжёлых трансформерных схем, чувствительность к качеству слияния признаков на границах, необходимость доменной донастройки для редких и мелких классов.

8. FasterSeg. Архитектуру подобрали автопоиском в пространстве с ветвями разных разрешений, чтобы учитывать и детали, и общий контекст. В процессе оптимизации применяется отдельная и детализированная регуляризация задержки, которая удерживает поиск от смещения к вариантам с низкой задержкой и низкой точностью. Система поддерживает совместный поиск с дистилляцией знаний от учителя к ученику, что повышает итоговую точность без утяжеления тракта инференса. Поиск ориентируется на оценку задержки для целевого устройства, что делает итоговую конфигурацию чувствительной к профилю устройства, но обеспечивает предсказуемую скорость.

На наборе Cityscapes при разрешении 1024x2048 точность достигает 71,5% mIoU при 163,9 кадра в секунду на NVIDIA GTX 1080 Ti, что соответствует применению в реальном времени [20]. Численные значения зависят от входного разрешения, движка инференса и особенностей предобработки. Архитектура допускает масштабирование по разрешению и числу каналов и интегрируется с приёмами ускорения, при сохранении общей структуры. Преимуществами FasterSeg являются: многомасштабные ветки, найденные автопоиском, и отдельный контроль задержки по блокам модели, прирост качества от совместного поиска с переносом знаний, высокая кадровая частота на Cityscapes. Недостатками являются: зависимость результата от профилирования конкретного GPU, чувствительность к настройкам поиска и прокси латентности, ограничение верхней точности относительно тяжёлых трансформерных схем, необходимость доменной донастройки при переносе.

9. TinyHMSeg. Модель использует модуль лёгкого контекстного слияния (Lightweight Context Fusion), который выравнивает и объединяет признаки на разных масштабах перед восстановлением карты классов. Лёгкий модуль глобального усиления (Lightweight Global Enhancement module) собирает общий контекст кадра и настраивает веса каналов, формируя небольшой выходной блок сегментации без тяжёлого декодера. Структура с несколькими уровнями разрешения в бэкбоне балансирует размер карт и число каналов, что снижает вычислительные затраты при сохранении устойчивых признаков. Такая схема сети позволяет удерживать низкую задержку и предсказуемый расход памяти на дискретных и встраиваемых графических процессорах.

На наборе Cityscapes точность достигает 71,4% mIoU при 172,4 кадра в секунду на NVIDIA GTX 1080 Ti, что указывает на пригодность для потоковой обработки в реальном времени [21]. Подход применим для покадровой обработки видео с фиксацией ключевых кадров в последующей постобработке и допускает доменную донастройку на целевых данных. Преимуществами TinyHMSeg являются: высокая скорость инференса при сопоставимой точности, компактная голова сегментации за счёт LCF и LGE, стабильное масштабирование по разрешению и памяти. Недостатками являются: снижение точности на границах и мелких объектах при низком внутреннем разрешении, зависимость метрик от оптимизаций инференса и предобучения, необходимость доменной донастройки при переносе на иные сцены.

10. STDC-Seg. Решение пересматривает двуветвевую идею BiSeNet, устраняя структурную избыточность за счёт сокращения каналов в ранних слоях и упрощения тракта декодирования. Базовый блок краткосрочной плотной конкатенации (short-term dense concatenate) формирует сжатые признаки, объединяя карты признаков соседних ступеней с последующим сведением каналов. В декодере предусмотрен модуль объединения деталей, который однопоточно вводит пространственную информацию в низкоуровневые слои. Далее низкоуровневые и высокоуровневые признаки объединяются для предсказания карты классов, что обеспечивает согласование границ и контекста при малых вычислительных затратах. На наборе Cityscapes точность достигает 71,9% mIoU при 250,4 кадра в секунду на GTX 1080 Ti, при повышенном входном разрешении достигается 76,8% mIoU при 97,0 кадра в секунду, что отражает масштабируемость по качеству и скорости [22]. Преимуществами STDC-Seg являются: высокая скорость инференса на практических разрешениях, компактные представления за счёт блока STDC, согласование деталей и контекста в простом декодере. Недостатками являются: ограничение верхней точности относительно трансформерных схем, чувствительность к выбору разрешения и степени сжатия каналов, необходимость доменной донастройки для редких и мелких классов.

Заключение

Сравнительный анализ показал, что среди современных моделей реального времени нет универсального решения, и каждая архитектура занимает свою нишу в зависимости от требований к скорости, качеству и ресурсам. Сети семейства BiSeNet и BiSeNet v2 ориентированы на высокую частоту кадров при полном входном разрешении, что делает их удобными для задач, где важна интерактивность и непрерывная потоковая обработка, например для мобильных и робототехнических платформ в городских сценах. PIDNet-S и SFNet демонстрируют повышенную точность на границах и тонких объектах и более всего подходят для сценариев, где критична детализация контуров дорожной разметки, инженерных объектов и мелких элементов инфраструктуры. DDRNet и DF1-Seg обеспечивают сбалансированный компромисс между точностью и задержкой на стандартных графических процессорах и могут рассматриваться как базовые варианты для систем дорожного анализа и промышленного мониторинга, где допустим средний уровень вычислительной нагрузки при высокой стабильности работы. SegFormer-V0 показывает целесообразность использования трансформерных подходов в компактных конфигурациях и особенно интересен для аэровизуальных наборов данных и задач бортовой обработки, где требуется широкий контекст при ограниченных ресурсах. TinyHMSeg и FasterSeg ориентированы на жёсткие ограничения по памяти и времени отклика, а потому целесообразны при развёртывании на встраиваемых платформах и специализированных устройствах, в том

числе с учётом аппаратно-ориентированного поиска архитектуры. STDC-Seg демонстрирует высокую скорость при сохранении конкурентного качества и подходит для систем, где важно масштабировать разрешение под конкретный режим эксплуатации без перепроектирования всей модели.

В совокупности рассмотренные решения формируют набор типовых архитектурных шаблонов, позволяющих подбирать модель под конкретный класс устройств, характеристики данных и допустимую задержку.

Список литературы

1. Мороз Е. С. Методы семантической сегментации в компьютерном зрении: обзор архитектур, функций потерь и современных подходов // *Донецкие чтения – 2025: образование, наука, инновации, культура и вызовы современности* : материалы X Международной научной конференции (Донецк, 2025 г.). Донецк, 2025.
2. Покинтелица А. Е. Применение клеточных автоматов для создания искусственных цифровых изображений // *Искусственный интеллект: теоретические аспекты и практическое применение: материалы Донецкого Международного круглого стола*. Донецк : ФГБНУ «ИПИИ», 2025. 296 с. С. 12–18.
3. Ермоленко Т. В. К вопросу о применении глубокого обучения для задачи перекрёстной геолокализации / Т. В. Ермоленко, Р. С. Хакимов // *Проблемы искусственного интеллекта*. 2024. № 4 (35). С. 12–28.
4. Пикалёв, Я. С. Обнаружение ключевых объектов и перекрёстная геолокация: анализ наборов данных и методологические аспекты // *Проблемы искусственного интеллекта*. 2024. № 4. С. 25–37. ISSN 2413-7383.
5. Павленко, Б. В. Методика создания набора аэрофотоснимков для задачи перекрёстной геолокации / Б. В. Павленко, Я. С. Пикалёв // *Проблемы искусственного интеллекта*. 2024. № 4. С. 101–112. ISSN 2413-7383.
6. Близно М. В. Сравнение различных методов аугментации на обучение модели нейронной сети // *Искусственный интеллект: теоретические аспекты и практическое применение: материалы Донецкого Международного круглого стола*. Донецк : ФГБНУ «ИПИИ», 2025. 296 с. С. 99–102.
7. Хакимов Р. С. К вопросу о разработке системы аннотирования данных для задач компьютерного зрения / Р. С. Хакимов, О. Л. Нижникова, М. В. Близно // *Проблемы искусственного интеллекта*. 2024. № 3 (34). С. 70–79. ISSN 2413-7383. DOI 10.24412/2413-7383-2024-3-70-79.
8. Акимов А. А. Предварительная обработка данных для машинного обучения / А. А. Акимов, Д. Р. Валитов, А. И. Кубряк // *Научное обозрение. Технические науки*. 2022. № 2. С. 26–31.
9. Кришнан Ш. Р. Улучшение обнаружения аномалий на видео с помощью усовершенствованной технологии UNET и техники каскадного скользящего окна / Ш. Р. Кришнан, П. Амудха // *Информатика и автоматизация*. 2024. № 6 (23). С. 1899–1930. ISSN 2713-3192. – DOI 10.15622/ia.23.6.12.
10. Хакимов Р. С. Обзор расширенных техник аугментации для набора данных изображений / Р. С. Хакимов, Б. В. Павленко, Я. С. Пикалёв // *Донецкие чтения 2024: материалы IX Международной научной конференции (Донецк, 15–17 октября 2024 г.)*. Т. 2. Донецк : Изд-во ДонГУ, 2024. С. 272–275. ISSN 2664-7362; ISSN 2664-7370.
11. Никитенко, К. А. Интерпретируемость нейросемантических моделей при их применении в прикладных областях / К. А. Никитенко, А. В. Звягинцева. // *Проблемы искусственного интеллекта*. 2025. № 2. С. 79–90. ISSN 2413-7383.
12. Солопов, М. В. Исследование применения моделей Mask R-CNN и Segment Anything Model (SAM) для инстанс-сегментации мезенхимных стволовых клеток на микрофотографиях / М. В. Солопов, Е. С. Чечехина, А. Г. Попандопуло, А. С. Кавелина, Г. В. Акопян, В. В. Турчин // *Проблемы искусственного интеллекта*. 2025. № 2(37). С. 21–29. DOI 10.24412/2413-7383-2025-2-37-21-29. EDN ODOKND. ISSN 2413-7383.
13. Yu C. BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation / C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, N. Sang // *Proc. ECCV*. 2018. Electronic resource. Mode of access: <https://www.alphaxiv.org/abs/1808.00897>
14. Xu J. PIDNet: A Real-time Semantic Segmentation Network Inspired by PID Controllers / J. Xu, Y. Tang, K. Chen [et al.]. – arXiv preprint, 2022. – Electronic resource. – Mode of access: <https://arxiv.org/abs/2206.02066>
15. Yin H. Deep Dual-resolution Networks for Real-time and Accurate Semantic Segmentation of Road Scenes / H. Yin, X. Shen, J. Fang [et al.]. arXiv preprint, 2021. Electronic resource. Mode of access: <https://arxiv.org/pdf/2101.06085>
16. Xie E. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers / E. Xie, W. Wang, Z. Yu [et al.]. arXiv preprint, 2024. Electronic resource. – Mode of access: <https://arxiv.org/abs/2410.01092v1>

17. Li H. DFNet: Deep Feature Aggregation Network for Real-time Semantic Segmentation / H. Li, P. Xiong, J. Fan, L. Sun. – arXiv preprint, 2019. – Electronic resource. – Mode of access: <https://arxiv.org/abs/1903.03777v2>
18. Li X. Semantic Flow for Fast and Accurate Scene Parsing (SFNet) / X. Li, H. Xiong, H. Fan [et al.] // Proc. ECCV. 2020. Electronic resource. Mode of access: https://www.ecva.net/papers/eccv_2020/papers_ECCV/papers/123460749.pdf
19. Yu C. BiSeNet V2: Bilateral Network with Guided Aggregation for Real-time Semantic Segmentation / C. Yu, C. Gao, J. Wang, G. Yu, N. Sang. arXiv preprint, 2020. Electronic resource. Mode of access: <https://arxiv.org/pdf/2004.02147>
20. Chen X. FasterSeg: Searching for Faster Real-time Semantic Segmentation / X. Chen, Y. Zhu, G. Lin [et al.]. – arXiv preprint, 2019. – Electronic resource. – Mode of access: <https://arxiv.org/abs/1912.10917>
21. When Humans Meet Machines: Towards Efficient Segmentation. – [s.l.], 2023. – Electronic resource. – Mode of access: <https://scispace.com/pdf/when-humans-meet-machines-towards-efficient-segmentation-4pir2uvpjc.pdf>
22. Fan M. Rethinking BiSeNet for Real-time Semantic Segmentation (STDC-Seg) / M. Fan, T. Wang, D. Gong, J. Yang. arXiv preprint, 2021. Electronic resource. – Mode of access: <https://arxiv.org/abs/2104.13188>

References

1. Moroz YE. S. Methods of semantic segmentation in computer vision: a review of architectures, loss functions and modern approaches / E. S. Moroz // Donetsk Readings – 2025: education, science, innovation, culture and challenges of the present: materials of the 10th International scientific conference (Donetsk, 2025). – Donetsk, 2025.
2. Pokintelitsa A. E. Application of cellular automata for creating artificial digital images / A. E. Pokintelitsa // Artificial Intelligence: theoretical aspects and practical application: proceedings of the Donetsk International Round Table. – Donetsk: FSBSI “Institute for Problems of Artificial Intelligence”, 2025. – 296 p. – P. 12–18.
3. Ermolienko T. V. On the application of deep learning to the problem of cross-view geolocation / T. V. Ermolienko, R. S. Khakimov // Problems of Artificial Intelligence. – 2024. – No. 4 (35). – P. 12–28.
4. Pikalev, Ya. S. Detection of key objects and cross-view geolocation: analysis of datasets and methodological aspects / Ya. S. Pikalev. // Problems of Artificial Intelligence. – 2024. – No. 4. – P. 25–37. – ISSN 2413-7383.
5. Pavlenko, B. V.; Pikalev, Ya. S. Methodology for creating a set of aerial images for the cross-view geolocation problem / B. V. Pavlenko, Ya. S. Pikalev. // Problems of Artificial Intelligence. – 2024. – No. 4. – P. 101–112. – ISSN 2413-7383.
6. Blizno M. V. Comparison of various augmentation methods for training a neural network model / M. V. Blizno // Artificial Intelligence: theoretical aspects and practical application: proceedings of the Donetsk International Round Table. – Donetsk: FSBSI “Institute for Problems of Artificial Intelligence”, 2025. – 296 p. – P. 99–102.
7. Khakimov R. S. On the development of a data annotation system for computer vision tasks / R. S. Khakimov, O. L. Nizhnikova, M. V. Blizno // Problems of Artificial Intelligence. – 2024. – No. 3 (34). – P. 70–79. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-3-70-79.
8. Akimov A. A. Preprocessing of data for machine learning / A. A. Akimov, D. R. Valitov, A. I. Kubryak // Scientific Review. Engineering Sciences. – 2022. – No. 2. – P. 26–31.
9. Krishnan S. R. Improving anomaly detection in video using an enhanced UNET technology and a cascaded sliding-window technique / S. R. Krishnan, P. Amudha // Informatics and Automation. – 2024. – No. 6 (23). – P. 1899–1930. – ISSN 2713-3192. – DOI 10.15622/ia.23.6.12.
10. Khakimov R. S. Review of advanced augmentation techniques for image datasets / R. S. Khakimov, B. V. Pavlenko, Ya. S. Pikalev // Donetsk Readings 2024: materials of the 9th International scientific conference (Donetsk, 15–17 October 2024). – Vol. 2. – Donetsk: DonNU Press, 2024. – P. 272–275. – ISSN 2664-7362 (Print); ISSN 2664-7370.
11. Nikitenko, K. A.; Zvyagintseva, A. V. Interpretability of neurosemantic models in their application to practical domains / K. A. Nikitenko, A. V. Zvyagintseva. // Problems of Artificial Intelligence. – 2025. – No. 2. – P. 79–90. – ISSN 2413-7383.
12. Solopov, M. V.; Chechekhina, E. S.; Popandopulo, A. G.; Kavelina, A. S.; Akopyan, G. V.; Turchin, V. V. Study of the application of Mask R-CNN and Segment Anything Model (SAM) for instance segmentation of mesenchymal stem cells in microphotographs / M. V. Solopov, E. S. Chechekhina, A. G. Popandopulo,

- A. S. Kavelina, G. V. Akopyan, V. V. Turchin. // Problems of Artificial Intelligence. – 2025. – No. 2(37). – P. 21–29. – DOI: 10.24412/2413-7383-2025-2-37-21-29. – EDN: ODOKND. – ISSN 2413-7383.
13. Yu C. BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation / C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, N. Sang // Proc. ECCV. – 2018. – Electronic resource. – Mode of access: <https://www.alphaxiv.org/abs/1808.00897>
 14. Xu J. PIDNet: A Real-time Semantic Segmentation Network Inspired by PID Controllers / J. Xu, Y. Tang, K. Chen [et al.]. arXiv preprint, 2022. Electronic resource. Mode of access: <https://arxiv.org/abs/2206.02066>
 15. Yin H. Deep Dual-resolution Networks for Real-time and Accurate Semantic Segmentation of Road Scenes / H. Yin, X. Shen, J. Fang [et al.]. arXiv preprint, 2021. Electronic resource. Mode of access: <https://arxiv.org/pdf/2101.06085>
 16. Xie E. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers / E. Xie, W. Wang, Z. Yu [et al.]. arXiv preprint, 2024. Electronic resource. Mode of access: <https://arxiv.org/abs/2410.01092v1>
 17. Li H. DFNet: Deep Feature Aggregation Network for Real-time Semantic Segmentation / H. Li, P. Xiong, J. Fan, L. Sun. – arXiv preprint, 2019. – Electronic resource. – Mode of access: <https://arxiv.org/abs/1903.03777v2>
 18. Li X. Semantic Flow for Fast and Accurate Scene Parsing (SFNet) / X. Li, H. Xiong, H. Fan [et al.] // Proc. ECCV. – 2020. – Electronic resource. – Mode of access: https://www.ecva.net/papers/eccv_2020/papers_ECCV/papers/123460749.pdf
 19. Yu C. BiSeNet V2: Bilateral Network with Guided Aggregation for Real-time Semantic Segmentation / C. Yu, C. Gao, J. Wang, G. Yu, N. Sang. – arXiv preprint, 2020. – Electronic resource. – Mode of access: <https://arxiv.org/pdf/2004.02147>
 20. Chen X. FasterSeg: Searching for Faster Real-time Semantic Segmentation / X. Chen, Y. Zhu, G. Lin [et al.]. – arXiv preprint, 2019. – Electronic resource. – Mode of access: <https://arxiv.org/abs/1912.10917>
 21. When Humans Meet Machines: Towards Efficient Segmentation / [no authors]. – [s.l.], 2023. – Electronic resource. – Mode of access: <https://scispace.com/pdf/when-humans-meet-machines-towards-efficient-segmentation-4pir2uvpjc.pdf>
 22. Fan M. Rethinking BiSeNet for Real-time Semantic Segmentation (STDC-Seg) / M. Fan, T. Wang, D. Gong, J. Yang. – arXiv preprint, 2021. – Electronic resource. – Mode of access: <https://arxiv.org/abs/2104.13188>

RESUME

YE. S. Moroz, Y.S. Pikaliyov

Analysis of Sota-Architectures for Semantic Augmentation

Background: The growth of real-time computer vision tasks under strict latency and resource constraints increases the demand for lightweight semantic segmentation architectures. It is important to compare on-board and embedded models, identify common structural principles, and relate them to the full processing pipeline – from data preparation and annotation to post-processing and the use of segmentation maps in navigation and control.

Materials and methods: A review and comparative analysis of lightweight real-time semantic segmentation networks was carried out.

Results: The analysis showed that lightweight real-time semantic segmentation architectures are built according to similar schemes with separate detail and context branches, multiscale feature fusion, and simplified decoders.

Conclusion: The proposed models can be used when choosing the architecture of segmentation networks in applied computer vision systems for a given class of devices and operating mode.

РЕЗЮМЕ

Е. С. Мороз, Я. С. Пикалёв

Анализ Sota-архитектур для семантической аугментации

Актуальность: рост задач компьютерного зрения реального времени при жёстких ограничениях по задержке и ресурсам усиливает требование к лёгким архитектурам семантической сегментации; важно сопоставлять бортовые и встраиваемые модели, выделять общие структурные принципы и увязывать их с полным циклом обработки - от подготовки и аннотирования данных до постобработки и использования карт сегментации в навигации и контроле.

Материалы и методы: выполнен обзор и сравнительный анализ по лёгким сетям семантической сегментации реального времени;

Результаты: анализ показал, что лёгкие архитектуры семантической сегментации реального времени строятся по сходным схемам с разделением ветвей деталей и контекста, многошкальным объединением признаков и упрощёнными декодерами.

Вывод: предложенные модели могут использоваться при выборе архитектуры сегментационных сетей в прикладных системах компьютерного зрения под заданный класс устройств и режим работы.

Мороз Егор Сергеевич – инженер-исследователь, Федеральное государственное бюджетное научное учреждение, «Институт проблем искусственного интеллекта».

Область научных интересов: анализ данных, распознавание образов, компьютерное зрение, эл. почта azidaan.moroz@yandex.com, адрес: г. Донецк, ул. Егорова, д. 16, кв. 1, телефон: +7(949) 381-67-00.

Пикалёв Ярослав Сергеевич – кандидат техн. наук, старший научный сотрудник лаборатории интеллектуальных систем и анализа данных Федерального государственного бюджетного научного учреждения «Институт проблем искусственного интеллекта»,

Область научных интересов: цифровая обработка сигналов, анализ данных, распознавание образов, обработка естественного языка, компьютерное зрение, машинное обучение, нейронные сети, г. Донецк, эл. почта: i@pikaliov.ru, адрес: 283085, ДНР, г. Донецк, ул. Отважных, д. 19, кв. 85, телефон: +7 (949) 428-73-88.

Статья поступила в редакцию 30.09.2025

Б. В. Павленко

Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта», г. Донецк
283048, г. Донецк, ул. Артема, 118 б

ПОДХОД К МУЛЬТИМОДАЛЬНОМУ ОБЪЕДИНЕНИЮ ДАННЫХ В ЗАДАЧАХ РЕГРЕССИИ И КЛАССИФИКАЦИИ НА СНИМКАХ С БПЛА

B.V. Pavlenko

Federal State Budgetary Scientific Institution «Institute of Artificial Intelligence Problems»
283048, Donetsk, Artema str, 118-b

AN APPROACH TO MULTIMODAL DATA FUSION IN TASKS OF REGRESSION AND CLASSIFICATION IN DRONE IMAGES

В данной работе предлагается метод мультимодального объединения признаков с использованием изображений с беспилотной воздушной платформы (БПЛА) и закодированных метаданных. Рассматривается применимость предлагаемого подхода к задачам регрессии значений положения БПЛА и классификации объектов на снимках с минимальным набором известных метаданных. Для регрессии введён механизм интерполяции значений высоты, позволяющий более точно связать разницу визуальных представлений снимка с его высотой и обеспечить большую устойчивость к изменениям высоты и углов поворота камеры БПЛА. Для регулирования вклада текстовой модальности использован метод, обеспечивающий случайное игнорирование текстовых признаков и уменьшающий переобучение на текстовую модальность дополнительно с регуляризацией ошибкой выравнивания признаков и коэффициентом ее влияния на общую ошибку модели. Для классификации применен тот же подход с сравнением нескольких вариантов текстовых шаблонов. Проведены эксперименты на выборках аэрофотосъёмки с вариацией высоты и углов ориентации камеры, которые показали, что предложенный подход превосходит простую конкатенацию признаков, подтверждены выдвинутые гипотезы об улучшении обобщающей способности модели с применением рассматриваемых методов.

Ключевые слова: мультимодальные нейронные сети, регрессия, классификация, объединение признаков, позиционирование, БПЛА.

In this paper, a method of multimodal feature combination is proposed using images from an unmanned aerial vehicle (UAV) and encoded metadata. The applicability of the proposed approach to the tasks of UAV position regression and object classification in images with a minimal set of known metadata is considered. For regression, a mechanism for interpolating altitude values is introduced, which allows for a more accurate association between the visual representation of an image and its altitude, and provides greater stability to changes in altitude and camera rotation angles of the UAV. To regulate the contribution of textual modality, a method is used that ensures random ignoring of textual features and reduces overfitting on textual modality in addition to regularization by the feature alignment error and its coefficient of influence on the overall model error. The same approach is used for classification, with the comparison of several variants of textual templates. Experiments were conducted on the following datasets.

Key words: multimodal neural networks, regression, classification, feature fusion, positioning, UAV.

Введение

Одним из ключевых вызовов в разработке автономных систем навигации для беспилотных летательных аппаратов (БПЛА) является обеспечение высокой точности оценки их положения при ограниченных вычислительных ресурсах [1], [2]. Классические методы навигации, основанные на GPS и инерциальных системах, показывают себя менее надежными в условиях сценариев ухудшения или полной потери сигнала, в том числе при полетах в городских каньонах или в условиях радиоэлектронного противодействия. Это стимулирует развитие альтернативных подходов, опирающихся на сенсорные данные с бортовых камер, методы компьютерного зрения и глубокого обучения.

В последние годы большое внимание уделяется исследованию архитектур, способных эффективно объединять визуальные признаки изображения с дополнительными источниками информации, семантическими описаниями сцены, геометками, метео-данными и прочими контекстными метаданными. Ряд работ [3-6] демонстрируют, что включение дополнительных семантических сигналов в модель приводит к росту точности предсказаний и улучшению устойчивости модели к шуму и вариативности входных данных, которая может быть выражена различиями в содержании изображаемой сцены или различиями перспективы. При этом данные результаты наблюдались не только в экспериментах в сфере компьютерного зрения для беспилотных устройств, перекрестной геолокализации, а в том числе в визуально-языковом обнаружении объектов и визуальном понимании сцены [5], [6].

В контексте использования семантических дополнений возникает задача обогащения открытых наборов данных [7], в случаях отсутствия таких метаданных, как описание местности, высота снимка, информация о положении камеры. При сборе синтетических наборов данных с использованием специализированных сред [8] полнота метаданных зависит от используемой среды. При поиске реальных наборов данных проблема неполноты метаданных может оставаться нерешенной. Использование таких адаптированных для мультимодального обучения наборов данных является важной частью экспериментов в сфере решения задач повышения автономности беспилотной навигации [9].

В данной работе демонстрируются результаты экспериментов с применением метода объединения признаков для регрессии значений положения БПЛА, и классификации объектов на снимках с дополнительной семантической и числовой информацией, передаваемой через текстовое описание. Целью исследования являлся поиск **легковесного мультимодального решения**, пригодного для использования в реальном времени на бортовых платформах с ограниченными вычислительными ресурсами.

Мы рассматриваем несколько вариантов объединения признаков для анализа их влияние на итоговую ошибку модели на целевых задачах, дополнительные методы улучшения процесса обучения, а также вариации текстовых шаблонов. Для оценки предложенного подхода проведены эксперименты с различными схемами объединения и методами регуляризации, а также исследовано влияние текстовой семантики на устойчивость модели в условиях, когда на режиме работы доступен только визуальный сигнал.

Наборы данных

Для экспериментов необходимы наборы данных, содержащих не только разнообразные снимки местности со спутника и БПЛА, но также и дополнительную семантику в виде метаданных позиционирования летательного аппарата.

GTA-UAV [10] – синтетический набор данных, созданный на основе игрового движка Grand Theft Auto V с использованием сценариев автоматического управления камерой и БПЛА. Набор данных включает в себя изображения высокого разрешения, снятые с различных высот, углов наклона и направлений полета, точные метки положения и ориентации (позиции в координатах, углы Эйлера и кватернионы), синхронизированные с каждым кадром, семантические маски и классы объектов (здания, дороги, растительность, транспорт) – что делает набор пригодным для задач регрессии, сегментации и обнаружения объектов.

VisLoc [11] – набор данных для задач визуальной локализации и регрессии положения. Набор содержит 6742 снимка с реальных бортовых камер в разнообразных сценариях полета – городские кварталы, пригороды, сельские территории, и учитывает вариации освещенности, шум сенсора, блики и другие артефакты. Высота полета представлена в диапазоне 400 – 2000 м.

UC Merced Land Use Dataset [12] – набор данных, созданный для задач классификации и анализа типов землепользования. Он состоит из 2100 изображений. Каждый снимок вручную аннотирован и отнесен к одному из 21 класса, таких как «жилая зона», «аэропорт», «ферма», «лес», «река», «дорога» и др. Изображения в наборе сбалансированы по классам (по 100 снимков каждого типа) и характеризуются разнообразием текстур, структуры и контекста сцены.

Методология

Благодаря способности трансформеров кодировать входные данные практически любой модальности, некоторые модели реализуют подход *single-stream multimodal transformer* или *unified encoder*. При таком подходе у модели отсутствуют отдельные базовые сети извлечения признаков (*backbone*) или отдельные энкодеры, данные кодируются одним энкодером, как например в ViLT [13]. Примерами иного подхода являются LXMERT [14], CLIP [15], использующие отдельные ветви кодирования для разных модальностей, с последующим применением перекрестного внимания (*Cross-Attention*) или контрастивного обучения.

В данной работе был использован подход с применением экстракторов признаков, и разделением сети на две ветви – обучающую и валидационную с двумя регрессорами. Изначально была выдвинута гипотеза о том, что внедрение дополнительных числовых данных повысит численные показатели модели в задаче регрессии, т.к. различия высоты снимков и углов наклона камеры явно отражаются в визуальном контексте изображения, а значит, можно подкрепить модель дополнительным контекстом. Отсюда следовала вторая гипотеза о применимости интерполяции значения высоты для повышения стабильности предсказаний для задачи регрессии. Она состоит из округления высоты до ближайшего шага (1) и последующим ограничением диапазона (2):

$$h_q = \text{round}\left(\frac{h}{d_{step}}\right) \cdot d_{step}, \quad (1)$$

где h – исходная высота, d_{step} – шаг интерполяции. Итоговая формула выглядит следующим образом:

$$h_{out} = \max(h_{minDs}, \min(h_{dsMax}, h_q)), \quad (2)$$

где h_{minDs} и h_{maxDs} – минимальное и максимальное значение высота в наборе данных.

Перед подачей в текстовый энкодер при обучении метаданные внедряются в текстовые шаблоны. Для регрессии, значения высоты снимка внедряются в текстовый шаблон вида «This is a drone aerial image taken from a height of h meters». Для классификации были разработаны следующие шаблоны: «The target is a <class> in an aerial photo», «An aerial photo contains a <class>», «An aerial view showing a <class>».

Для снижения затрат на обучение был введен механизм text-dropout, при котором с заданной вероятностью $p_{textDrop}$ текстовые признаки игнорировались и модель получала на вход только изображение. В соответствии с этим адаптировалась итоговая функция потерь (3) т.к. при подаче обеих модальностей к ошибке регрессии L_{reg} прибавляется ошибка выравнивания признаков L_{align} – L2-регуляризация между векторными представлениями изображения v_{img} и текста v_{text} умноженная на коэффициент влияния λ (4).

$$L = \begin{cases} L_{reg}, & p_{textDrop} = true \\ L_{reg} + \lambda \cdot L_{align}, & p_{textDrop} = false \end{cases} \quad (3)$$

где L_{reg} – комбинация Hubert Loss и Focal Loss, которая обеспечивает меньшую чувствительность к выбросам и дисбалансу примеров, а $\lambda \cdot L_{align}$:

$$L_{align} = 1 - \frac{\langle v_{img}, v_{text} \rangle}{\|v_{img}\| \cdot \|v_{text}\|}, \quad (4)$$

Архитектура

Модели состоят из двух ветвей обработки входных данных, которые работают в разных режимах в зависимости от режима обучения или теста. Каждая ветвь имеет свой экстрактор признаков для своей модальности: сверточную сеть (CNN) StripNet-small [16] для изображений и языковую модель MobileClip2-s0 [17] с частичной разморозкой, за счет чего обучаемыми остаются только слои текстовой проекции и последний блок трансформера. После экстрактора следует блок информативного усиления признаков – шея – включающая в себя блок, вдохновленный архитектурой UAV-YOLOv12 [18] для выделения релевантных признаков при сохранении низкой вычислительной стоимости (рисунок 2). Multi-Pool агрегирует разномасштабные признаки перед объединением. Такое мульти-адаптивное сжатие позволило избежать проблемы дорогих вычислений, возникшей при подаче на разные ветки и шею сырых тензоров из CNN с огромными ядрами.

В конце модель разветвляется на две головы в зависимости от режима обучения или теста. В режиме обучения используются текстовый экстрактор и объединение, результат которого подается в голову для объединенных признаков. В тестовом режиме SLM не работает и текстовые признаки не извлекаются, что существенно экономит ресурсы и дает ощутимый прирост в скорости. Общая архитектура представлена на рисунке 1.

Шея (рисунок 2) состоит из сверток, C3k2-блоков и блоков A2C2F-Attention (рисунок 3), комбинирующих пространственное и каналное внимание, позволяя модели фокусироваться на важных областях и подавлять шум в фоновых регионах. C3k2-блоки представляют собой модифицированные C3-блоки (Cross Stage Partial) в виде бутылочных горлышек с двумя свертками, где используются два сверточных ядра ($k=2$) для баланса между глубиной сети и эффективностью вычислений. Они помогают лучше интегрировать контекст и сохранить детальные признаки для мелких объектов, что критично для обнаружения на UAV-снимках.

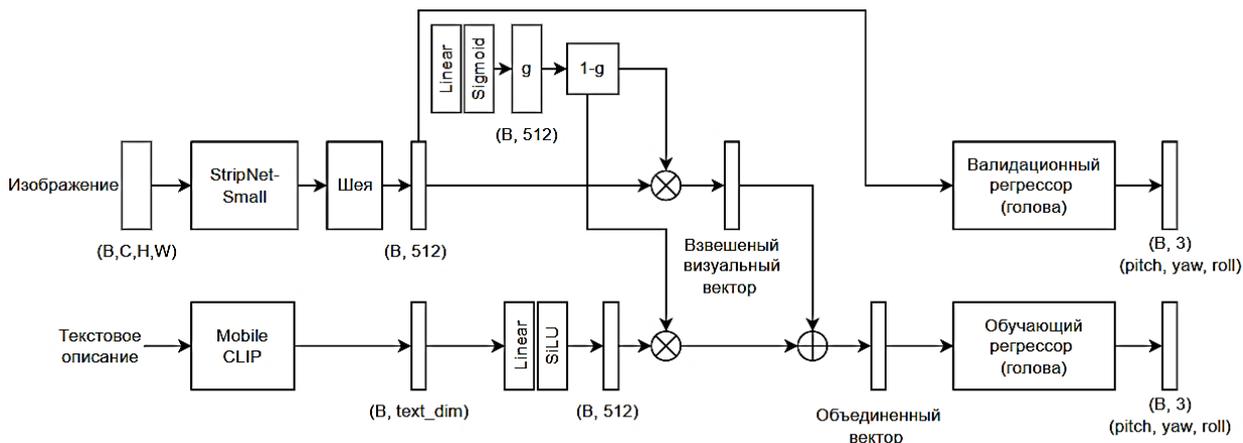


Рисунок 1 – Архитектура регрессионной модели с отдельными ветвями обработки объединенных и визуальных признаков

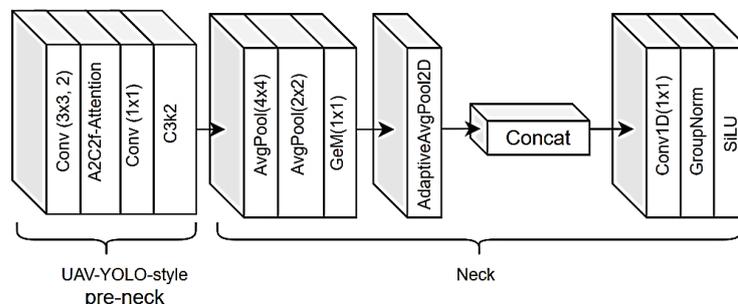


Рисунок 2 – Архитектура шеи с блоком предобработки признаков с блоками YOLO и адаптивным усредняющим сжатием разноразмерных карт признаков

Голова AngleHead (рисунок 3) принимает на вход размерности (B, D) результат объединения или признаки изображения после шеи. Так как для объединения не использовалась конкатенация, размерность выходного вектора не удваивалась, оставаясь равной 512 в зависимости от режима модели. Выход головы имеет размерность равную 3 по числу регрессионных параметров.

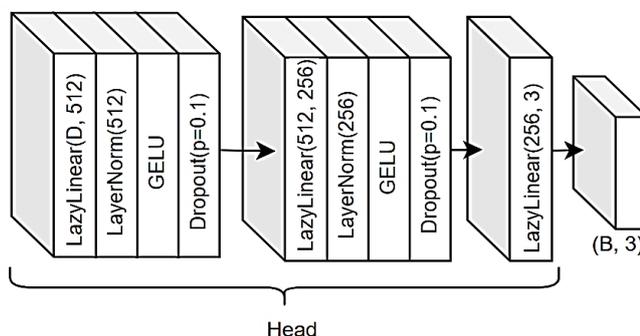


Рисунок 3 – Подробная схема головы-регрессора значений углов

Обучение

При обучении сравнивалось объединение через конкатенацию (5) или gate-сумму с коэффициентом g (6), которое происходит после шеи, что позволяет работать с компактными и семантически информативными признаками:

$$fused = concat(v_{img}, v_{text}), \quad (5)$$

$$fused = v_{img}g + v_{text}(1 - g), \quad (6)$$

Обучение модели проводилось в течение 20 эпох с $batch-size=32$, что позволило достичь стабильной сходимости без переобучения. В качестве оптимизатора использовался Ranger [19], объединяющий преимущества Rectified Adam (RAdam) [20] и Lookahead [21], что обеспечивает более плавную и надёжную оптимизацию за счёт адаптивной скорости обучения и усреднения весов по «медленным» и «быстрым» шагам. Для управления скоростью обучения применялся планировщик скорости обучения WarmupReduceLRonPlateauScheduler, выполняющий постепенный «прогрев» скорости обучения (learning rate) на начальных итерациях, а затем автоматически снижал его при стабилизации валидационной метрики, способствуя лучшему обучению и повышению финальных метрик модели. Для управления влиянием ошибки выравнивания признаков в режиме обучения использовался коэффициент λ при различных значениях $p_{textDropout}$. Для более стабильной до-настройки (fine-tuning) текстового энкодера было принято решение использовать «частичную разморозку». Для сохранения чувствительности к домену оставлен последний блок трансформера, а для адаптации векторных представлений оставлены слои проекции. Реализованная таким образом до-настройка позволяет не терять предобученные представления модели.

Эксперименты

Эксперименты проводились с использованием StripNet в качестве CNN-экстрактора на основании результатов из таблицы 1:

Таблица 1 – Сравнение показателей моделей StripNet-small и InceptionNext-tiny на наборах данных GTA-UAV и VisLoc.

Backbone	Набор данных	test loss	train loss	test MAE	train MAE
StripNet-small	GTA-UAV	0.05	0.05	0.1	0.11
InceptionNext-tiny	GTA-UAV	0.1	0.08	0.21	0.17
StripNet-small	VisLoc	0.11	0.10	0.22	0.22
InceptionNext-tiny	VisLoc	0.11	0.11	0.21	0.24

Основные эксперименты состояли в сравнении методов объединения конкатенацией (3) и суммированием с gate-коэффициентом g (4) при различных значениях $p_{textDropout}$ и интерполяции высоты с шагом равным 5 или 10. Результаты показали, что $p_{textDropout} > 0.5$ перестает давать улучшения показателей.

В первом наборе экспериментов модель обучалась на 50 эпохах для VisLoc с объединением конкатенацией и ошибками регрессии (3) (значение λ в описании к изображениям указывает на прибавление L_{align} к ошибке).

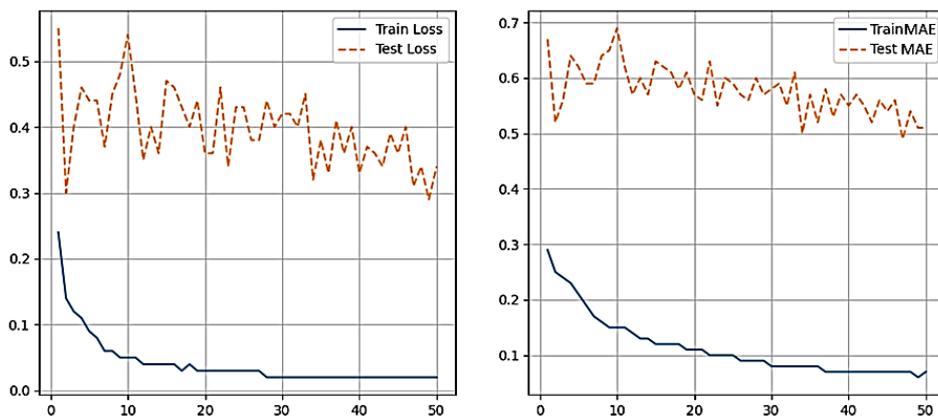


Рисунок 4 – Изменение ошибки регрессии L_{reg} и MAE без применения text-dropout в течении обучения

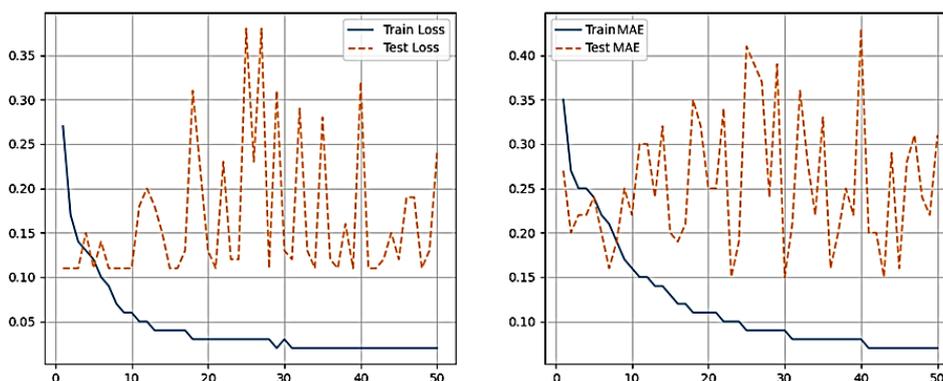


Рисунок 5 – Изменение ошибки регрессии L_{reg} и MAE с применением text-dropout с вероятностью $p_{textDropout} = 0.5$

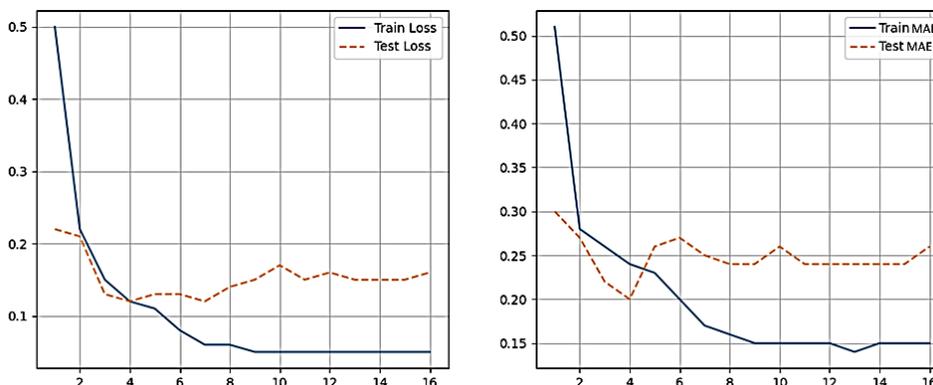


Рисунок 6 – Изменение ошибки регрессии L_{reg} и MAE с применением text-dropout с вероятностью $p_{textDropout} = 0.3$, коэффициентом влияния ошибки выравнивания признаков $\lambda = 0.1$, шагом интерполяции высоты $d_{step} = 5$

Во втором наборе экспериментов модель обучалась на 50 эпохах с объединением суммированием с весовым gate-коэффициентом и ошибкой регрессии $L_{reg} + \lambda \cdot L_{align}$ (1). Также использовался механизм интерполяции высоты со значением d_{step} равным 5 и 10.

Таблица 3 – Минимальные значения ошибки и MAE при gate-суммировании

$p_{textDropout}$	Train Loss	Train MAE	Test Loss	Test MAE	λ	d_{step}
0.2	0.03	0.10	0.11	0.19	0.5	10
0.2	0.03	0.10	0.13	0.15	0.5	10
0.3	0.02	0.08	0.11	0.17	0.1	5
0.3	0.04	0.13	0.12	0.20	0.5	5

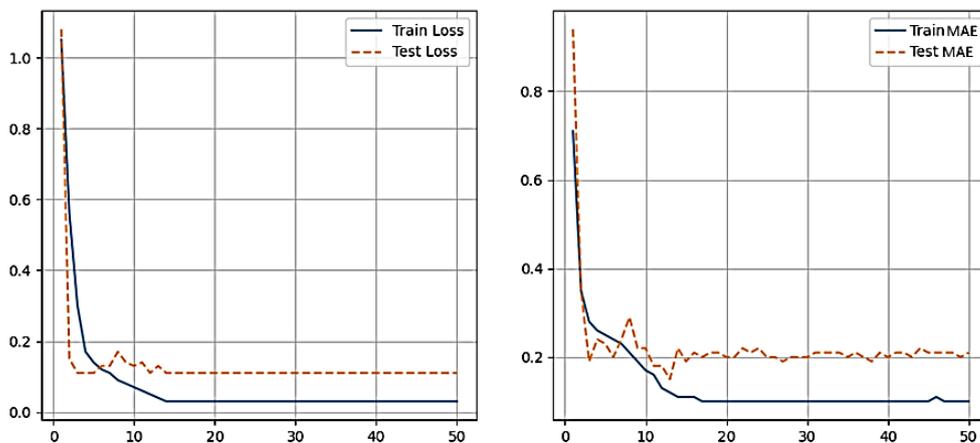


Рисунок 7 – Изменение ошибки регрессии L_{reg} и MAE с применением text-dropout с вероятностью $p_{textDropout} = 0.2$, коэффициентом влияния ошибки выравнивания признаков $\lambda = 0.5$, шагом интерполяции высоты $d_{step} = 10$

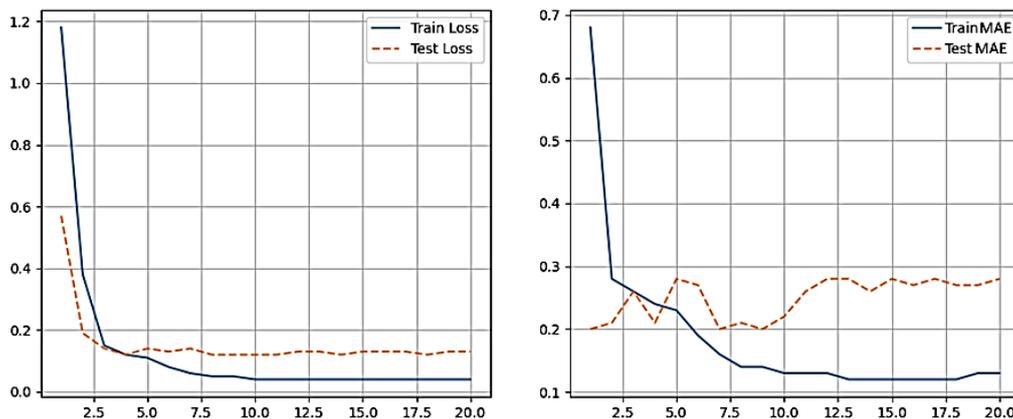


Рисунок 8 – Изменение ошибки регрессии L_{reg} и MAE с применением text-dropout с вероятностью $p_{textDropout} = 0.3$, коэффициентом влияния ошибки выравнивания признаков $\lambda = 0.5$, шагом интерполяции высоты $d_{step} = 5$

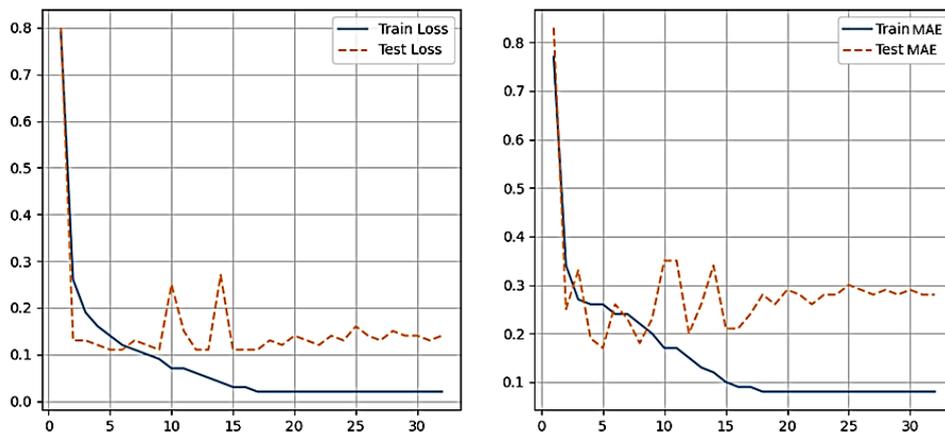


Рисунок 9 – Изменение ошибки регрессии L_{reg} и MAE с применением text-dropout с вероятностью $p_{textDropout} = 0.3$, коэффициентом влияния ошибки выравнивания признаков $\lambda = 0.1$, шагом интерполяции высоты $d_{step} = 5$

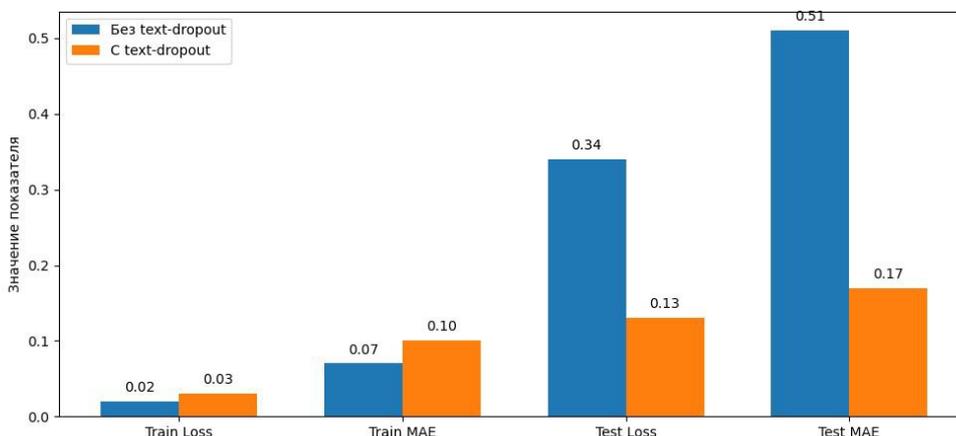


Рисунок 10 – Сравнение показателей ошибки обучения и валидации при обычном обучении и с применением метода text-dropout

Эксперименты показали, что внедрение числовых метаданных в текстовые шаблоны позволяет связать различия визуального контекста в зависимости от высоты с ее значением, что положительно отразилось на значениях ошибки при обучении. Однако, из-за отсутствия ветки текстовых признаков в тестовом режиме, ошибка была крайне нестабильна.

Для достижения хороших показателей были применены такие методики, как случайное обнуление текстовых признаков при обучении, интерполяция высоты с разным значением шага, а также введение дополнительной ошибки выравнивания признаков с коэффициентом влияния на общую ошибку.

На графиках можно отчетливо видеть, что при использовании конкатенации с различными конфигурациями $p_{textDropout}$, λ , d_{step} снижение ошибки на обучении прекращается приблизительно к 30 эпохе и далее уходит на плато. Также наблюдается высокая нестабильность показателей с одинаковой динамикой на тестовых данных на всем диапазоне значений $p_{textDropout}$. Более лучшую сходимость при обучении показал метод использования gate-суммирования, при значениях $p_{textDropout} = 0.2$, $\lambda = 0.5$, $d_{step} = 10$ и $p_{textDropout} = 0.3$, $\lambda = 0.5$, $d_{step} = 5$.

Полученные результаты наталкивают на гипотезу о том, что увеличение $p_{textDropout}$ с 0.2 до 0.3 коррелирует с уменьшением шага интерполяции высоты с 10 до 5. Объединение суммированием показало себя лучше, в том числе с точки зрения вычислений, т.к. при суммировании размерность объединенного вектора признаков не увеличивается. Использование $\lambda=0.5$ для регулирования вклада ошибки выравнивания признаков дало лучшую стабильность и динамику общей ошибки. Гипотетически, наиболее оптимальные значения λ находятся в окрестности 0.5.

Заключение

В настоящей работе предложен метод мультимодальной регрессии положения БПЛА, основанный на совместной обработке визуальных признаков и текстовых закодированных метаданных. Предложенный подход учитывает высоту съёмки как важный фактор для точности предсказания и интегрирует её в текстовый шаблон, что позволило повысить обобщающую способность модели и уменьшить чувствительность к изменению ракурсов.

Применение регуляризации через функцию выравнивания признаков и случайное обнуление текстовых признаков способствовало стабилизации процесса обучения и снижению риска переобучения на текстовую модальность. Эксперименты показали, что методы адаптивного объединения признаков с использованием механизма *gate*-суммирования демонстрируют преимущество над простой конкатенацией как по точности, так и по вычислительной эффективности. Анализ параметров модели выявил оптимальные настройки коэффициента выравнивания (~ 0.5) и шага интерполяции высоты, которые обеспечили наилучший баланс между точностью регрессии и стабильностью обучения.

Результаты подтверждают гипотезу о важности согласования числовых и визуальных признаков при решении задач позиционирования. В перспективе планируется расширить подход за счёт учёта дополнительных сенсорных данных и изучения влияния различных стратегий регуляризации внимания на устойчивость модели в условиях шумных измерений.

Список литературы

1. Пикалёв Я. С. Обнаружение ключевых объектов и перекрёстная геолокализация: Анализ наборов данных и методологические перспективы // Проблемы искусственного интеллекта. 2024. Т. 35. № 4. С. 25-37.
2. Пикалёв Я. С., Ермоленко Т. В. О нейронных архитектурах извлечения признаков для задачи распознавания объектов на устройствах с ограниченной вычислительной мощностью // Проблемы искусственного интеллекта. 2023. № 3 (30). С. 44-54.
3. Sheng K., Zhan H., Xie Z., Chen H., Xu Y., Tang L., Luo B. GeoText-1652: A Benchmark for Language-Aware Cross-View Geo-Localization // European Conference on Computer Vision (ECCV). 2024.
4. Xu Y., Zhang Z., He W., Wang J., Luo B. A Survey of Open-Vocabulary Object Detection for UAV Imagery // Drones. 2025. Vol. 9, No. 1. P. 12. DOI: 10.3390/drones9010012.
5. Sun X. et al. Spatial-LLaVA: Enhancing Large Language Models with Spatial Referring Expressions for Visual Understanding // arXiv preprint arXiv:2505.12194. – 2025.
6. Ye J. et al. Where am I? Cross-View Geo-localization with Natural Language Descriptions // arXiv preprint arXiv:2412.17007. 2024.
7. Павленко, Б. В. Методика создания набора аэрофотоснимков для задачи перекрёстной геолокализации / Б.В. Павленко, Я.С. Пикалёв // Проблемы искусственного интеллекта. 2024. №4(35). С. 101-112. DOI:10.24412/2413-7383-2024-4-101-112.

8. Пикалёв Я. С. Обнаружение ключевых объектов и перекрёстная геолокализация: Анализ наборов данных и методологические перспективы // Проблемы искусственного интеллекта. 2024. Т. 35. №. 4. С. 25-37.
9. Зуев В. М., Иванова С. Б. Оценка собственного местоположения аппарата на основе анализа видеозображения // Проблемы искусственного интеллекта. 2024. Т. 33. №. 2. С. 21-28.
10. Ji Y. et al. Game4loc: A uav geo-localization benchmark from game data // Proceedings of the AAAI Conference on Artificial Intelligence. 2025. Т. 39. №. 4. С. 3913-3921.
11. Xu W. et al. Uav-visloc: A large-scale dataset for uav visual localization // arXiv preprint arXiv:2405.11936. 2024.19:06
12. Yang Y., Newsam S. Bag-of-visual-words and spatial extensions for land-use classification // Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. 2010. – С. 270-279. Kim W., Son B., Kim I. ViLT: Vision-and-Language Transformer Without Convolution or Region Supervision // arXiv preprint, arXiv:2102.03334, 2021.
13. Huang S. et al. Language is not all you need: aligning perception with language models. arXiv // Preprint posted online February. 2023. Т. 27.
14. Tan H., Bansal M. LXMERT: Learning cross-modality encoder representations from transformers [Электронный ресурс] // arXiv preprint arXiv:1908.07490. 2019. URL: <https://arxiv.org/abs/1908.07490> (дата обращения: 18.09.2025).
15. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I. Learning Transferable Visual Models From Natural Language Supervision // arXiv preprint arXiv:2103.00020, 2021. URL: <https://arxiv.org/abs/2103.00020>
16. Qu G. et al. Stripnet: Towards topology consistent strip structure segmentation // Proceedings of the 26th ACM international conference on Multimedia. 2018. С. 283-291.
17. Faghri F. et al. MobileCLIP2: Improving Multi-Modal Reinforced Training // arXiv preprint arXiv:2508.20691. 2025.
18. Cui B., Liu Z., Yang Q. UAV-YOLO12: A Multi-Scale Road Segmentation Model for UAV Remote Sensing Imagery // Drones. 2025. Т. 9. №. 8. С. 533.
19. Less Wright, Nestor Demeure et al. Ranger: synergistic combination of RAdam + Lookahead for the best of both [Электронный ресурс] // GitHub / arXiv preprint. 2019. URL: <https://arxiv.org/abs/2004.01461> (дата обращения: 18.09.2025).
20. Liu L., Jiang H., He P., Chen W., Liu X., Gao J., Han J. On the Variance of the Adaptive Learning Rate and Beyond [Электронный ресурс] // arXiv preprint. – 2019. – № arXiv:1908.03265. – URL: <https://arxiv.org/abs/1908.03265>. (дата обращения: 18.06.2025).
21. Zhang M. R., Lucas J., Hinton G., Ba J. Lookahead Optimizer: k steps forward, 1 step back [Электронный ресурс] // Advances in Neural Information Processing Systems, NeurIPS 2019. – 2019. – Vol. 32. – P. 9593-9604. – URL: <https://arxiv.org/abs/1907.08610>. (дата обращения: 18.06.2025).

References

1. Pikalyov, Ya. S. Key Object Detection and Cross-Geolocation: Dataset Analysis and Methodological Prospects // Problems of Artificial Intelligence. - 2024. - Vol. 35. - No. 4. - Pp. 25-37.
2. Pikalyov, Ya. S., Ermoolenko, T.V. On Neural Feature Extraction Architectures for Object Recognition on Devices with Limited Computing Power // Problems of Artificial Intelligence. - 2023. - No. 3 (30). - Pp. 44-45
3. Sheng K., Zhan H., Xie Z., Chen H., Xu Y., Tang L., Luo B. GeoText-1652: A Benchmark for Language-Aware Cross-View Geo-Localization // European Conference on Computer Vision (ECCV). – 2024.
4. Xu Y., Zhang Z., He W., Wang J., Luo B. A Survey of Open-Vocabulary Object Detection for UAV Imagery // Drones. – 2025. – Vol. 9, No. 1. – P. 12. – DOI: 10.3390/drones9010012.
5. Sun X. et al. Spatial-LLaVA: Enhancing Large Language Models with Spatial Referring Expressions for Visual Understanding // arXiv preprint arXiv:2505.12194. – 2025.
6. Ye J. et al. Where am I? Cross-View Geo-localization with Natural Language Descriptions // arXiv preprint arXiv:2412.17007. – 2024.
7. Pavlenko, B.V. Methodology for Creating a Set of Aerial Photographs for the Task of Cross-Geolocation / B. V. Pavlenko, Ya. S. Pikalev // Problems of Artificial Intelligence. - 2024. - No. 4 (35). - Pp. 101-112. - DOI: 10.24412/2413-7383-2024-4-101-112.
8. Pikalyov, Ya. S. Detection of Key Objects and Cross-Geolocation: Analysis of Datasets and Methodological Prospects // Problems of Artificial Intelligence. - 2024. - Vol. 35. - No. 4. - Pp. 25-37.

9. Zuev, V.M., Ivanova S. B. Estimation of the Device's Own Location Based on Video Image Analysis // Problems of Artificial Intelligence. – 2024. – Т. 33. – No. 2. – Pp. 21-28.
10. Ji Y. et al. Game4loc: A uav geo-localization benchmark from game data //Proceedings of the AAAI Conference on Artificial Intelligence. – 2025. – Т. 39. – No. 4. – С. 3913-3921.
11. Xu W. et al. Uav-visloc: A large-scale dataset for uav visual localization //arXiv preprint arXiv:2405.11936. – 2024.19:06
12. Yang Y., Newsam S. Bag-of-visual-words and spatial extensions for land-use classification //Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems. – 2010. – С. 270-279. Kim W., Son B., Kim I. ViLT: Vision-and-Language Transformer Without Convolution or Region Supervision // arXiv preprint, arXiv:2102.03334, 2021.
13. Huang S. et al. Language is not all you need: aligning perception with language models. arXiv //Preprint posted online February. – 2023. – Т. 27.
14. Tan H., Bansal M. LXMERT: Learning cross-modality encoder representations from transformers [Электронный ресурс] // arXiv preprint arXiv:1908.07490. – 2019. – URL: <https://arxiv.org/abs/1908.07490> (дата обращения: 18.09.2025).
15. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I. Learning Transferable Visual Models From Natural Language Supervision // arXiv preprint arXiv:2103.00020, 2021. URL: <https://arxiv.org/abs/2103.00020>
16. Qu G. et al. Stripnet: Towards topology consistent strip structure segmentation //Proceedings of the 26th ACM international conference on Multimedia. – 2018. – С. 283-291.
17. Faghri F. et al. MobileCLIP2: Improving Multi-Modal Reinforced Training //arXiv preprint arXiv:2508.20691. – 2025.
18. Cui B., Liu Z., Yang Q. UAV-YOLO12: A Multi-Scale Road Segmentation Model for UAV Remote Sensing Imagery //Drones. – 2025. – Т. 9. – No. 8. – С. 533.
19. Less Wright, Nestor Demeure et al. Ranger: synergistic combination of RAdam + Lookahead for the best of both [Электронный ресурс] // GitHub / arXiv preprint. – 2019. – URL: <https://arxiv.org/abs/2004.01461> (дата обращения: 18.09.2025).
20. Liu L., Jiang H., He P., Chen W., Liu X., Gao J., Han J. On the Variance of the Adaptive Learning Rate and Beyond [Электронный ресурс] // arXiv preprint. – 2019. – No arXiv:1908.03265. – URL: <https://arxiv.org/abs/1908.03265>. (дата обращения: 18.06.2025).
21. Zhang M. R., Lucas J., Hinton G., Ba J. Lookahead Optimizer: k steps forward, 1 step back [Электронный ресурс] // Advances in Neural Information Processing Systems, NeurIPS 2019. – 2019. – Vol. 32. – P. 9593-9604. – URL: <https://arxiv.org/abs/1907.08610>. (дата обращения: 18.06.2025).

RESUME

B. V. Pavlenko

An approach to multimodal data fusion in the problem of UAV position value regression

Research into multimodal learning models and methods is currently one of the most active areas. Along with the development and implementation of intelligent unmanned systems, the need for mobile and precise intelligent positioning systems is growing. The use of additional semantics in the form of metadata allows for improving the quantitative and qualitative performance of models. Since existing datasets do not always provide rich numerical metadata on UAV position and altitude, a solution can be found in a multimodal approach to regressing the corresponding values, predicted by clarifying the context and meaning of the numerical value through text descriptions. Experiments have demonstrated the suitability of this approach for determining missing UAV positioning values based on aerial photographs and altitude using interpolation during the training process.

РЕЗЮМЕ

Б.В. Павленко

Подход к мультимодальному объединению данных в задаче регрессии значений положений БПЛА

Исследования моделей и методов мультимодального обучения являются одними из наиболее активных на сегодняшний день. Совместно с развитием и внедрением интеллектуальных беспилотных систем растет необходимость в мобильных и точных интеллектуальных системах позиционирования. Использование дополнительной семантики в виде метаданных позволяет повышать количественные и качественные показатели моделей. В силу того, что имеющиеся наборы данных не всегда располагают богатыми числовыми метаданными о положении и высоте БЛА, решением может быть мультимодальный подход к регрессии соответствующих значений, предсказываемых за счет уточнения контекста и смысла числового значения через текстовое описание. Эксперименты показали пригодность подхода к определению недостающих значений позиционирования БЛА по аэрофотоснимку и высоте с применением интерполяции в процессе обучения.

Павленко Богдан Викторович – младший научный сотрудник, аспирант, ФГБНУ «Институт проблем искусственного интеллекта», 283048, г. Донецк, ул. Артема, д. 118 б, телефон +7(949) 438-6450, bogdanpav12000@mail.ru. *Область научных интересов:* системы распознавания образов, нейронные сети, мультимодальные модели.

Статья поступила в редакцию 10.09.2025.

УДК 004.932.2

DOI 10.24412/2413-7383-2025-4-39-84-97

Устенко В.Ю., Близно М.В.

Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта», Российская Федерация г. Донецк,
283048, г. Донецк, ул. Артема, 118 б

МЕТОДИКА РАСШИРЕНИЯ ДАННЫХ ДЛЯ СЕГМЕНТАЦИИ КЛЮЧЕВЫХ ОБЪЕКТОВ НА АЭРОФОТОСНИМКАХ БПЛА

Ustenko V.Y. Blizno M.V.

Federal State Budgetary Scientific Institution «Institute of Artificial Intelligence Problems»
283048, Russian Federation, Donetsk, Artema str, 118-b

DATA AUGMENTATION METHODOLOGY FOR SEGMENTING KEY OBJECTS IN UAV AERIAL IMAGERY

В данной работе авторами предложена методика (StageAug), использующая ограниченное количество ресурсов сервера машинного обучения, повышающая робастность. Эксперименты проводятся на синтетическом наборе данных SynDrone (дорога, природные объекты, застройка, препятствия, вода, иные объекты), с использованием предварительно обученной нейросетевой архитектуры StripNet, предназначенной для извлечения высокоуровневых признаков. Ключевым элементом – поэтапный конвейер аугментаций (методов расширения данных): от мягких фотометрических преобразований до Copy-Paste редких объектов, мозаики $2 \times 2/3 \times 3$. В результате экспериментов было выявлено, что предложенная авторами методика стабилизирует ранние эпохи и обеспечивает $mIoU=0.6264$, $FWIoU=0.9565$. Результаты исследования подтверждают гипотезу: гармонизация аугментаций и использование насыщенного StripNet-FPN действительно помогает сгладить синтетико-реальный разрыв и повысить робастность.

Ключевые слова: семантическая сегментация, StripNet, FPN, SynDrone, БПЛА, аэрофотосъемка, аугментация, мозаика.

In this work, the authors propose a method (StageAug) that uses a limited amount of machine-learning server resources while increasing robustness. Experiments are conducted on the synthetic SynDrone dataset (road, natural objects, buildings, obstacles, water, other objects) using the pretrained StripNet neural architecture designed for extracting high-level features. The key element is a staged augmentation pipeline: from mild photometric transformations to Copy-Paste of rare objects and $2 \times 2/3 \times 3$ mosaic compositions. The experiments show that the proposed method stabilizes early epochs and achieves $mIoU = 0.6264$ and $FWIoU = 0.9565$. The results confirm the hypothesis that harmonizing augmentations and using the enhanced StripNet-FPN indeed help to reduce the synthetic-to-real gap and improve robustness.

Keywords: semantic segmentation, StripNet, FPN, SynDrone, UAV, aerial imagery, augmentation, mosaic.

Введение

Аэрофотосъёмка с помощью беспилотных летательных аппаратов (БПЛА) является ключевым источником данных для задач мониторинга инфраструктуры, реагирования в чрезвычайных ситуациях и картографирования городских пространств [1]. Семантическая сегментация дорожных и природных объектов позволяет автоматизировать анализ изображений с камеры БПЛА, повышая скорость и качество, но не лишена недостатков. Реальные и синтетические снимки с камеры БПЛА, полученные при различных условиях, формируют доменный сдвиг, снижающий устойчивость стандартных сегментационных систем [2].

Большинство архитектур моделей глубокого обучения для семантической сегментации объектов на аэрофотоснимках опираются на большие предварительно обученные нейросетевые архитектуры, предназначенные для извлечения высокоуровневых признаков (backbone) и единообразные аугментации [3]. Такие методы показывают высокое влияние на качество обучения модели глубокого обучения на синтетических наборах данных, что часто выражается переобучением, при смешении реальных и синтетических сцен SynDrone [4], плохо охватывают редкие мета-классы и перегружают вычислительные ресурсы при прямом обучении.

Слабые аугментации (weak) – это минимальные преобразования, сохраняющие геометрию и основное семантическое содержание сцены. Главная их цель – повысить устойчивость модели семантического анализа к базовым вариациям данных без внесения выраженных артефактов [7]. К таким преобразованиям относятся зеркальное отражение, небольшие повороты, мягкие фотометрические сдвиги (яркость, контраст, насыщенность, тон), гауссовский шум, лёгкое размытие, случайные небольшие обнуления областей, кадрирование и нормализация. Weak-аугментации применяют на ранних этапах обучения для регуляризации и контроля переобучения.

Сильные аугментации (strong) существенно изменяют изображение, расширяя распределение данных и создавая агрессивные вариации, которые помогают бороться с доменным сдвигом, шумом сенсора, низким качеством съёмки и изменчивыми погодными условиями. В отличие от weak-аугментаций, здесь допускаются большие фотометрические и геометрические смещения: масштабирование, вращение и сдвиг, сильные регулировки HSV, моделирование шума, JPEG-деградация, искусственное снижение и восстановление разрешения, туман, синтетические тени [8]. Эти преобразования допускают заметное отличие от исходного изображения, но сохраняют семантику класса. Strong-аугментации широко используются в semi-supervised-подходах и при обучении сложных моделей.

Продвинутые аугментации (advanced, ADV) – это специально построенные возмущения, направленные на максимизацию ошибки модели при минимальной визуальной заметности. Аугментация строится с использованием градиента модели или с учётом изменения предсказаний. Такие методы моделируют преднамеренные атакующие воздействия и существенно повышают робастность модели к advanced примерам. Advanced-аугментации особенно важны в сценариях с требованиями к безопасности и устойчивости [9].

В работе предложены две гипотезы: (1) комбинирование weak, strong, advanced преобразований с Copy-Paste, мозаикой сокращает разрыв между синтетикой и реальностью; (2) использование предварительно обученной нейросетевой модели StripNet снижает требуемое число эпох и стабилизирует ранние стадии обучения.

Постановка задачи

Цель данной работы заключается в разработке методики расширения данных для аэрофотоснимков в задаче семантической сегментации, обеспечивающей выделение ключевых классов (в соответствии с таблицей 2) в условиях доменного сдвига, наличия шумов, разных высот, углов поворота и т.п. Для достижения этой цели были поставлены и решены следующие задачи:

1. Провести системный анализ существующих методов аугментации, определить их преимущества и ограничения с точки зрения устойчивости к синтетико-реальным сдвигам и затрат ресурсов;
2. Выбрать нейросетевую архитектуру для задачи семантической сегментации ключевых объектов;
3. Разработать методику расширения данных (AugUAV) для задачи семантической сегментации аэрофотоснимков;
4. Выбрать ключевые классы для задачи семантической сегментации аэрофотоснимков. Переформатировать сегментационные карты в наборе данных SynDrone на основе выбранных ключевых классов;
5. Провести численные эксперименты по сравнению стандартных методов расширения данных и AugUAV на тех же конфигурациях модели.

Математическое описание задачи

Пусть имеется выборка (1), где $(x_i \in R^{3 \times H \times W})$ – RGB-снимок с борта БПЛА (высота $(H=W=384)$ после приведения), а $(y_i \in 0 \dots 255^{H \times W})$ – пиксельная разметка ключевых классов (в соответствии с таблицей 2). Требуется обучить параметрическое отображение (2), которое максимизирует вероятность корректного класса для каждого пикселя, при этом выдерживая доменный сдвиг между реальными и синтетическими сценами (разные высоты, сезоны и освещённость) и ограничение на вычислительные ресурсы.

Формально задача сводится к минимизации (3) с учётом маски игнорируемых пикселей и дополнительного регуляризатора DiceLoss для сбалансирования классов. Решение считается успешным, если на валидационном наборе достигаются высокие значения основных количественных показателей семантической сегментации: средняя пересекающаяся площадь (4) средняя точность классов (5) и взвешенная по частотам FWIoU. Целевой уровень – $(mIoU \geq 0.62)$, $(mAcc \geq 0.66)$, $(FWIoU \geq 0.95)$, что демонстрирует адаптивность модели к смешанному домену SynDrone.

$$(\mathcal{D} = (x_i, y_i)_{i=1}^N), \quad (1)$$

$$(f_{\theta}: R^{3 \times H \times W} \rightarrow [0,1]^{5 \times H \times W}), \quad (2)$$

$$L_{CE} = - \sum_{i=1}^N \sum_{p \in \Omega} \log f_{\theta}^{(y_i^p)}(x_i)p, \quad (3)$$

$$\left(mIoU = \frac{1}{5} \sum_c = 0^4 \frac{TP_c}{TP_c + FP_c + FN_c} \right), \quad (4)$$

$$\left(mAcc = \frac{1}{5} \sum_{c=0}^4 \frac{TP_c}{TP_c + FN_c} \right), \quad (5)$$

Связанные работы

Семантическая сегментация при малых выборках и редких классах.

В условиях ограниченных данных используются few-shot методы (мета-обучение, прототипы) и усиленная аугментация, включая синтетические данные, что позволяет сегментировать новые и редкие классы с приемлемой точностью.

Few-shot сегментация решает проблему малого числа размеченных данных [10]. One-shot модель Shaban et al. дала ~25% прироста mIoU на новых классах VOC, показав эффективность мета-обучения [11]. Для редких классов применяют аугментации и синтетические данные: добавление 5–100k примеров существенно снижает ошибку без ущерба другим классам [12]. В итоге комбинация мета-обучения, прототипных представлений и синтетических аугментаций обеспечивает приемлемую точность на новых и редких классах.

Сегментация городских сцен с беспилотников

UAV-сегментация осложнена наклонной съёмкой и сильной вариабельностью масштабов объектов [13]. Специализированный датасет UAVid (300 4K-кадров, 8 классов) выявил проблемы временной согласованности и масштабной инвариантности; базовые модели дают ~50% mIoU, улучшения достигаются многоуровневыми архитектурами и 3D-CRF [14]. Дополнительные подходы включают объединение разнородных UAV-датасетов через CSN для повышения устойчивости к доменным различиям [15], а также лёгкие модели реального времени с глобально-локальным вниманием на базе ResNet-18 [16]. Специализированные датасеты и контекстные архитектуры остаются ключевыми для повышения качества сегментации UAV-сцен.

Поэтапные аугментации и обучением с постепенным усложнением (curriculum learning)

Curriculum-стратегии уменьшают разрыв между синтетическими и реальными данными: в [17] модель сначала решает простые подзадачи (глобальные и локальные распределения меток), затем использует их как регуляризаторы при финальном обучении. Для аугментаций действует тот же принцип: в [18] сложность Colorful Cutout увеличивается по мере обучения, улучшая обобщение. В детекции аналогично вводят warm-up без тяжёлых трансформаций, включая их позже [19]. Эти поэтапные схемы стабилизируют ранние эпохи и повышают итоговую точность.

Доменная адаптация и обобщение (синтетические наборы данных против реальных)

Сегментация страдает от различий между синтетическими и реальными данными [20]. Несупервизируемая адаптация выравнивает распределения через дискриминаторы [21], работающие по выходным маскам и многослойным признакам, улучшая перенос структуры сцены [22]. Другой подход — стиливые преобразования: циклическая стилизация под целевой домен [23] и генерация “трудных” стиливых вариаций (AdvStyle, MixStyle и др.) [24], повышающих инвариантность. Эти методы либо адаптируют сеть к целевому домену, либо формируют модель, устойчивую к любым стиливым сдвигам без доступа к целевым данным [17].

Предобученные и специальные backbone в сегментации

Сегментация выигрывает от мощных энкодеров: предобученные сети вроде VGG [10], ResNet [10] и их версии в DeepLab/UNet ускоряют обучение и повышают точность [14]. Для мобильных сценариев применяют облегчённые MobileNet/Xception. Специализированные backbone (HRNet, LSKNet [25], StripNet [5]) учитывают особенности задачи: сохранение высокого разрешения, адаптивный контекст или топология длинных объектов. В итоге современные подходы либо адаптируют проверенные классификационные модели, либо создают энкодеры под конкретные требования сегментации.

Сопоставление с предлагаемым подходом

Существующие направления по отдельности решают проблему, связанную с доменным сдвигом, устойчивостью и постепенным усложнением обучения, но редко объединяются в единую систему. Предлагаемая методика AugUAV совмещает специализированный backbone StripNet, адаптивный FPN-декодер и поэтапные аугментации. Модель начинает обучение на простых сценах, затем используется подход Copy-Paste для редких классов и Mosaic-композиции, что повышает обобщающую способность без перегрузки вычислений. Такой curriculum-подход стабилизирует ранние эпохи, ускоряет сходимость и улучшает качество сегментации ключевых объектов на аэрофотоснимках.

Набор данных

В работе для экспериментов используется набор данных SynDrone, включающий реальные съёмки городов (Town01–Town10HD, высоты 20/50/80 м) и соответствующие синтетические сцены, рендеренные в идентичной геометрии. Каждая запись описывается путями к RGB-кадру и семантической маске, анотациями по углам (yaw/pitch/roll), высоте и доменным признакам (день/ночь, сезон). Итоговые CSV (syndrone_train.csv, syndrone_test.csv) содержат десятки тысяч изображений, которые после ремапа в 5 метакатегорий (Road, Nature, Construction, Obstacle, Water) и добавления индекса 255 для Void подходят для задач пиксельной сегментации. Главное преимущество SynDrone – контролируемый доменный сдвиг: синтетические сцены покрывают редкие погодные и сезонные комбинации, тогда как реальные кадры отражают шумы сенсора и артефакты аэрофотосъёмки. Это позволяет проверять гипотезу о пользе поэтапных аугментаций: метод должен одинаково хорошо сегментировать и «чистые» рендеры, и noisy-реальность. Кроме того, фиксированные ID для train/test, в соответствии с таблицей 1, делают эксперимент воспроизводимым и дают возможность отслеживать метрики (mIoU/mAcc/FWIoU) на неизменной валидации. Авторами в работе используется единый формат входа. RGB-кадр 384×384 пикселей (кадр приводится к квадрату в процессе предобработки; оригинальный SynDrone содержит неравномерные разрешения, но вы нормализуете до 384×384).

Сегментационные карты хранятся как индексированные 8-битные PNG (по одному каналу), где каждый пиксель содержит ID тонкого класса SynDrone; преобразование в 5 мета-категорий выполняется в датасете.

Таблица 1 – разделение на выборки.

Подвыборка	Количество кадров
Train	60 000
Test	12 000
Всего	72 000

Выбор ключевых классов

В задачах сегментации аэрофотоснимков с БПЛА критически важно обеспечить устойчивую классификацию крупных и неизменяемых структур сцены, которые формируют геометрию городской среды и сохраняют визуальную консистентность при изменении высоты, угла съёмки и погодных условий. Полный набор SynDrone включает 105 классов, однако лишь небольшая их часть представляет собой действительно статичные, структурно стабильные и многообразные объекты среды, тогда как значительное число категорий относится к динамическим объектам – транспортным

средствам, пешеходам, велосипедистам, мотоциклам и т.п. Такие объекты обладают крайне малым размером в кадре, особенно на высотах 50–80 м; имеют нестабильную форму, что приводит к высокой внутриклассовой вариативности; отсутствуют в ряде реальных UAV-датасетов либо представлены неполно; занимают менее 1% пикселей сцены, что приводит к дисбалансу классов. Таким образом, динамические категории не обладают свойствами, необходимыми для устойчивой сегментации, и их исключение повышает обобщающую способность модели, устраняет шум редких классов и улучшает перенос на реальные данные.

В связи с этим в работе предлагается следующая система метаклассов, включающая только статичные и геометрически устойчивые элементы.

1. *Road* включает дорожное покрытие, разметку, тротуары и рельсы — объекты, формирующие опорную структуру сцены и обладающие высокой стабильностью.

2. *Nature* объединяет растительность, газоны, почву и природные поверхности, которые занимают значительную часть городской среды и хорошо различимы в UAV-данных.

3. *Construction* включает здания, стены, крыши, мосты и другие крупные архитектурные элементы, которые являются полностью статичными и определяют пространственную структуру города.

4. *Obstacle* аккумулирует малые, но статичные элементы инфраструктуры — столбы, знаки, светофоры, ограничители. Их объединение в один класс позволяет сохранить важную информацию о потенциальных препятствиях без перегрузки модели мелкими категориями.

5. *Water* выделен в отдельный класс из-за особых оптических свойств водных поверхностей и их критической важности для UAV-безопасности.

В исходной разметке SynDrone используется «тонкая» таксономия Carla/Cityscapes [26]: ID покрывают диапазон 0–105. Авторами было произведено переформатирование данных классов в 5 ключевых классов, в соответствии с таблицей 2.

Таблица 2 – Соотношение классов SynDrone к ключевым классам

Мета-класс	Fine-ID, входящие в мета-класс	Карта цветов	Объекты входящие в мета-класс
Road	6, 7, 8, 16, 27, 28, 62, 90, 91, 97, 98	[128, 64, 128]	дорога, тротуар, парковки, бордюры, разметка и дефекты покрытия
Nature	9, 10, 14, 22, 23, 24, 25, 36, 37, 81, 82, 88, 89	[107, 142, 35]	трава, кусты, деревья, грунт и прочие растительные поверхности
Construction	1, 2, 11, 15, 35, 55, 56, 57, 58, 59, 60, 61, 63, 64, 65, 66, 67, 75, 76, 80, 84, 86, 92, 93, 94	[70, 70, 70]	здания, стены, мосты, крыши и другие элементы капитальных сооружений
Obstacle	3, 4, 5, 12, 17, 18, 19, 26, 29, 30, 31, 32, 33, 34, 38, 39, 43, 44, 45, 68, 69, 70, 71, 72, 73, 74, 77, 78, 83, 85, 87, 95, 96, 99	[255, 255, 0]	ограждения, столбы, знаки, камни, уличная мебель, различные малые объекты и вертикальные препятствия
Water	21	[45, 60, 150]	любые водные поверхности
Void / Ignore	0, 13, 20, 40–42, 46–54, 79, 100–105	[96, 96, 96]	небо, люди, транспорт, прочие нецелевые классы

Методология

Следуем гипотезе, что сочетание предобученного визуального энкодера и поэтапного пайплайна аугментаций снижает синтетико-реальный сдвиг SynDrone:

- 1) подготовка данных с weak, strong, advanced аугментациями;

- 2) извлечение пирамидальных признаков StripNet;
- 3) адаптивный FPN-декодер;
- 4) пятиканальная сегментационная голова с комбинированной функцией потерь.

Энкодер признаков. StripNetExtractor (вариант StripNet Small) выдаёт три карты признаков (F_3, F_4, F_5). Модель переносится на GPU и может обучаться end-to-end или частично фиксироваться – гипотеза (1): предобученный бекбон ускоряет выход из «холодного старта» и уменьшает вариативность на ранних эпохах.

Голова и выходы. FPN-декодер выравнивает числа каналов латеральными 1×1 свёртками, выполняет top-down суммирование с билинейным апсемплингом, сглаживает Conv-BN-ReLU и завершает последовательностью Conv-BN-ReLU \rightarrow Dropout2d $\rightarrow 1 \times 1$ Conv. Результат – логиты ($Z \in R^{5 \times H \times W}$), которые апсемплируются до разрешения входа и нормализуются softmax.

Регуляризации и спецмеханизмы. Набор данных управляет стадиями аугментаций: get_weak_aug (мягкие геометрия+фотометрика), get_strong_aug, get_advanced_aug (Copy-Paste редких классов, мозаика $2 \times 2/3 \times 3$). В обучении используются Ranger+Lookahead, AMP GradScaler, клиппинг градиентов.

Функция потерь. Оптимизируем

$$\mathcal{L} = \lambda_{CE} \mathcal{L}_{CE} + \lambda_{Dice} \mathcal{L}_{Dice},$$

где (\mathcal{L}_{CE}) – кросс-энтропия с *label smoothing* и игнорированием класса 255, а

$$\mathcal{L}_{Dice} = 1 - \frac{2 \sum_c \sum_p \hat{Y}_{c,p} Y_{c,p}}{\sum_c \sum_p \hat{Y}_{c,p} + Y_{c,p} + \epsilon}$$

балансирует редкие метакатегории. Такая архитектура подтверждает гипотезу (2): stage-подход + StripNet-FPN дают mIoU 0.626 при умеренной нагрузке CPU.

Архитектура

Backbone (в соответствии с рисунком 2) основан на StripNet (конфигурация архитектуры Small): модуль StripNetExtractor загружает предобученную модель и возвращает три пирамидальные карты признаков. Блок завернут в BackboneLoader, который отвечает за перенос на GPU, управление заморозкой слоёв и унификацию формата выхода (список $f3-f5$), что обеспечивает единую точку входа для всех верхних уровней.

Декодер (в соответствии с рисунком 2). Поверх StripNet располагается FPN-голова. Каждый уровень признаков проходит через CNN 1×1 для выравнивания числа каналов; далее выполняется классическое top-down-суммирование с билинейным апсемплингом и сглаживание Conv-BN-ReLU (3×3). Итоговый тензор из 256 каналов поступает в сегментационную голову (в соответствии с рисунком 3), реализованную последовательностью формируя пятиканальную карту классов.

Классификационная ветвь (в соответствии с рисунком 1). После декодера логиты поднимаются до исходного разрешения 384×384 билинейным апсемплингом, формируя итоговые сегментационные предсказания.

Функциональный блок потерь (в соответствии с рисунком 4). Система использует две параллельные ветви: CrossEntropy с *label smoothing* (учитывает `ignore_index=255`) и DiceLoss. Такой состав оптимизации балансирует пиксельную и региональную точность.

Пайплайн аугментаций AugUAV (в соответствии с рисунком 5). AugUAV включает три параллельных набора аугментаций (weak/strong/advanced), выбираемых в зависимости от эпохи. Weak включает флипы и цветовую коррекцию; strong

добавляет аффинные преобразования и цветовые искажения; advanced расширяется за счет Copy-Paste редких классов, мозаик $2 \times 2/3 \times 3$. Буфер выступает дополнительной ветвью для Copy-Paste и мозаики.

Стратегия обучения (в соответствии с рисунком 6). Класс TrainerSynDroneSegmentation поднимает две ветви данных (train/test DataLoader), управляет AMP-микшированием через GradScaler, выполняет цикл обучения и валидации, ведёт мониторинг (лог-файлы и CSV), а также сохраняет веса по лучшему mIoU. Для глубокого анализа предусмотрен модуль расчёта IoU по каждому метаклассу и выгрузки в отдельный CSV.

Конфигурирование (в соответствии с рисунком 6).

Такой модульный стек (StripNet → FPN → CE+Dice → многостадийные аугментации) обеспечивает устойчивость к синтетико-реальным доменным сдвигам, корректный учёт мелких объектов и стабилизацию ранних эпох за счёт продвинутых операций Copy-Paste и мозаик (в соответствии с рисунком 1, в соответствии с рисунком 5).

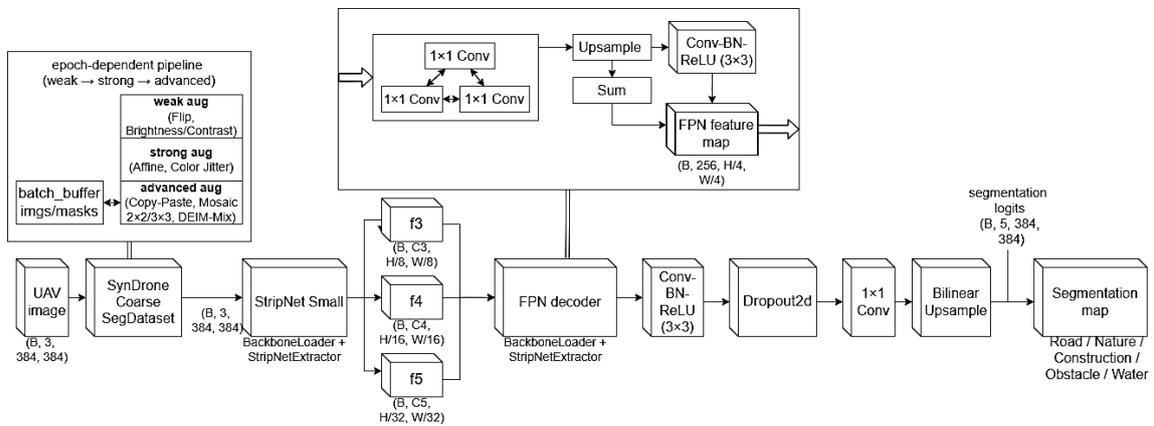


Рисунок 1 – Общая архитектура модели сегментации

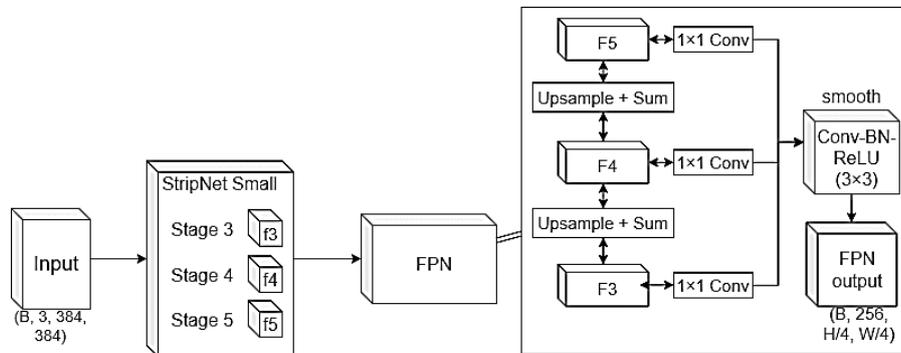


Рисунок 2 – Архитектура backbone и FPN-декодера

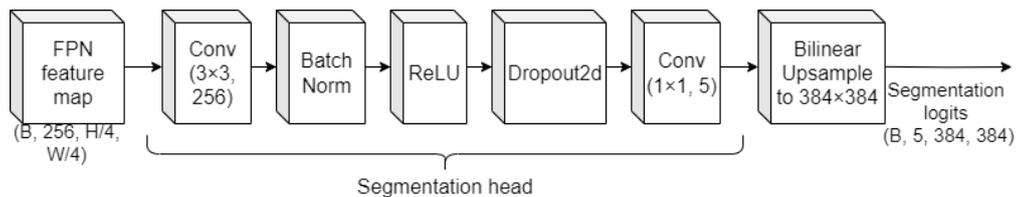


Рисунок 3 – Архитектура классификационной головы

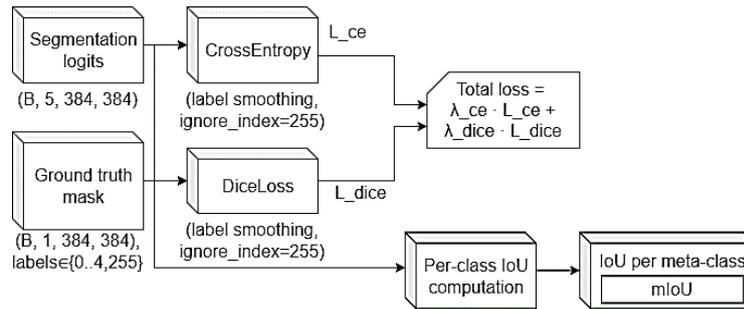


Рисунок 4 – Блок функций потерь и метрик

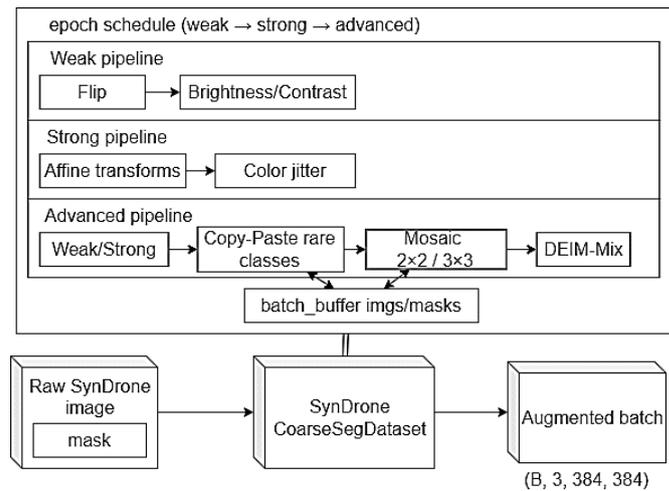


Рисунок 5 – Пайплайн аугментаций и буфера

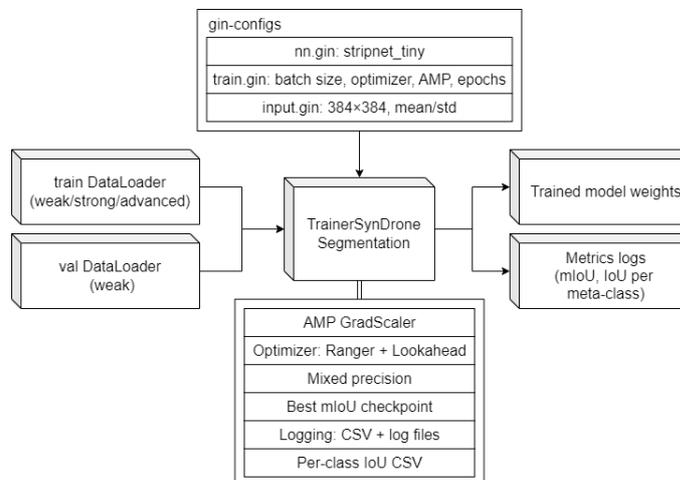


Рисунок 6 – Стратегия обучения и конфигурация

Эксперименты

Настройки. Реализация на PyTorch. Используем оптимизатор Ranger с Lookahead и ReduceLRonPlateau. Начальный lr=1e-4 (warmup шаги 0, initial_lr формально 1e-10), weight_decay 0.01, $\beta_1/\beta_2=0.9/0.999$, eps=1e-8. Batch size: 16 (train) и 16 (val), 10 эпох, GradScaler AMP (fp16), клиппинг градиентов до 1.0. Все эксперименты выполняются на одном GPU, CPU-загрузка зависит от стратегии аугментаций.

Стратегия обучения. Тонкая настройка StripNet позволяет дообучать все уровни. Поэтапный тренинг каждой эпохи заново инициализирует SynDroneCoarseSegDataset и задаёт current_epoch, чтобы переключать weak, strong, advanced аугментации. Классический запуск использует «жёсткий» набор аугментаций без переключений.

Разделение данных. Подготовленные CSV (syndrone_train.csv, syndrone_test.csv) формируются скриптом SynDroneSegRegDatasetPrep; train/val раздел зафиксирован списками ID из исходного набора и одинаков для всех экспериментов. Тестирование выполняется на real-поднаборе «test».

Бейзлайны. Сравниваются два режима:

- 1) классическое обучение (однородные сильные аугментации, высокая загрузка CPU);
- 2) поэтапное обучение AugUAV (weak, strong, advanced pipeline, умеренная загрузка CPU). Архитектура и гиперпараметры идентичны, отличается только логика аугментаций.

Результаты. Классический режим достигает mIoU=0.6246, mAcc=0.6620, FWIoU=0.9555 (в соответствии с таблицей 3); поэтапный – 0.6264/0.6636/0.9565 (в соответствии с таблицей 4). Преимущество второго проявляется уже с 4-й эпохи.

Таблица 3 – результаты обучения классическим способом

Epoch	Train Loss	Val Loss	mIoU	mAcc	FWIoU
1	2.4489	2.3327	0.1389	0.2470	0.3412
2	2.4488	2.3920	0.1263	0.2325	0.3030
3	0.8562	0.7949	0.5540	0.6008	0.9159
4	0.6425	0.7482	0.5814	0.6246	0.9335
5	0.5913	0.7252	0.5950	0.6348	0.9418
6	0.5630	0.7072	0.6065	0.6473	0.9471
7	0.5423	0.6991	0.6123	0.6513	0.9504
8	0.5302	0.6913	0.6171	0.6548	0.9526
9	0.5204	0.6855	0.6210	0.6572	0.9539
10	0.5131	0.6803	0.6246	0.6620	0.9555

Таблица 4 – результаты обучение предложенной авторами методикой

Epoch	Train Loss	Val Loss	mIoU	mAcc	FWIoU
1	2.4420	2.3878	0.1059	0.2072	0.2600
2	2.4415	2.3923	0.1029	0.2041	0.2527
3	0.8507	0.7967	0.5516	0.5942	0.9170
4	0.6445	0.7477	0.5802	0.6197	0.9347
5	0.5907	0.7234	0.5963	0.6338	0.9432
6	0.5592	0.7074	0.6076	0.6469	0.9481
7	0.5416	0.6960	0.6135	0.6507	0.9510
8	0.5269	0.6896	0.6186	0.6562	0.9526
9	0.5177	0.6835	0.6228	0.6608	0.9548
10	0.5093	0.6776	0.6264	0.6636	0.9565

На основе полученных результатов численных экспериментов можно сделать вывод, что:

- 1) все улучшения после 3-й эпохи стабильны и однонаправленны;
- 2) Δ Val Loss положительная только на эпохах 1–3, далее поэтапный режим всегда лучше (меньше loss);
- 3) mIoU, mAcc, FWIoU демонстрируют устойчивое преимущество на поздних эпохах, что соответствует гипотезе о постепенном разогреве модели;
- 4) Train loss почти всегда чуть ниже, что указывает на более стабильную оптимизацию.

Проведенные численные эксперименты позволяют сделать достаточно строгую интерпретацию, подтверждающую рабочую гипотезу: поэтапная стратегия обучения (weak, strong, advanced) не ухудшает динамику сходимости и демонстрирует систематические, пусть и предельно малые, улучшения качества при том же числе эпох и умеренной нагрузке на вычислительные ресурсы. Минимальное систематическое снижение валидационной ошибки свидетельствует о более плавной адаптации к доменному сдвигу, что критично для SynDrone.

По совокупности наблюдений предложенная авторами методика демонстрирует устойчивые, структурно согласованные улучшения всех ключевых метрик сегментации при сохранении вычислительных требований на умеренном уровне. Поэтапный pipeline действительно стабилизирует обучение: после провального старта ($mIoU \approx 0.1$) обе схемы выходят к 0.55 на 3-й эпохе, но staged-режим быстрее сокращает разрыв между `train_loss` и `val_loss`, а финальные метрики (0.6264 $mIoU$, 0.6636 $mAcc$, 0.9565 $FWIoU$) немного превосходят классический запуск. Видимо, soft, strong, advanced аугментации позволяют модели сначала освоить базовую геометрию, а затем безопасно включить Copy-Paste/мозаики.

Заключение

В настоящей работе рассмотрена задача сегментации пяти мета-категорий на смешанных реальных/синтетических UAV-сценах SynDrone. Эксперименты показали, что сочетание StripNet-backbone, адаптивного FPN и методики аугментаций AugUAV стабилизирует обучение при ограниченных ресурсах.

Результаты численных экспериментов показали, что AugUAV умеренно нагружает CPU и подходит для GPU конфигураций; классический метод расширения данных работает в режиме максимальной вычислительной занятости, достигая 100-процентной утилизации ресурсов. С другой стороны, AugUAV требует буферизации батчей для мозаики, поэтому при ограниченной памяти на CPU/RAM могут возникать трудности. В целом подход применим к UAV-сценариям, где важен баланс между качеством и ресурсами. Таким образом, AugUAV сохраняет умеренную нагрузку CPU и даёт небольшое, но стабильное усиление качества ($mIoU$ 0.6264 vs 0.6246, $FWIoU$ 0.9565 vs 0.9555) без изменения архитектуры при обучении на 10 эпохах.

Дальнейшие планы:

- 1) добавить мультимодальные подсказки (текст/сенсорные метаданные) для модели семантической сегментации;
- 2) увеличить количество эпох при проведении численных экспериментов;
- 3) оптимизировать объектные аугментации на уровне C++/CUDA.

Результаты подтверждают гипотезу о том, что методика расширения данных AufUAV с использованием поэтапных разноуровневых аугментаций сглаживает синтетико-реальный сдвиг ($mIoU$ до 0.6264), тогда как при классическом запуске прирост ниже и ресурсы расходуются неэффективно.

Список литературы

1. Пикалёв, Я. С. Обнаружение ключевых объектов и перекрёстная геолокализация: Анализ наборов данных и методологические перспективы / Я.С. Пикалёв [Электронный ресурс] // Проблемы искусственного интеллекта. – 2024. – №4(35). – С. 25-37. DOI: 10.24412/2413-7383-2024-4-25-37. URL: http://paijournal.guiaidn.ru/download_pai/2024_4/3_%D0%9F%D0%B8%D0%BA%D0%B0%D0%BB%D0%B5%D0%B2.pdf (дата обращения: 17.11.2025).

2. Ермоленко, Т. В. К вопросу о применении глубокого обучения к задачи перекрёстной геолокализации / Т.В. Ермоленко, Р.С. Хакимов [Электронный ресурс] // Проблемы искусственного интеллекта. – 2024. – №4(35). – С.4-15. – DOI:10.24412/2413-7383-2024-4-4-15. URL: http://paijournal.guide.ru/download_pai/2024_4/1_%D0%95%D1%80%D0%BC%D0%BE%D0%BB%D0%B5%D0%BD%D0%BA%D0%BE_%D0%A5%D0%B0%D0%BA%D0%B8%D0%BC%D0%BE%D0%B2.pdf (дата обращения: 17.11.2025).
3. Кравченко С. В. и др. Проблемы детектирования объекта на изображении в задачах глубокого обучения в области компьютерного зрения на основе свёрточных нейронных сетей // Инновации и инвестиции. 2020. №6. . 2020. С. 194–197.
4. Rizzoli G. и др. SynDrone -- Multi-modal UAV Dataset for Urban Scenarios [Электронный ресурс]. 2023. URL: <https://arxiv.org/pdf/2308.10491> (дата обращения: 14.11.2025).
5. Qu G. и др. StripNet [Электронный ресурс] // Proceedings of the 26th ACM international conference on Multimedia. 2018. С. 283–291. URL: <https://dl.acm.org/doi/10.1145/3240508.3240553> (дата обращения: 16.11.2025).
6. Lin T.-Y. и др. Feature Pyramid Networks for Object Detection [Электронный ресурс]. 2017. URL: <https://arxiv.org/pdf/1612.03144v2> (дата обращения: 10.11.2025).
7. Berthelot D. и др. MixMatch: A Holistic Approach to Semi-Supervised Learning [Электронный ресурс]. 2019. URL: <https://arxiv.org/pdf/1905.02249> (дата обращения: 18.11.2025).
8. Foi A. и др. Practical Poissonian-Gaussian Noise Modeling and Fitting for Single-Image Raw-Data [Электронный ресурс] // IEEE Transactions on Image Processing. 2008. С. 1737–1754. URL: https://www.researchgate.net/publication/23249686_Practical_Poissonian-Gaussian_Noise_Modeling_and_Fitting_for_Single-Image_Raw-Data (дата обращения: 18.11.2025).
9. Miyato T. и др. Virtual Adversarial Training: A Regularization Method for Supervised and Semi-Supervised Learning [Электронный ресурс] // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019. Т. 41, № 8. С. 1979–1993. URL: <https://ieeexplore.ieee.org/document/8417973> (дата обращения: 18.11.2025).
10. Catalano N., Matteucci M. Few Shot Semantic Segmentation: a review of methodologies, benchmarks, and open challenges [Электронный ресурс]. 2024. URL: <https://arxiv.org/pdf/2304.05832v2> (дата обращения: 16.11.2025).
11. Shaban A. и др. One-Shot Learning for Semantic Segmentation [Электронный ресурс]. 2017. URL: <https://arxiv.org/pdf/1709.03410> (дата обращения: 16.11.2025).
12. Veery S. и др. Synthetic Examples Improve Generalization for Rare Classes [Электронный ресурс]. 2019. URL: <https://arxiv.org/pdf/1904.05916> (дата обращения: 16.11.2025).
13. Зуев, В. М. Сравнение обнаружения объектов средствами искусственного интеллекта в сравнении с классическими методами / В.М. Зуев [Электронный ресурс] // Проблемы искусственного интеллекта. – 2024. – Т. 34 (3). – С. 4-10. URL: http://paijournal.guide.ru/download_pai/2024_3/3_%D0%97%D1%83%D0%B5%D0%B2.pdf (дата обращения: 17.11.2025).
14. Lyu Y. и др. UAVid: A Semantic Segmentation Dataset for UAV Imagery [Электронный ресурс]. 2020. URL: <https://arxiv.org/pdf/1810.10438> (дата обращения: 16.11.2025).
15. Song A. Deep Learning-Based Semantic Segmentation of Urban Areas Using Heterogeneous Unmanned Aerial Vehicle Datasets [Электронный ресурс] // Aerospace. 2023. С. 880. URL: <https://www.mdpi.com/2226-4310/10/10/880#:~:text=areas,be%20used%20on%20diverse%20data> (дата обращения: 16.11.2025).
16. Zhang Z., Li G. UAV Imagery Real-Time Semantic Segmentation with Global–Local Information Attention [Электронный ресурс] // Sensors. 2025. С. 1786. URL: <https://www.mdpi.com/1424-8220/25/6/1786#:~:text=semantic%20segmentation%20of%20drones%20that,branch%20compresses%20and%20extracts%20global> (дата обращения: 16.11.2025).
17. Zhang Y., David P., Gong B. Curriculum Domain Adaptation for Semantic Segmentation of Urban Scenes [Электронный ресурс]. 2018. URL: <https://arxiv.org/pdf/1707.09465> (дата обращения: 16.11.2025).
18. Choi J., Kim Y. Colorful Cutout: Enhancing Image Data Augmentation with Curriculum Learning [Электронный ресурс]. 2024. URL: <https://arxiv.org/pdf/2403.20012> (дата обращения: 16.11.2025).
19. Huang S. и др. DEIM: DETR with Improved Matching for Fast Convergence [Электронный ресурс]. 2025. URL: <https://www.themoonlight.io/en/review/deim-detr-with-improved-matching-for-fast-convergence> (дата обращения: 16.11.2025).

20. Зуев, В. М. Подготовка данных для обучения нейронной сети, управляющей движением механизма / В.М. Зуев, О.А. Бутов, А.А. Никитина, С.И. Уланов [Электронный ресурс] // Материалы Донецкого международного круглого стола «Искусственный интеллект: теоретические аспекты и практическое применение. ИИ – 2021». – 2021. – С. 92-95. ГУ «ИПИИ». URL: http://paijournal.guiaidn.ru/download_pai/2021_2/2_%D0%97%D1%83%D0%B5%D0%B2_%D0%91%D1%83%D1%82%D0%BE%D0%B2_%D0%98%D0%B2%D0%B0%D0%BD%D0%BE%D0%B2%D0%B0_%D0%9D%D0%B8%D0%BA%D1%82%D0%B8%D0%BD%D0%B0_%D0%A3%D0%BB%D0%B0%D0%BD%D0%BE%D0%B2.pdf (дата обращения: 17.11.2025).
21. Пикалёв, Я. С. О нейронных архитектурах извлечения признаков для задачи распознавания объектов на устройствах с ограниченной вычислительной мощностью / Я.С. Пикалёв [Электронный ресурс] // Проблемы искусственного интеллекта. – 2023. – № 2413-7383. – С. 44-54. – DOI: 10.34757/2413-7383.2023.30.3.004. – ISSN: 2413-7383. URL: http://paijournal.guiaidn.ru/download_pai/2023_3/4_%D0%9F%D0%B8%D0%BA%D0%B0%D0%BB%D0%B5%D0%B2_%D0%95%D1%80%D0%BC%D0%BE%D0%BB%D0%B5%D0%BD%D0%BA%D0%BE.pdf (дата обращения: 17.11.2025).
22. Tsai Y.-H. и др. Learning to Adapt Structured Output Space for Semantic Segmentation [Электронный ресурс]. 2020. URL: <https://arxiv.org/pdf/1802.10349> (дата обращения: 16.11.2025).
23. Hoffman J. и др. CyCADA: Cycle-Consistent Adversarial Domain Adaptation [Электронный ресурс]. 2017. URL: <https://arxiv.org/pdf/1711.03213> (дата обращения: 16.11.2025).
24. Zhong Z. и др. Adversarial Style Augmentation for Domain Generalized Urban-Scene Segmentation [Электронный ресурс]. 2022. URL: <https://arxiv.org/pdf/2207.04892> (дата обращения: 16.11.2025).
25. Li Y. и др. LSKNet: A Foundation Lightweight Backbone for Remote Sensing [Электронный ресурс]. 2025. URL: <https://arxiv.org/pdf/2403.11735> (дата обращения: 16.11.2025).
26. Szántó M. и др. Building Maps Using Monocular Image-feeds from Windshield-mounted Cameras in a Simulator Environment [Электронный ресурс] // Periodica Polytechnica Civil Engineering. 2023. URL: <https://doi.org/10.3311/PPci.21500> (дата обращения: 10.11.2025).

RESUME

V. Yu. Ustenko, Blizno M.V.

Data Augmentation Method for Segmenting Key Objects in UAV Aerial Imagery

Semantic segmentation of urban UAV scenes is complicated by the domain gap between real and synthetic SynDrone data and by the high computational demands of training. This study examines the combination of a pre-trained StripNet backbone and an adaptive FPN decoder with the authors' proposed AugUAV data augmentation methodology for semantic segmentation of five key classes: Road, Nature, Construction, Obstacle, and Water.

The methodology enables a staged training regime: the model is first exposed to mild transformations and then progressively to more complex object-level perturbations. This increases robustness to seasonal and illumination shifts. Experimental results show that AugUAV stabilizes early training epochs and achieves mIoU = 0.6264 and FWIoU = 0.9565 with moderate resource consumption, improving quantitative metrics compared to standard augmentation pipelines.

The findings support the underlying hypothesis: harmonizing augmentations effectively mitigates the synthetic-to-real gap and makes AugUAV suitable for training semantic segmentation systems targeting key objects in UAV aerial imagery.

РЕЗЮМЕ

В.Ю. Устенко, М.В. Близно

Методика расширения данных для сегментации ключевых объектов на аэрофотоснимках БПЛА

Семантическая сегментация городских сцен с БПЛА осложняется доменным разрывом между реальными и синтетическими данными SynDrope и высокими требованиями к вычислительным ресурсам. В работе исследуется сочетание предобученного StripNet-backbone и адаптивного FPN-декодера с предложенной авторами методики расширения данных AugUAV для задачи семантической сегментации (для пяти ключевых классов: Road, Nature, Construction, Obstacle, Water). Данная методика позволяет сначала мягко обучить модель, затем постепенно вводить сложные объектные трансформации, что повышает устойчивость к сезонным и световым сдвигам. Эксперименты показывают, что AugUAV стабилизирует ранние эпохи и обеспечивает $mIoU=0.6264$, $FWIoU=0.9565$ при умеренном потреблении ресурсов, демонстрируя улучшение количественных показателей по сравнению со стандартной методикой расширения данных. Результаты подтверждают гипотезу: гармонизация аугментаций действительно помогает сгладить синтетикореальный разрыв и делает AugUAV пригодной для использования при обучении систем семантической сегментации ключевых объектов на аэрофотоснимках.

Устенко Владимир Юрьевич – младший научный сотрудник, лаборатория интеллектуальных систем и анализа данных, аспирант, Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта». *Область научных интересов:* системы распознавания образов, эл. почта ustenko.vova@ya.ru, адрес: 283048, Российская Федерация, г. Донецк, ул. Артема, д. 118 б.

Близно Максим Витальевич – младший научный сотрудник, лаборатория интеллектуальных систем и анализа данных, Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта». *Область научных интересов:* системы распознавания образов, адрес: 283048, Российская Федерация, г. Донецк, ул. Артема, д. 118 б.

Статья поступила в редакцию 23.10.2025.

Раздел 2 *Section 2*

Математическое
моделирование,
численные методы
и комплексы программ

*Mathematical modeling,
numerical methods
and software packages*

УДК 004.8+533.65.013.622

DOI 10.24412/2413-7383-2025-4-39-99-111

Д. Л. Остапущенко, А. Э. Воронов, В. В. Черных, М. О. Петренко
Федеральное государственное бюджетное образовательное учреждение высшего
образования «Луганский государственный университет имени Владимира Даля»
291034, Луганская Народная Республика, город Луганск, квартал Молодежный, дом 20А

ЧИСЛЕННОЕ РЕШЕНИЕ ЗАДАЧИ АЭРОДИНАМИКИ ПРИ МОДЕЛИРОВАНИИ ПОЛЕТА БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ С ИНТЕЛЛЕКТУАЛЬНЫМИ СИСТЕМАМИ УПРАВЛЕНИЯ

D. L. Ostapushchenko, A. E. Voronov, V. V. Chernykh, M. O. Petrenko
Federal State Budgetary Educational Institution of Higher Education
«Lugansk Vladimir Dahl State University»
291034, Lugansk People's Republic, Lugansk, Molodezhny District, 20A

NUMERICAL SOLUTION OF THE AERODYNAMICS PROBLEM IN MODELING THE FLIGHT OF UNMANNED AERIAL VEHICLES WITH INTELLIGENT CONTROL SYSTEMS

Разработка интеллектуальных систем управления беспилотными летательными аппаратами, особенно тех, которые используют методы обучения с подкреплением, требует физически адекватной модели объекта управления. Работа посвящена построению математической модели процесса обтекания беспилотного летательного аппарата воздушной средой, основанная на теории потенциала, теории и численных методах решения интегральных уравнений. Данная модель предназначена для определения фазовой траектории летательного аппарата в режиме реального времени на различных этапах создания интеллектуальных систем управления полетом.

Ключевые слова: беспилотный летательный аппарат, интеллектуальные системы управления, искусственный интеллект, обучение с подкреплением, аэродинамика, интегральные уравнения, динамика полета

Developing intelligent control systems for unmanned aerial vehicles, especially those using reinforcement learning methods, requires a physically adequate model of the controlled object. The work is devoted to the construction of a mathematical model of the process of air flow around an unmanned aerial vehicle, based on potential theory, theory and numerical methods for solving integral equations. This model is designed to determine the phase trajectory of an aircraft in real time at various stages of the creation of intelligent flight control systems.

Keywords: unmanned aerial vehicle, intelligent control systems, artificial intelligence, reinforcement learning, aerodynamics, integral equations, flight dynamics

Введение

Одной из основных задач при разработке беспилотных летательных аппаратов (БПЛА) является создание автоматических систем управления полетом. Наиболее гибкие механизмы автоматического управления БПЛА при выполнении различных задач в сложных условиях могут быть обеспечены с использованием интеллектуальных систем управления (ИСУ).

Во многих случаях для создания ИСУ используются методы искусственного интеллекта (ИИ), входящие в группу методов, основанных на обучении с подкреплением (*RL, Reinforcement Learning,*) [1-4]. Эти методы предполагают использование на этапе поиска оптимальной программы управления физически адекватной модели объекта управления, позволяющей осуществлять многократное решение задачи динамики объекта при различных программах управления в условиях приближенных к реальности в режиме реального времени.

Целью данной работы является построение физически адекватной математической модели процесса обтекания беспилотных летательных аппаратов воздушной средой, предназначенной для определения сил и моментов при расчете фазовой траектории летательного аппарата в режиме реального времени на различных этапах создания и использования интеллектуальных систем управления полетом.

В большинстве случаев БПЛА представляет собой твердое тело [5], положение которого в пространстве характеризуется шестью степенями свободы. Движение такого тела хорошо описывается решением второго закона Ньютона и динамического уравнения Эйлера совместно с уравнениями, описывающими взаимодействие поверхности тела с обтекающими его потоками среды [6-10]. Решение задачи взаимодействия поверхности тела с обтекающими его потоками среды ставит своей целью определение сил и моментов аэродинамической природы, вклад которых оказывает существенное влияние на динамику движения аппарата в пределах атмосферы. Мгновенная картина обтекания определяется распределением скоростей движения по поверхности, ограничивающей тело. При этом на поверхности должны выполняться граничные условия, в простейшем случае условия непротекания.

Представляется перспективным при решении задачи обтекания в режиме реального времени использовать панельные методы, основанные на теории потенциала. При использовании данной группы методов предполагается, что поверхность тела вносит искажение в первичный невозмущенный поток среды за счет поля скоростей, создаваемых вторичными источниками (двойной слой источников) [11]. Вторичные источники располагаются на поверхности тела и некоторых дополнительных поверхностях, определяемых спецификой используемой расчетной схемы. При этом решение задачи нахождения их плотности сводится к численному решению поверхностного интегрального уравнения, которое путем разбиения поверхности тела на элементарные площадки сводится к системе алгебраических уравнений.

Постановка задачи исследования

Рассмотрим задачу обтекания летательного аппарата воздушной средой в связанной системе координат $Oxuz$. Внутренний объем летательного аппарата обозначим Ω . Относительно характера взаимодействия поверхности летательного аппарата $\partial\Omega$ с набегающим потоком воздуха на данном этапе выполнения работы примем ряд предположений. Будем считать невозмущенный набегающий воздушный поток \vec{V}^∞ однородным и постоянным во времени. Будем считать, что он равен по величине воздушной

скорости, и противоположен ей по направлению. Примем гипотезу стационарности, согласно которой будем считать, что силы и моменты, действующие на летательный аппарат в неустановившемся полете, совпадают с силами и моментами, которые бы действовали на него в стационарном воздушном потоке. Угловыми движениями летательного аппарата пренебрегаем. Среду будем считать несжимаемой и идеальной, т.е. вязкостью также пренебрежем.

Методы решения задачи исследования

Методы исследований, применяемые в работе, основываются на современных достижениях в области аэродинамики и динамики полета летательных аппаратов, векторном и тензорном анализе, теории потенциала, теории и численных методах решения дифференциальных и интегральных уравнений.

Возмущенный летательным аппаратом воздушный поток представим в виде суммы набегающего воздушного потока \vec{V}^∞ и возмущения \vec{V}^* , которое будем считать потенциальным

$$\vec{V} = \vec{V}^* + \vec{V}^\infty = \text{grad } \varphi = \text{grad } \varphi^* + \text{grad } \varphi^\infty; \varphi^\infty = \vec{V}^\infty \cdot \vec{r} + C. \quad (1)$$

Учитывая уравнение неразрывности несжимаемой среды $\text{div } \vec{V} = 0$, видим, что потенциалы φ , φ^* и φ^∞ удовлетворяют уравнению Лапласа

$$\Delta \varphi = \Delta \varphi^* + \Delta \varphi^\infty = \Delta \varphi^* = 0. \quad (2)$$

На поверхности $\partial\Omega$ выполняется условие непротекания

$$\vec{V} \cdot \vec{n} = \text{grad } \varphi \cdot \vec{n} = \frac{\partial \varphi}{\partial n} = 0. \quad (3)$$

Потенциал во внутренних точках Ω будем считать постоянным (для определенности нулевым), это соответствует тому, что внутренние точки Ω являются неподвижными в связанной системе координат $Oxuz$. Тогда, при подходе к поверхности $\partial\Omega$ с внутренней стороны, нормальная производная такого потенциала также обращается в ноль.

Расположим на поверхности $\partial\Omega$ двойной слой с поверхностной плотностью τ . Его потенциал в точке наблюдения Q задается формулой

$$\varphi(Q) = \frac{1}{4\pi} \iint_{\partial\Omega} \tau(M) \frac{\vec{n}_M \cdot \vec{r}_{MQ}}{r_{MQ}^3} dS_M + \varphi^\infty(Q), \quad (2)$$

где \vec{r}_{MQ} – вектор, проведенный из точки интегрирования M на поверхности $\partial\Omega$ в точку наблюдения Q ; $r_{MQ} = \|\vec{r}_{MQ}\|$; \vec{n}_M – внешняя нормаль к поверхности $\partial\Omega$ в точке интегрирования M .

Потенциал двойного слоя удовлетворяет уравнению Лапласа во всех точках вне поверхности $\partial\Omega$, имеет непрерывную нормальную производную. При переходе через поверхность потенциал двойного слоя претерпевает разрывы и его предельные значения в точке наблюдения Q , помещенной на поверхность $\partial\Omega$, описываются соотношениями [16], [17]:

$$\lim_{P \rightarrow Q} \varphi(P) = \pm \frac{\tau(Q)}{2} + \frac{1}{4\pi} \iint_{\partial\Omega} \tau(M) \frac{\vec{n}_M \cdot \vec{r}_{MQ}}{r_{MQ}^3} dS_M, \quad (2)$$

где знак «+» берется при стремлении $P \rightarrow Q$ с внешней стороны поверхности $\partial\Omega$; «-» – с внутренней. Во внутренних точках области Ω потенциал φ тождественно равен нулю, поэтому

$$\lim_{P \rightarrow Q} \varphi(P) = -\frac{\tau(Q)}{2} + \frac{1}{4\pi} \iint_{\partial\Omega} \tau(M) \frac{\vec{n}_M \cdot \vec{r}_{MQ}}{r_{MQ}^3} dS_M + \varphi^\infty(Q) = 0, \quad (3)$$

что позволяет записать уравнение Фредгольма второго рода относительно плотности двойного слоя τ

$$\tau(Q) = \frac{1}{2\pi} \iint_{\partial\Omega} \tau(M) \frac{\vec{n}_M \cdot \vec{r}_{MQ}}{r_{MQ}^3} dS_M + 2\varphi^\infty(Q). \quad (4)$$

После нахождения из данного уравнения поверхностной плотности двойного слоя τ скорость движения среды в точке Q окружающего тело пространства может быть вычислена по формуле

$$\vec{V}(Q) = \frac{1}{4\pi} \text{grad}_Q \iint_{\partial\Omega} \tau(M) \frac{\vec{n}_M \cdot \vec{r}_{MQ}}{r_{MQ}^3} dS_M + \vec{V}^\infty(Q). \quad (5)$$

Для численного решения уравнения (4) приближенно представим поверхность $\partial\Omega$ в виде теоретико-множественного объединения плоских элементарных площадок Δ_i

$$\partial\Omega = \cup_{i=1}^N \Delta_i, \quad (6)$$

пересекающихся только по своим границам

$$\Delta_i \cap \Delta_j \in \partial\Delta_i \cap \partial\Delta_j \text{ при } i \neq j. \quad (7)$$

Пример дискретной геометрической модели представлен на рис. 1.

В данной работе в качестве элементарных площадок используются треугольники. Будем считать плотность двойного слоя в пределах элементарных площадок постоянной. Поочередно помещая точку наблюдения Q_i в центры элементарных площадок Δ_i , получим

$$\tau_i = \frac{1}{2\pi} \sum_{j=1}^N \tau_j \iint_{\Delta_j} \frac{\vec{n}_M \cdot \vec{r}_{MQ_i}}{r_{MQ_i}^3} dS_M + 2\varphi_i^\infty, \quad (8)$$

$$\tau_i = \tau(Q_i), \varphi_i^\infty = \varphi(Q_i), i = \overline{1, N}.$$

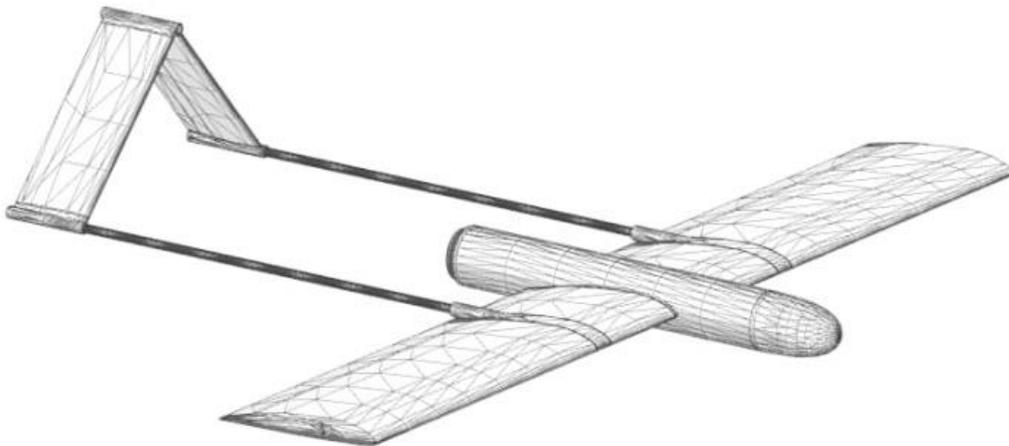


Рисунок 1 – Дискретная геометрическая модель летательного аппарата

Тогда при $i = j$

$$\iint_{\Delta_j} \frac{\vec{n}_M \cdot \vec{r}_{MQ_i}}{r_{MQ_i}^3} dS_M = 0. \quad (9)$$

В свою очередь при $i \neq j$

$$\iint_{\Delta_j} \frac{\vec{n}_M \cdot \vec{r}_{MQ_i}}{r_{MQ_i}^3} dS_M = -\theta(Q_i, \Delta_j), \quad (10)$$

где $\theta(Q, \Delta)$ – величина телесного угла, под которым виден треугольник Δ из точки наблюдения Q .

Для пространственного треугольника ABC и точки наблюдения Q значение величины телесного угла могут быть найдены из соотношений:

$$T = \frac{(\vec{r}_{QA}, \vec{r}_{QB}, \vec{r}_{QC})}{r_{QA}r_{QB}r_{QC} + (\vec{r}_{QA} \cdot \vec{r}_{QB})r_{QC} + (\vec{r}_{QB} \cdot \vec{r}_{QC})r_{QA} + (\vec{r}_{QC} \cdot \vec{r}_{QA})r_{QB}}, \quad (11)$$

$$\theta(Q, \Delta) = 2 \arctan(T).$$

Вводя обозначения

$$\mathbf{x} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \dots \\ \tau_N \end{pmatrix}; \mathbf{A} = \begin{pmatrix} 1 & a_{12} & \dots & a_{1N} \\ a_{21} & 1 & \dots & a_{2N} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{pmatrix}; \mathbf{b} = \begin{pmatrix} \varphi_1^\infty \\ \varphi_2^\infty \\ \dots \\ \varphi_N^\infty \end{pmatrix}, \quad (12)$$

где

$$a_{ij} = \frac{1}{2\pi} \theta(Q_i, \Delta_j), \quad i, j = \overline{1, N}, \quad i \neq j, \quad (13)$$

систему уравнений можно записать в матричном виде

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}. \quad (14)$$

Для решения полученной системы используется метод минимизации обобщенной невязки GMRES, который относится к итерационным методам и позволяет при принятии в качестве начального приближения достаточно близкого к решению вектора \mathbf{x} сделать процедуру решения достаточно эффективной с вычислительной точки зрения. Это позволяет на каждом временном шаге численного решения динамического уравнения брать в качестве начального приближение вектор \mathbf{x} , вычисленный на прошлом временном шаге.

После решения полученной системы уравнений и нахождения τ_i скорость движения среды может быть найдена по формуле

$$\vec{V}(Q) = \frac{1}{4\pi} \text{grag}_Q \sum_{i=1}^N \tau_i \iint_{\Delta_j} \frac{\vec{n}_M \cdot \vec{r}_{MQ}}{r_{MQ}^3} dS_M + V(Q), \quad (15)$$

которая может быть преобразована к виду

$$\vec{V}(Q) = -\frac{1}{4\pi} \sum_{i=1}^N \tau_i \oint_{\partial \Delta_i} \frac{d\vec{l}_M \times \vec{r}_{MQ}}{r_{MQ}^3} + V(Q), \quad (16)$$

где интегрирование осуществляется по сторонам треугольника.

Вычисление интеграла по одной из сторон треугольника ABC может быть осуществлено аналитически. Запишем формулу для стороны AB

$$\frac{\tau}{4\pi} \int_A^B \frac{d\vec{l}_M \times \vec{r}_{MQ}}{r_{MQ}^3} = \tau \frac{\cos \alpha_{AQ} - \cos \alpha_{BQ}}{4\pi r} \vec{e}, \quad (17)$$

где

$$\vec{s} = \vec{r}_{AB}/r_{AB}, \quad \vec{r} = \vec{r}_{AQ} - (\vec{r}_{AQ} \cdot \vec{s})\vec{s}, \quad \vec{n} = \vec{r}/r, \quad \vec{e} = \vec{s} \times \vec{n}, \quad (18)$$

$$\cos \alpha_{AQ} = (\vec{r}_{AQ} \cdot \vec{s})/r_{AQ}, \quad \cos \alpha_{BQ} = (\vec{r}_{BQ} \cdot \vec{s})/r_{BQ}.$$

Соответствующие формулы для остальных сторон могут быть получены путем циклической подстановки наименований вершин.

Для нахождения сил и моментов воспользуемся уравнением Бернулли [13-15] для струйки воздуха, обтекающей тело, выбрав два сечения струйки: одно на бесконечности, второе в точке M вблизи поверхности $\partial \Omega$

$$p^\infty + \frac{\rho(V^\infty)^2}{2} = p(M) + \frac{\rho(V(M))^2}{2}. \quad (19)$$

Данное соотношение удобно переписать в виде

$$p(M) = \frac{\rho(V^\infty)^2}{2} \left(1 - \frac{(V(M))^2}{(V^\infty)^2} \right) + p^\infty, \quad (20)$$

где величина

$$c_p(M) = 1 - \frac{(V(M))^2}{(V^\infty)^2} \quad (21)$$

называется коэффициентом давления.

Тогда полная аэродинамическая сила и момент в связанной системе координат $Oxuz$ могут быть найдены по формулам

$$\vec{F}^a = - \oint_{\partial\Omega} p(M) \vec{n}_M dS_M \text{ и } \vec{M}^a = - \oint_{\partial\Omega} p(M) \vec{r}_M \times \vec{n}_M dS_M. \quad (22)$$

Для приближенного учета возникающих на поверхности летательного аппарата касательных сил воспользуемся следующей эвристической конструкцией. Будем считать, что касательные напряжения направлены по вектору скорости потока вблизи поверхности летательного аппарата и подчиняются закону Амонтона-Кулона. Вводя в рассмотрение вектор $\vec{v}_M = \vec{V}(M)/V(M)$, можно записать

$$\vec{F}^a = \oint_{\partial\Omega} \mu(V(M)) p(M) \vec{v}_M dS_M \text{ и } \vec{M}^a = \oint_{\partial\Omega} \mu(V(M)) p(M) \vec{r}_M \times \vec{v}_M dS_M, \quad (23)$$

где $\mu(V)$ – эвристическая зависимость коэффициента трения от абсолютной величины скорости обтекания, в простейшем случае $\mu = const$.

Полученные результаты исследования

Для верификации результатов решения задачи обтекания тела воздушной средой проведем сравнение результатов расчета, полученных с использованием описанной модели, с результатами, полученными аналитически [19].

В качестве первого примера рассмотрим обтекание шара радиусом a потоком невязкой жидкости или невязкого газа. Осесимметричное решение в системе координат $RO\theta$ имеет вид:

$$V_R = V^\infty \left[1 - \left(\frac{a}{R} \right)^3 \right] \cos(\theta); V_\theta = V^\infty \left[1 + \frac{1}{2} \left(\frac{a}{R} \right)^3 \right] \sin(\theta). \quad (24)$$

Распределение скорости обтекания на поверхности сферы определяется зависимостями

$$V_R = 0; V_\theta = \frac{3}{2} V^\infty \sin(\theta). \quad (25)$$

Следовательно, коэффициент давления определяется формулой

$$c_p = 1 - \frac{(V_\theta)^2}{(V^\infty)^2} = 1 - \frac{9}{4} \sin^2(\theta). \quad (26)$$

Осуществим расчет поля скоростей для шара радиусом $a = 1$ м при $V^\infty = 100$ м/с. Дискретная геометрическая модель поверхности шара и распределение поля скоростей обтекающей его воздушной среды представлены на рис. 2а. Для установления погрешностей вычисления определялись скорость обтекания вблизи поверхности и распределение коэффициента давления. Эпюра коэффициента давления представлена на рисунке 2б. Расчет проводился при различном числе разбиений поверхности шара. Поскольку рассчитываемые величины проходят через нулевое значение, определялись их абсолютные погрешности, значения которых приведены в таблице 1.

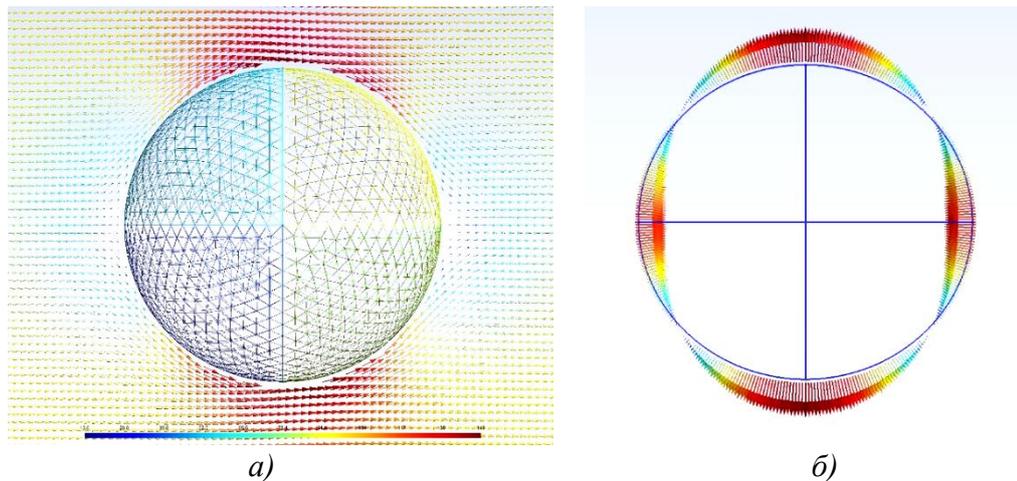


Рисунок 2 – Результаты решения задачи обтекания шара воздушной средой: *a* – дискретная геометрическая модель поверхности шара и распределение поля скоростей воздушной среды; *б* – эпюра коэффициента давления

Таблица 1 – Погрешности расчета поля скоростей и коэффициента давления

Количество элементов разбиения	12 552	48 062	57 722
Абсолютная погрешность скорости, м/с	10,35	5,84	5,81
Абсолютная погрешность коэффициента давления	0,297	0,107	0,099

В качестве второго примера рассмотрим обтекание цилиндра радиусом a потоком невязкой жидкости или невязкого газа. Решение в системе координат $RO\theta$ имеет вид:

$$V_R = V^\infty \left[1 - \left(\frac{a}{R}\right)^2 \right] \cos(\theta); V_\theta = V^\infty \left[1 + \left(\frac{a}{R}\right)^2 \right] \sin(\theta). \quad (27)$$

Скорость на поверхности цилиндра:

$$V_R = 0; V_\theta = 2V^\infty \sin(\theta). \quad (28)$$

Следовательно, коэффициент давления определяется формулой

$$c_p = 1 - \frac{(V_\theta)^2}{(V^\infty)^2} = 1 - 4 \sin^2(\theta). \quad (29)$$

Расчет проводился для цилиндра радиусом $a = 1$ м при $V^\infty = 100$ м/с. Абсолютные погрешности расчета скорости и коэффициента давления не превысили 5 м/с и 0,3 соответственно. Дискретная геометрическая модель поверхности цилиндра и результат расчета распределения поля скоростей воздушной среды представлены на рис. 3.

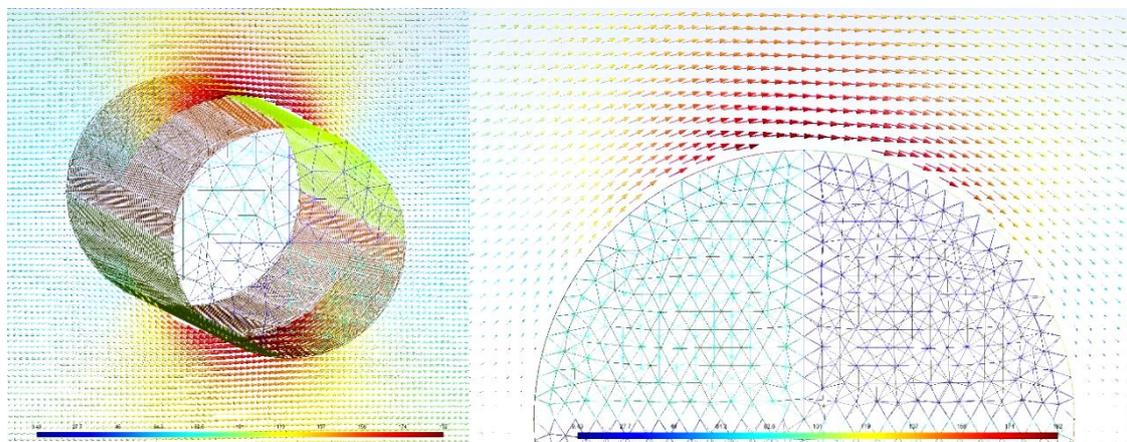


Рисунок 3 – Дискретная геометрическая модель поверхности цилиндра и результат расчета распределения поля скоростей воздушной среды

Достаточно высокая точность совпадения численных расчетов с результатами аналитических расчетов, проведенных на этапе верификации, позволяет сделать вывод о том, что реализованная в работе модель может быть использована при решении практически значимых задач, для которых не существует аналитических решений.

Проведем численный расчет для задачи обтекания воздушным потоком при $V^\infty = 100$ м/с фюзеляжа БпЛА. Дискретная геометрическая модель и результаты расчета поля скоростей представлены на рис. 4.

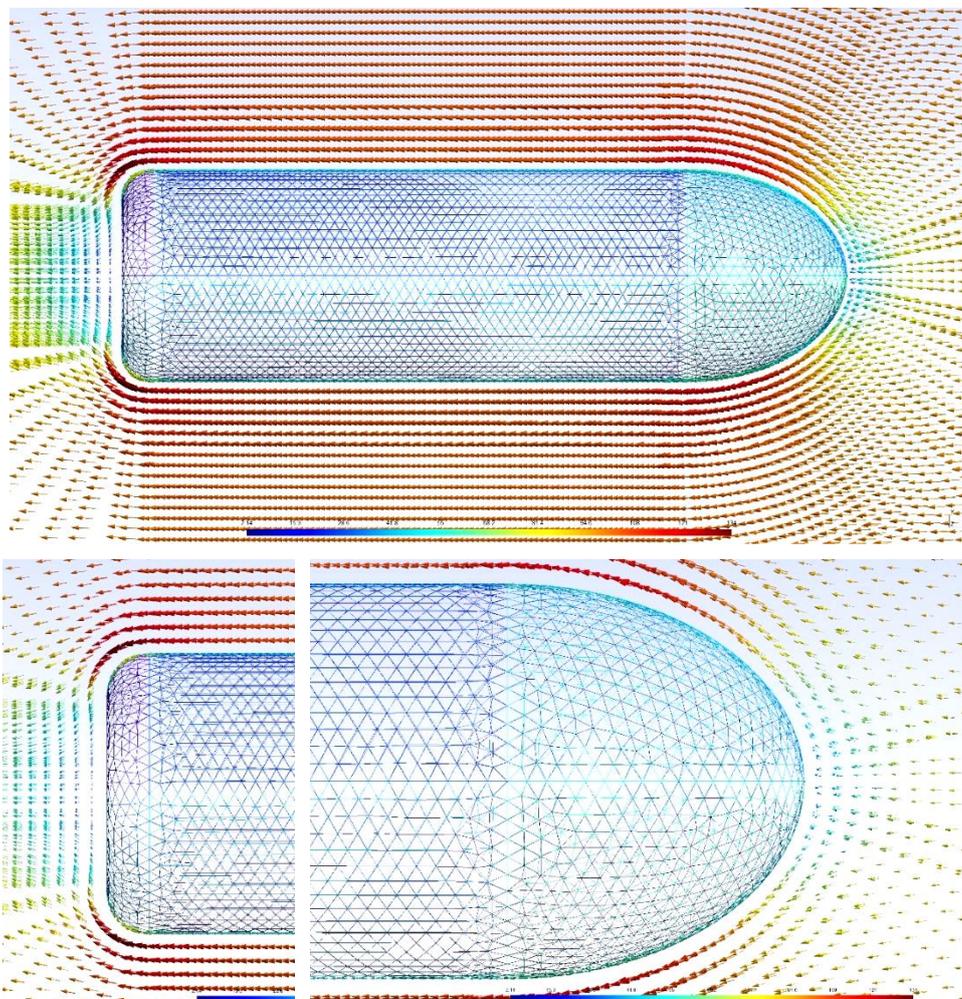


Рисунок 4 – Дискретная геометрическая модель поверхности фюзеляжа и результат расчета распределения поля скоростей воздушной среды

Заключение

В настоящей работе получили дальнейшее развитие методы синтеза управляющих программ беспилотных летательных аппаратов, основанные на применении технологий машинного обучения с подкреплением, за счёт построения математической модели, включающей в себя рациональное сочетание приближенных бессеточных вычислительных схем и эвристических приемов, и позволяющей осуществлять решение задач аэробаллистики в режиме реального времени с учетом реальной формы поверхности летательного аппарата. Помимо этого, по итогам проведенных исследований были достигнуты следующие результаты.

1. Построена физически адекватная математическая модель процесса обтекания беспилотного летательного аппарата воздушной средой, основанная на теории потенциала, теории и численных методах решения интегральных уравнений.

2. Полученная модель предназначена для определения фазовой траектории летательного аппарата для различных зависимостей значений управляющих параметров от времени, а, следовательно, различных программ управления. На основании результатов моделирования становится возможным осуществление приближенного оценивания соответствия фазовой траектории летательного аппарата заданию на полет, исходя из заранее определенных метрик.

3. Построенная модель предназначена для использования как на этапе предполетной подготовки при синтезе программ управления летательным аппаратом с использованием обучения с подкреплением, так и в процессе полета с целью прогнозирования полетной траектории бортовым оборудованием летательного аппарата при использовании методов обучения с подкреплением, основанных на моделях.

4. Является актуальным дальнейшее развитие математической модели взаимодействия поверхности летательного аппарата с воздушной средой за счет учета в ней большего числа физически значимых факторов, например, более точного учета вязкости воздушной среды, углового движения летательного аппарата, завихрений потока, явления срыва потока и т.д.

5. Поскольку в рамках предлагаемого подхода предполагается многократное применение математической модели, при ее развитии необходимо тщательно соблюсти баланс между требованиями физической адекватности и требуемыми для осуществления процесса моделирования вычислительными ресурсами. Существенным требованием является возможность выполнения расчетов в режиме реального времени.

Список литературы

1. Саттон Р.С. Обучение с подкреплением / Р.С. Саттон, Э.Г. Барто: пер. с англ. М.: БИНОМ. Лаборатория знаний, 2017. 399 с.
2. Лапань М. Глубокое обучение с подкреплением. AlphaGo и другие технологии / М. Лапань. СПб.: Питер, 2020. 496 с.
3. Уиндер Ф. Обучение с подкреплением для реальных задач /Ф. Уиндер: пер. с англ. СПб.: БХВ-Петербург, 2023. 400 с.
4. Моралес М. Грокаем Глубокое обучение с подкреплением / М. Моралес. СПб.: Питер, 2023. 464 с.
5. Остапущенко Д.Л., Воронов А.Э., Петренко М.О., Черных В.В. Численное решение динамических уравнений движения летательного аппарата в задачах разработки интеллектуальных систем управления полетом // Вестник Луганского государственного университета имени Владимира Даля. 2024. № 8(86). С. 35-42.
6. Боднер В.А. Теория автоматического управления полетом / В.А. Боднер. М.: Наука, 1964. 700 с.
7. Дрябкин В.В. Бортовые системы управления полетом / Ю.В. Байбородин, В.В. Дрябкин, Е.Г. Сменковский, С.Г. Унгурян. М.: Транспорт, 1975. –336 с.
8. Жаков А.М. Управление баллистическими ракетами и космическими объектами / А.М. Жаков. – М.: Воениздат, 1974. 261 с.
9. Павловский М.А. Теоретическая механика. Динамика: Учебник / М.А. Павловский, Л.Ю. Акинфиева, О.Ф. Бойчук. К.: Выща шк., 1990. 480 с.
10. Федорченко А.М. Теоретическая физика. Классическая механика / А.М. Федорченко. К.: Вища школа. Головное изд-во, 1983. 351 с.
11. Гальченко В.Я. Численный анализ пространственной конфигурации магнитных полей объектов сложной геометрической формы с учетом нелинейных характеристик веществ / В.Я. Гальченко, Д.Л. Остапущенко // Информационные технологии. 2008. №8. С. 43-49.
12. Афонин П.М. Беспилотные летательные аппараты / П.М. Афонин, И.С. Голубев, Н.И. Колоткови др.; Под ред. Л.С. Чернобровкина. М.: Машиностроение. – 440 с.

13. Павленко В.Ф. Боевая авиационная техника: Летательные аппараты, силовые установки и их эксплуатация / В.Ф. Павленко, А.А. Дьяченко, В.И. Жулев и др.; Под ред. В.Ф. Павленко. М.: Воениздат, 1984. 320 с.
14. Самойлович Г.С. Гидроаэромеханика: Учебник для вузов / Г.С. Самойлович. М.: Машиностроение, 1980. 280 с.
15. Голубева О.В. Курс механики сплошных сред. Учеб. Пособие для педвузов / О.В. Голубева. М.: Высшая школа, 1972. 368 с.
16. Курбатов П.А. Численный расчет электромагнитных полей / П.А. Курбатов, С.А. Аринчин. М.: Энергоатомиздат, 1984. 168 с.
17. Припадчев А.Д., Горбунов А.А., Магдин А.Г. Моделирование устойчивости и управляемости летательных аппаратов. = Инфра-Инженерия, 2021. 116 с.
18. О.Е. Лукьянов, Д.В. Золотов, О.У. Эспиноса Барсенас, В.А. Комаров Определение аэродинамических характеристик малоразмерных беспилотных летательных аппаратов в лётном эксперименте // Вестник Самарского университета. Аэрокосмическая техника, технологии и машиностроение. 2023. №3. URL: <https://cyberleninka.ru/article/n/opredelenie-aerodinamicheskikh-harakteristik-malorazmernih-bespilotnyh-letatelnyh-apparatov-v-lyotnom-eksperimente> (дата обращения: 02.12.2025).
19. ГОСТ Р 57700.8-2018. Численное моделирование физических процессов. Численное моделирование дозвуковых течений вязких жидкостей и газов. Верификация ПО: национальный стандарт Российской Федерации: издание официальное : утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 6 февраля 2018 г. № 48-ст. М.: Стандартинформ, 2018. 15 с.
20. Уланов, С.И. Анализ возможностей применения летательных аппаратов в контроле состояния местности / С.И. Уланов, О.А. Криводубский, А.А. Никитина Проблемы искусственного интеллекта (International Peer-Reviewed Scientific Journal «Problems of Artificial Intelligence»), ISSN 2413-7383. 2023. № 3(30).
21. Моржов В.И., Ермачков Ю.А. Моделирование физических процессов в авиации. Учебное пособие – Инфра-Инженерия, 2021. 160 с.
22. Уланов, С.И. Анализ современных разработок беспилотных летательных аппаратов / С.И. Уланов, О.А. Криводубский, А.А. Никитина. Проблемы искусственного интеллекта (International Peer-Reviewed Scientific Journal «Problems of Artificial Intelligence»), ISSN 2413-7383. 2024. № 4(35).
23. Исмаилов Куат Кайратович Определение аэродинамических характеристик беспилотного летательного аппарата самолетного типа аналитическими методами // Вестн. Том. гос. ун-та. Математика и механика. 2022. №78. URL: <https://cyberleninka.ru/article/n/opredelenie-aerodinamicheskikh-harakteristik-bespilotnogo-letatelno-go-apparata-samoletnogo-tipa-analiticheskimi-metodami> (дата обращения: 02.12.2025).
24. Савищенко Николай Петрович, Апевалов Игорь Владимирович, Дёма Игорь Анатольевич, Попов Александр Сергеевич Экспериментальные исследования аэродинамических характеристик беспилотного летательного аппарата // Известия ТулГУ. Технические науки. 2020. №2. URL: <https://cyberleninka.ru/article/n/eksperimentalnye-issledovaniya-aerodinamicheskikh-harakteristik-bespilotnogo-letatelno-go-apparata> (дата обращения: 02.12.2025).
25. Гвоздева В.А. Интеллектуальные технологии в беспилотных системах. Учебник. ИНФРА-М, 2021. 176 с.

References

1. Sutton, R.S. Reinforcement Learning, E.G. Barto, translated from English. Moscow: BINOM. Knowledge Laboratory, 2017, 399 p.
2. Lapan, M. Deep Reinforcement Learning: AlphaGo and Other Technologies. St. Petersburg: Piter, 2020, 496 p.
3. Winder, F. Reinforcement Learning for Real-World Problems. St. Petersburg: BHV-Petersburg, 2023, 400 p.
4. Morales, M. Grokaem: Deep Reinforcement Learning. St. Petersburg: Piter, 2023, 464 p.
5. Ostapushchenko D.L., Voronov A.E., Petrenko M.O., Chernykh V.V. Numerical solution of dynamic equations of aircraft motion in problems of developing intelligent flight control systems // Bulletin of Vladimir Dahl Luhansk State University. - 2024. - No. 8 (86). - P. 35-42.
6. Bodner V.A. Theory of automatic flight control / V.A. Bodner. - Moscow: Nauka, 1964. - 700 p.
7. Dryabkin V.V. On-board flight control systems / Yu.V. Bayborodin, V.V. Drabkin, E.G. Smenkovsky, S.G. Unguryan. - Moscow: Transport, 1975. - 336 p.
8. Zhakov A.M. Control of Ballistic Missiles and Space Objects / A. M. Zhakov. Moscow: Voениzdat, 1974. 261 p.
9. Pavlovsky M. A. Theoretical Mechanics. Dynamics: Textbook / M. A. Pavlovsky, L. Yu. Akinfieva, O. F. Boychuk. - Kiev: Vyshcha shkola, 1990. - 480 p.
10. Fedorchenko A. M. Theoretical Physics. Classical Mechanics / A. M. Fedorchenko. - Kiev: Vyshcha shkola. Main publishing house, 1983. - 351 p.

11. Galchenko V. Ya. Numerical Analysis of the Spatial Configuration of Magnetic Fields of Objects of Complex Geometric Shape Taking into Account the Nonlinear Characteristics of Substances / V. Ya. Galchenko, D. L. Ostapushchenko // Information Technologies. – 2008. – №8. – P. 43-49.
12. Afonin P.M. Unmanned Aerial Vehicles / P.M. Afonin, I.S. Golubev, N.I. Kolotkov, and others; Ed. by L.S. Chernobrovkin. – Moscow: Mashinostroenie. – 440 p.
13. Pavlenko V.F. Combat Aviation Equipment: Aircraft, Power Plants, and Their Operation / V.F. Pavlenko, A.A. Dyachenko, V.I. Zhulev, and others; Ed. by V.F. Pavlenko. – Moscow: Voenizdat, 1984. – 320 p.
14. Samoylovich G.S. Hydroaeromechanics: Textbook for Universities / G.S. Samoylovich - Moscow: Mashinostroenie, 1980. - 280 p.
15. Golubeva O.V. Course in Continuum Mechanics. Textbook. Manual for pedagogical universities / O.V. Golubeva - Moscow: Vysshaya shkola, 1972. - 368 p.
16. Kurbatov P.A. Numerical calculation of electromagnetic fields / P.A. Kurbatov, S.A. Arinchin. - Moscow: Energoatomizdat, 1984. - 168 p.
17. Pripadchev A.D., Gorbunov A.A., Magdin A.G. Modeling of Stability and Controllability of Aircraft. = Infra-Engineering, 2021. – 116 p.
18. O.E. Lukyanov, D.V. Zolotov, O.U. Espinosa Barcenas, V.A. Komarov Determination of Aerodynamic Characteristics of Small-Sized Unmanned Aerial Vehicles in a Flight Experiment // Bulletin of Samara University. Aerospace Engineering, Technology and Mechanical Engineering. 2023. No. 3. URL: <https://cyberleninka.ru/article/n/opredelenie-aerodinamicheskikh-harakteristik-malorazmernyh-bespilotnyh-letatelnyh-apparatov-v-lyotnom-eksperimente> (date of access: 02.12.2025).
19. GOST R 57700.8-2018. Numerical modeling of physical processes. Numerical modeling of subsonic flows of viscous liquids and gases. Software verification: national standard of the Russian Federation: official publication: approved and put into effect by Order of the Federal Agency for Technical Regulation and Metrology dated February 6, 2018, No. 48-st. – Moscow: Standartinform, 2018. – 15 p.
20. Ulanov, S.I. Analysis of the possibilities of using aircraft in terrain monitoring / S.I. Ulanov, O.A. Krivodubsky, A.A. Nikitina. International Peer-Reviewed Scientific Journal "Problems of Artificial Intelligence", ISSN 2413-7383. 2023. No. 3(30).
21. Morzhov, V.I., Ermachkov, Yu.A. Modeling of Physical Processes in Aviation. Tutorial - Infra-Engineering, 2021. - 160 p.
22. Ulanov, S.I. Analysis of Modern Developments of Unmanned Aerial Vehicles / S.I. Ulanov, O.A. Krivodubsky, A.A. Nikitina. International Peer-Reviewed Scientific Journal "Problems of Artificial Intelligence", ISSN 2413-7383. 2024. No. 4(35).
23. Ismailov Kuat Kairatovich Determination of the aerodynamic characteristics of an unmanned aerial vehicle of an airplane type by analytical methods // Vestn. Tomsk. state University. Mathematics and mechanics. 2022. No. 78. URL: <https://cyberleninka.ru/article/n/opredelenie-aerodinamicheskikh-harakteristik-bespilotnogo-letatel'nogo-apparata-samoletnogo-tipa-analiticheskimi-metodami> (date of access: 02.12.2025).
24. Savishchenko Nikolay Petrovich, Apevalov Igor Vladimirovich, Dema Igor Anatolyevich, Popov Aleksandr Sergeevich Experimental studies of the aerodynamic characteristics of an unmanned aerial vehicle // Bulletin of Tula State University. Technical sciences. 2020. No. 2. URL: <https://cyberleninka.ru/article/n/eksperimentalnye-issledovaniya-aerodinamicheskikh-harakteristik-bespilotnogo-letatel'nogo-apparata> (Accessed: 02.12.2025).
25. Gvozdeva V.A. Intelligent Technologies in Unmanned Systems. Textbook. INFRA-M, 2021. – 176 p.

RESUME

D. L. Ostapushchenko, A. E. Voronov, V. V. Chernykh, M. O. Petrenko
Numerical solution of the aerodynamics problem in modeling the flight
of unmanned aerial vehicles with intelligent control systems

The development of modern unmanned aerial vehicles (UAVs) requires the creation of sophisticated automatic flight control systems capable of operating in complex and changing conditions. To achieve high flexibility and adaptability, intelligent control systems, particularly those using artificial intelligence and reinforcement learning, are increasingly being used. These methods require a physically adequate mathematical model of the controlled object to perform repeated, rapid simulations of flight dynamics under conditions close to real-world conditions. The primary objective of this study is to develop such a model that would accurately and in real time determine the aerodynamic forces and moments acting on the UAV surface.

This study utilizes a mathematical modeling approach based on potential theory and numerical methods for solving integral equations. The aerodynamic problem is formulated using a panel method, where the continuous surface of the UAV is approximated by a discrete set of planar triangular elements. To efficiently solve the resulting system of linear algebraic equations, the generalized residual minimization method (GMRES) is applied. Aerodynamic forces are calculated using the Bernoulli equation to determine the pressure coefficient, and a heuristic approach based on the Amontons-Coulomb law is introduced to estimate the tangential friction forces.

The proposed model was thoroughly verified by comparing the obtained numerical results with known analytical solutions for canonical geometric shapes: a sphere and a cylinder. The verification process showed that increasing the sampling density significantly improves accuracy. Subsequently, the model was successfully applied to calculate the velocity field distribution around the complex geometry of a UAV fuselage, confirming its practical applicability to objects for which analytical solutions do not exist.

The resulting model provides the necessary balance between physical accuracy and computational efficiency, making it suitable for real-time applications—a mandatory requirement for dynamic flight control problems. It is designed for dual use: for autonomous reinforcement learning of agents and for online trajectory prediction by onboard UAV systems.

РЕЗЮМЕ

*Д. Л. Остапущенко, А. Э. Воронов, В. В. Черных, М. О. Петренко
Численное решение задачи аэродинамики при моделировании полета
беспилотных летательных аппаратов с интеллектуальными системами
управления*

Разработка современных беспилотных летательных аппаратов (БПЛА) требует создания совершенных систем автоматического управления полетом, способных функционировать в сложных и меняющихся условиях. Для достижения высокой гибкости и адаптивности все чаще применяются интеллектуальные системы управления, в частности те, которые используют методы искусственного интеллекта и обучения с подкреплением. Эти методы требуют наличия физически адекватной математической модели объекта управления для выполнения многократного быстрого моделирования динамики полета в условиях, приближенных к реальным. Основная задача, решаемая в данной работе, заключается в построении такой модели, которая позволяла бы точно и в режиме реального времени определять аэродинамические силы и моменты, действующие на поверхность БПЛА.

В исследовании используется подход математического моделирования, основанный на теории потенциала и численных методах решения интегральных уравнений. Аэродинамическая задача сформулирована с использованием панельного метода, где непрерывная поверхность БПЛА аппроксимируется дискретным набором плоских треугольных элементов. Для эффективного решения полученной системы линейных алгебраических уравнений применяется метод минимизации обобщенной невязки (GMRES). Аэродинамические силы рассчитываются с использованием уравнения Бернулли для определения коэффициента давления, а для оценки касательных сил трения введен эвристический подход, основанный на законе Амонтона-Кулона.

Предложенная модель была тщательно верифицирована путем сравнения полученных численных результатов с известными аналитическими решениями для канонических геометрических форм: шара и цилиндра. Процесс верификации показал, что увеличение плотности дискретизации значительно повышает точность. Впоследствии модель

была успешно применена для расчета распределения поля скоростей вокруг сложной геометрии фюзеляжа БПЛА, что подтвердило ее практическую применимость для объектов, для которых не существует аналитических решений.

Полученная модель обеспечивает необходимый баланс между физической точностью и вычислительной эффективностью, что делает ее пригодной для приложений реального времени — обязательного требования для задач динамического управления полетом. Она предназначена для двойного использования: при автономном обучении агентов с подкреплением и для онлайн-прогнозировании траектории бортовыми системами БПЛА.

Остапушенко Дмитрий Леонидович – доцент кафедры прикладной математики, заведующий МНИЛ «Лаборатория интеллектуальных систем управления», ФГБОУ ВО «Луганский государственный университет имени Владимира Даля», 291034, Луганская Народная Республика, г. Луганск, квартал Молодежный, дом 20А, телефон +7(959)560-35-93, ostapuschenko.dmitriy@mail.ru. *Область научных интересов:* математическое моделирование, неразрушающий контроль, численные методы решения задач электродинамики и аэродинамики, эвристические и мультиагентные алгоритмы, поведенческий искусственный интеллект и обучение с подкреплением

Воронов Артур Эдуардович – доцент кафедры автоматизации и компьютерно-интегрированных технологий, ФГБОУ ВО «Луганский государственный университет имени Владимира Даля», 291034, Луганская Народная Республика, г. Луганск, квартал Молодежный, дом 20А, телефон +7(959)161-06-51, osooler@ua.ru. *Область научных интересов:* интерфейсы взаимодействия человека и машины, UX/UI-дизайн автоматизированных систем

Черных Виктория Валерьевна – доцент кафедры информационных и управляющих систем, ведущий научный сотрудник МНИЛ «Лаборатория интеллектуальных систем управления», ФГБОУ ВО «Луганский государственный университет имени Владимира Даля», 291034, Луганская Народная Республика, г. Луганск, квартал Молодежный, дом 20А, телефон +7(959)138-33-19, viktoria.chernykh@mail.ru. *Область научных интересов:* искусственный интеллект, информационные системы и технологии в различных сферах деятельности, интеллектуальные системы машинного обучения

Петренко Максим Олегович – младший научный сотрудник МНИЛ «Лаборатория интеллектуальных систем управления», ФГБОУ ВО «Луганский государственный университет имени Владимира Даля», 291034, Луганская Народная Республика, г. Луганск, квартал Молодежный, дом 20А, телефон +7(959) 130-31-47, max_petrenko_2002@mail.ru. *Область научных интересов:* бортовые системы беспилотных летательных аппаратов, искусственные нейронные сети, системы визуальной и инерциальной навигации

Статья поступила в редакцию 08.11.2025.

УДК 621.771.07:004.42

DOI 10.24412/2413-7383-2025-4-39-112-123

В.С. Солод

Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта», г. Донецк
283048, г. Донецк, ул. Артема, 118 б

АЛГОРИТМ И ПРОГРАММА ОПТИМИЗАЦИИ РЕЖИМА УСКОРЕННОГО ОХЛАЖДЕНИЯ АРМАТУРНЫХ ПРОФИЛЕЙ В ТЕРМОУСТАНОВКЕ КОНСТРУКЦИИ «THERMEX»

V. S. Solod

Federal State Budgetary Scientific Institution «Institute of Artificial Intelligence Problems»
283048, Donetsk, Artema str, 118-b

ALGORITHM AND PROGRAM FOR OPTIMIZING THE ACCELERATED COOLING OPERATION MODE OF REINFORCEMENT BARS IN THE THERMAL UNIT OF THE "THERMEX" DESIGN

Работа посвящена разработке программных средств оптимизации режима термоупрочнения арматурных профилей с целью получения заданных механических свойств арматуры с минимальным расходом воды на охлаждение при соблюдении технологических ограничений с учетом химического состава (химсостава) стали. Постановка задачи оптимизации с учетом химического состава (химсостава) стали и ограничений осуществлена впервые. Целевая функция – условный предел текучести, к ограничениям относятся: допускаемые давления воды на форсунках, отсутствие кипения воды в секции, наличие гидро-транспортирования проката через установку. Приведена блок-схема алгоритма, описан порядок выполнения программы, приведены примеры оптимизации режимов.

Ключевые слова: ускоренное охлаждение, арматурный профиль, математическая модель, алгоритм, программа оптимизация, предел текучести.

The objective of this study is to develop software tools for optimizing the thermal hardening process for reinforcing sections with minimal cooling water consumption while adhering to process constraints. This is the first time the optimization problem has been formulated taking into account the chemical composition of the steel and the constraints. The objective function is the yield strength, and constraints include permissible water pressures at the nozzles, the absence of water boiling in the section, and the presence of hydro transportation of the rolled product through the unit. A flowchart of the algorithm is shown, the program execution procedure is described, and examples of optimization modes are given.

Keywords: accelerated cooling, rebar section, mathematical model, algorithm, optimization program, yield strength.

Современные условия производства сортового проката характеризуются высокими скоростями прокатки, малыми паузами между деформацией на заключительной стадии процесса, высокими температурами конца прокатки (1020-1100°C). Кроме того, дефицит охлаждающей воды на ряде существующих станов и дороговизна строительства и эксплуатации систем водоснабжения, регенерации и рециркуляции для обеспечения охлаждающих установок — все это создает жесткие требования к технологии и устройствам ускоренного охлаждения, а именно: - технология охлаждения проката в технологическом потоке стана должна обеспечивать получение требуемой структуры и механических свойств металла, равномерность охлаждения по периметру, сечению и длине проката, устройства ускоренного охлаждения должны развивать высокую интенсивность теплообмена и иметь при этом малую длину.

Цель работы - разработка программных средств для оптимизации режима термоупрочнения арматурных профилей с минимальным расходом воды на охлаждение при соблюдении технологических ограничений. Все охлаждающие системы должны обеспечивать эффективное использование хладагента и его минимальный расход; - конструкция охлаждающих устройств должна способствовать транспортировке проката по технологической линии [1].

История вопроса

Вопросам моделирования ускоренного охлаждения посвящено много работ. В частности, можно отметить ранние работы Национальной металлургической академии и Института черной металлургии Украины [2-6], Донецкого политехнического института [7], ученых НПО «Доникс» [8], специалистов ОАО «БМЗ» [9], ученых Череповецкого государственного университета [10], Екатеринбургского ОАО ВНИИМТ [11], СибГИУ [12] и др.

Вопросы оптимизации технологии ускоренного охлаждения рассматривают применительно к конкретным профилям и установкам [13] или маркам стали [14], по концептуальному подходу работы отсутствуют. В работе [8] описана концепция построения автоматизированной системы управления процессом упрочнения арматурных профилей на базе синтеза математических моделей комбинированного охлаждения, структурообразования и механических свойств. Наиболее детальное описание системы автоматизации управления технологией в работах [11], [15-17]. Однако, вопросы оптимизации режимов охлаждения не нашли отражения. Поэтому постановка задачи оптимизации режимов охлаждения является актуальной.

В проектном сортаменте среднесортно-мелкосортного стана 390 Макеевского металлургического завода [18] присутствует прокат арматурный для железобетонных конструкций классов от А400 до А800 номеров от 10 до 32. За чистовой клетью стана расположен участок ускоренного охлаждения (УУО), предназначенный для термической обработки проката с прокатного нагрева конструкции «Thermex». Результаты моделирования УУО стана 390 с помощью простой ИНС на основе перцептрона процесса ускоренного охлаждения арматурного проката показали, что приемлемое решение получить не удастся [19]. Ошибки прогноза $b_{0.2}$ превышают допустимые значения (2-3%), которые гарантируют получение заданных свойств проката по его длине в пределах, допускаемых стандартом. Поэтому принято решение использовать прямой расчет по физико-математической модели [20] с алгоритмом оптимизации, учитывающим все ограничения.

Постановка задачи. При оптимизации входными параметрами, кроме параметров конструкции камер охлаждения, которые считаются постоянными и в расчете не изменяются, являются: - класс арматуры- АClass; (А400С, А500С, А600С А800С), определяемый величиной условного предела текучести $\sigma_{0.2}$, МПа;

- условный диаметр профиля $-d_0$, мм; (10, 12, 14, 16, 18, 20, 22, 25, 28, 32);
- количество ниток проката N;
- скорость прокатки- V_{np} , м/с (5.0...14.5);
- температура конца прокатки t_{np} , °С (980...1100);
- химсостав основных химических элементов стали (C, Mn, Si), %;
- температура воды на входе в установку $t_{в.вх.о}$, °С (обычно 20-40).

Приняты следующие ограничения, устанавливаемые как настроечные параметры:

- максимально и минимально возможные давления воды на форсунках P_{max} и P_{min} ;
- максимальная среднemasсовая температура воды на выходе из каждой секции $t_{в.ср.i}$;
- минимальная скорость воды по условию гидротранспортирования $V_{в.ср}=K_{mp} \cdot V_{np}$;
- величина допустимого разброса предела текучести σ_m от оптимального уровня $\Delta\sigma$;
- величина оптимального уровня предела текучести для каждого класса прочности σ_{opt} .

В качестве выходных параметров приняты:

- количество $N_{вкл}$, калибр K и длина включенных секций, $L_{охл.i}$;
- давление воды на форсунках включенных секций $P_{\phi i}$, атм;
- механические свойства: пределы текучести $\sigma_{0.2}(\sigma_m)$ и прочности σ_b , Мпа; относительное удлинение δ , %;
- средняя температура воды на выходе каждой секции $t_{в.вых.i}$, °С;
- скорость воды по секциям, м/с; ($V_{в.max}$, $V_{в.ср}$, $V_{в.min}$)
- расход воды по секциям - W_i м³/ч;
- суммарное время охлаждения - $\tau_{охл}$, с;
- общая длина охлаждения $-L_{охл}$, м;
- общий расход воды на охлаждение - W , м³/ч;
- максимальная производительность установки - Q , т/ч;
- температура самоотпуска - t_{co} (за термоустановкой, после окончания выравнивания по сечению) °С;
- количество итераций при поиске оптимума.

Результаты

Разработка алгоритма оптимизации при наличии ряда ограничений выполнена впервые и облегчается при использовании свойств математической модели, установленных прямым расчетом, фиксируя значения одних входных параметров и варьируя другие. Для установления свойств модели проведен расчет охлаждения арматуры №20 из стали СтЗпс со следующим химсоставом С-0.22%, Мн-0.55%, Si-0.07% при прокатке со скоростью $V_{np}=12.5$ м/с, при температуре прокатки $t_{np}=1082$ °С. Расчет произведен при переменной сумме давлений ΣP в 4-х секциях. Данные для моделирования показаны в таблице 1, результаты – на рисунке 1.

Анализ данных в соответствии с рисунком 1 показывает линейный характер влияния ΣP на температуру воды на выходе камер охлаждения $t_{в.вых}$, среднюю скорость воды и почти линейный - на предел текучести $\sigma_{0.2}$. Это означает, что поиск оптимального

решения возможен используя градиенты $dP/db_{0.2}$, $dP/dV_{в.ср}$ и $dP/dt_{в.вых.i}$. В соответствии с рисунком 2 показана блок-схема алгоритма оптимизации технологии ускоренного охлаждения с использованием прямого расчета по модели «Cooling» [21].

Таблица 1 – Данные моделирования

$\sigma_{0.2}$	σ_b	P1	P2	P3	P4	ΣP , атм	тв.вых1	тв.вых2	тв.вых3	тв.вых4
556	646	22	22	22	22	88	66	57	51	47
547	635	20	20	20	20	80	67	58	52	48
538	625	18	18	18	18	72	68	59	53	49
520	605	16	16	16	16	64	69	60	54	50

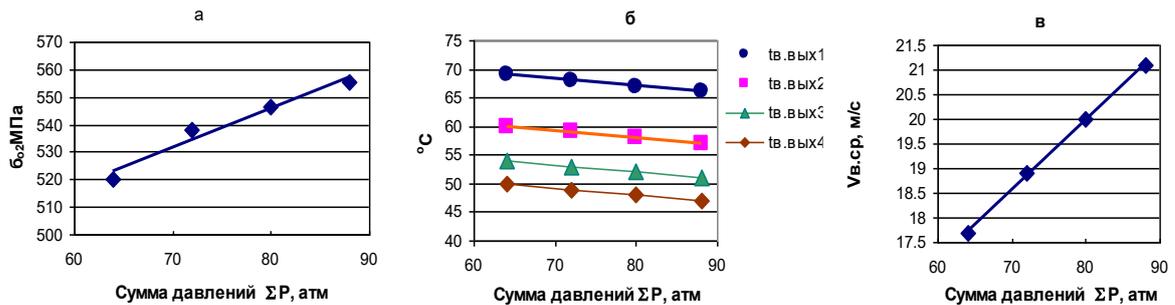


Рисунок 1 –Влияние суммы давлений воды на форсунках 4-х секций установки при охлаждении арматуры №20: а- на предел текучести; б- на температуру воды на выходе камеры охлаждения; в – на среднюю скорость воды на УУО стана 390.

В соответствии с таблицей 2 приведены результаты сравнения расхода воды на охлаждение $\Delta W/W$ в существующих и оптимизированных (опт) режимах охлаждения

Таблица 2 – Результаты сравнения параметров охлаждения существующих и оптимальных режимов (класс А500С)

№ опыта	doxN мм	Vпр м/с	tпр °C	C %	Mn %	Si %	$\sigma_{0.2}$	σ_b	Ω %	P1	P2	P3	P4	P5	W м ³ /ч	$\Delta W/W$ %
							Мпа			атм						
1	25x1	9.0	1090	0.17	0.65	0.3	593	689	18	18	18	18	18	0	562	-
							542	629	20	22	22	22	0	0	471	16%
2	20x1	12.5	1100	0.17	0.6	0.1	585	679	18	22	22	22	22	22	542	
							524	609	20	22	22	22	22	0	434	20%
3	10x4	10.0	1090	0.21	0.6	0.3	575	668	18	19.6	0	0	0	0	118	
							553	643	19	15.4	0	0	0	0	100	15%
4	12x4	9.0	1060	0.24	0.65	0.15	607	705	17	16	0	0	0	0	104	
							543	631	20	11.5	0	0	0	0	83	20%

Последовательность выполнения программы следующая. На первом этапе определяются тип и размеры камер охлаждения (длина и количество элементов, длина участка выравнивания). Тип камеры зависит от диаметра арматуры d_o , всего имеется 3 типа с минимальным диаметром 33, 42 и 48мм в длинном L и коротком S исполнении. Автоматически подключаются ограничения и определяется диапазон оптимальных значений b_{min} и b_{max} . Определение количества включенных секций осуществляется следующим образом: включается одна длинная секция, давление на форсунке принимается равным p_{max} и производится расчет σ_m . Количество включенных секций увеличивается на одну $inc(N_{вкл})$ в случае выполнения неравенства $b_m < b_{min}$.

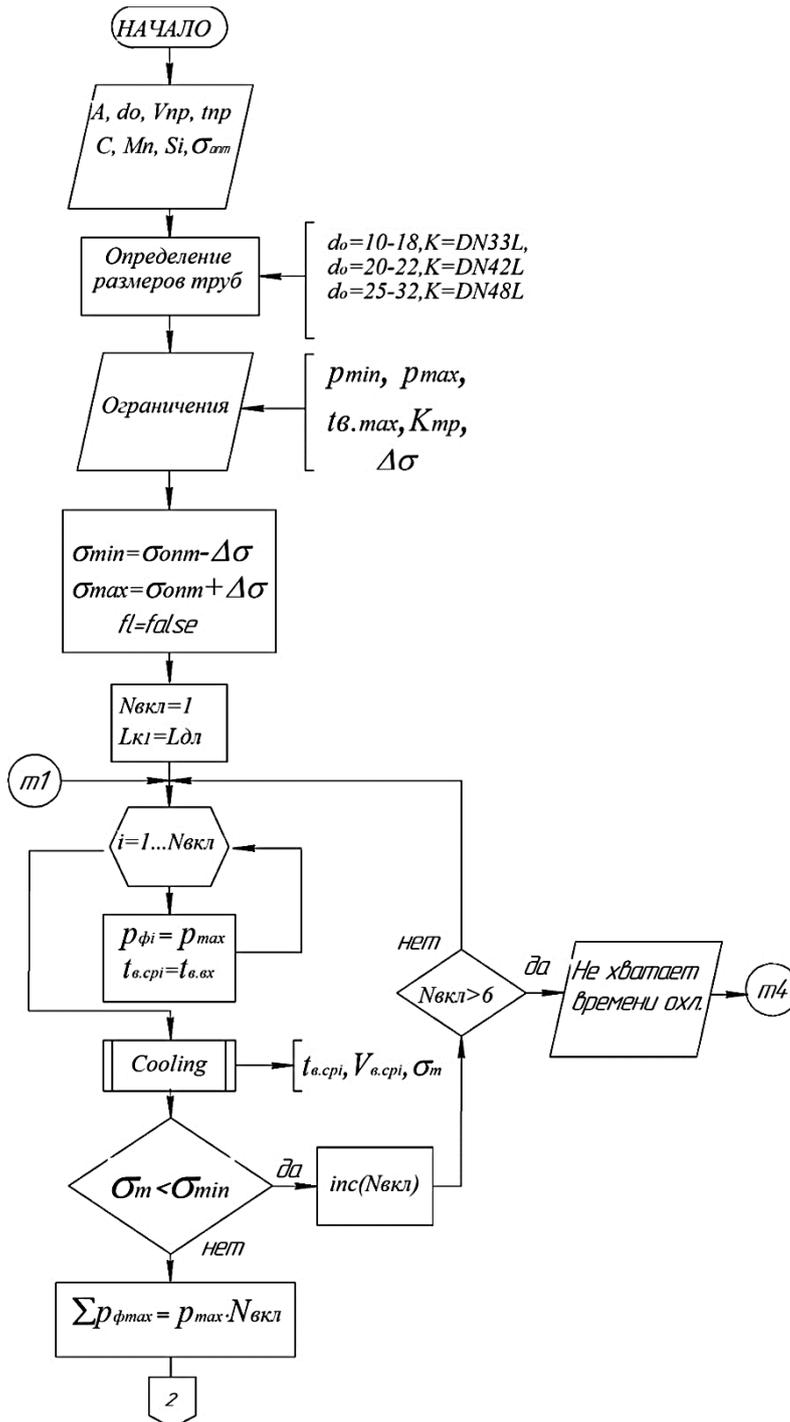


Рисунок 2 - Блок-схема алгоритма оптимизации режима термоупрочнения арматурных профилей (Лист 1)

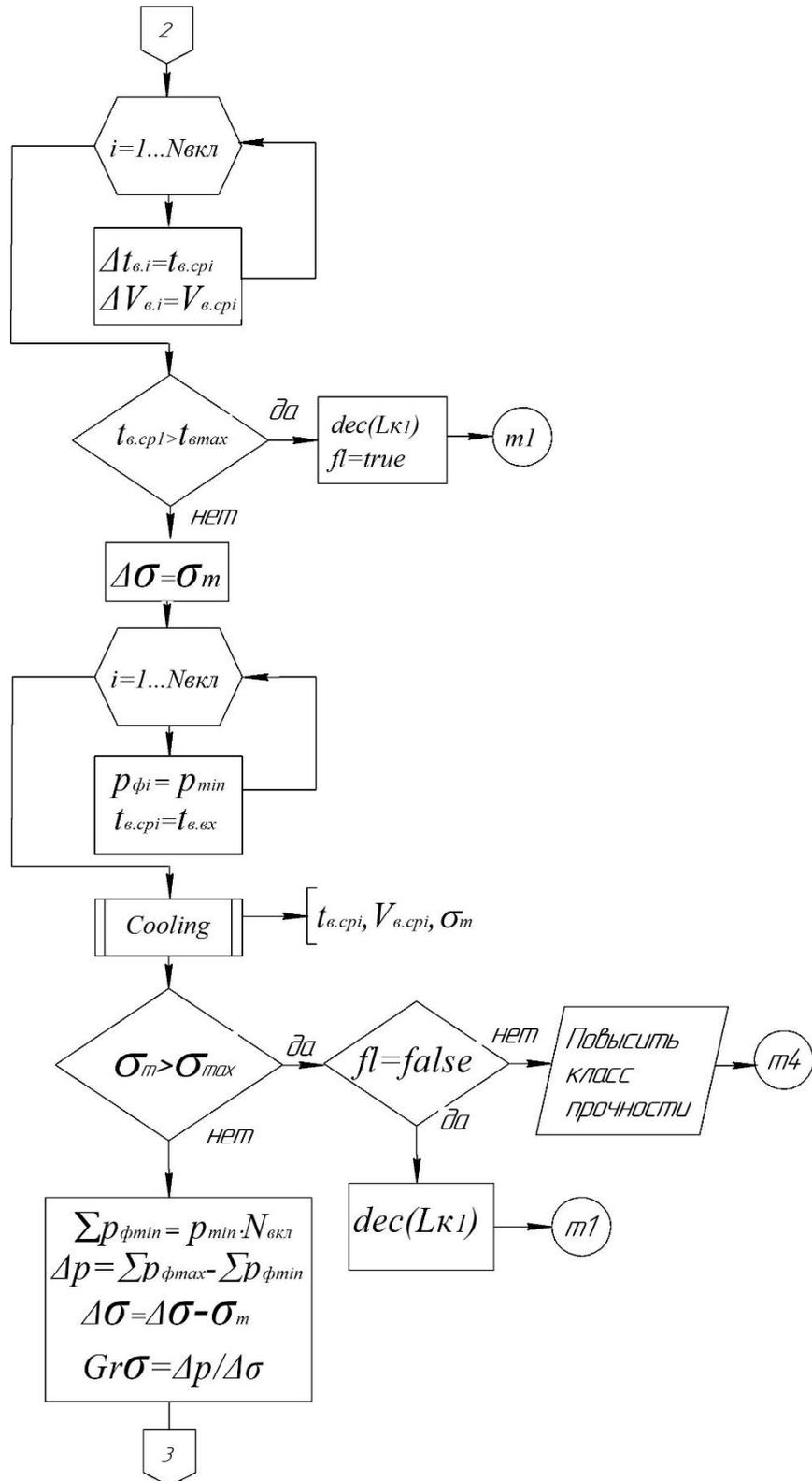


Рисунок 2 - продолжение (Лист 2)

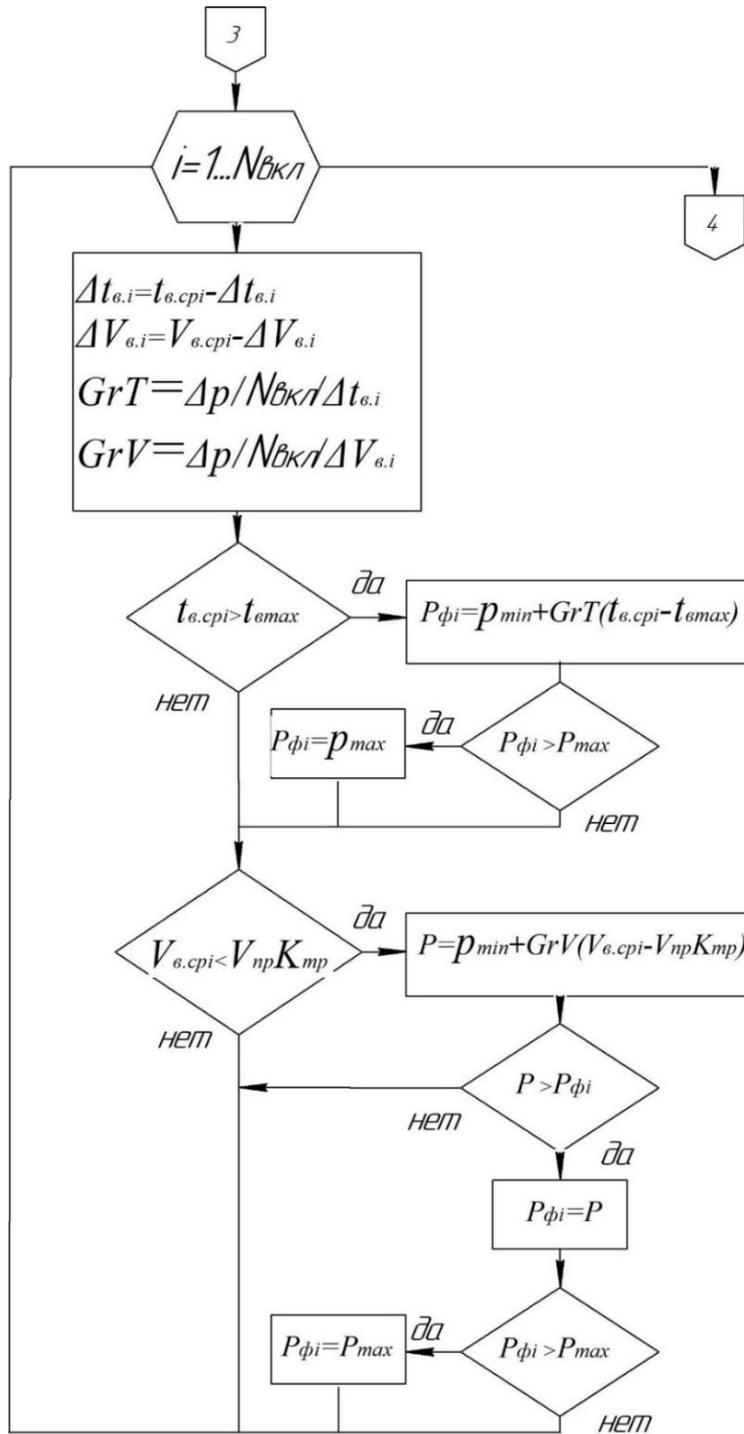


Рисунок 2 – продолжение (Лист 3)

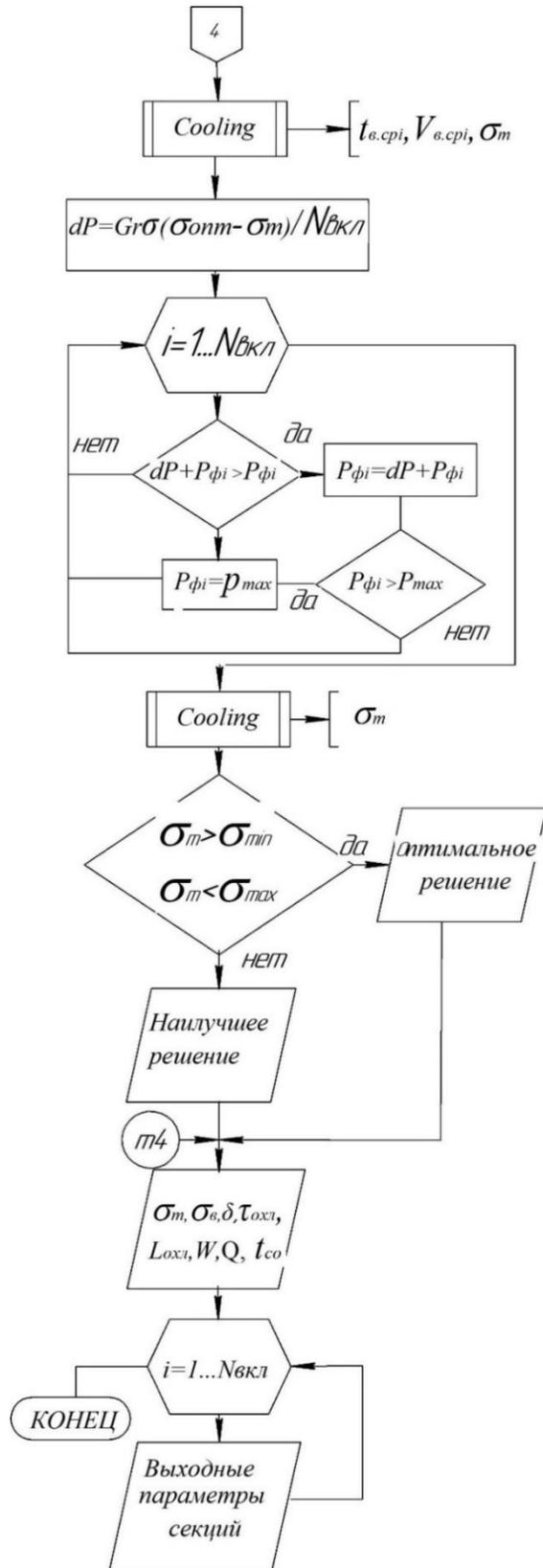


Рисунок 2 - окончание (Лист 4)

Если температура воды на выходе 1-й секции превышает допускаемую $t_{в,max}$, выбирается короткая камера охлаждения $L_{kl} dec(L_{kl})$. Производится пересчет режима при $p_{фор.i} = p_{max}$.

Далее осуществляется расчет режима при $p_{фор.i} = p_{min}$ при известном количестве и длине включенных секций.

По известным значениям σ_m , $t_{в,ср.i}$, $V_{в,ср.i}$ при крайних значениях давлений на форсунках определяются градиенты $Gr\sigma$, GrT , GrV , с помощью которых уточняется режим охлаждения на УУО.

По результатам расчета на экран выводятся сообщения или рекомендации оператору по изменению исходных данных для улучшения результата.

Заключение

На основе физико-математической модели ускоренного охлаждения арматурных профилей из низкоуглеродистых и низколегированных сталей в потоке прокатного стана впервые разработаны алгоритм и программа оптимизации режима упрочнения. Программа обеспечивает расчет оптимальных по расходу воды настроек для получения прочностных свойств на уровне близком к нижней границе, определяемой стандартами, и повышение пластических свойств арматуры

Практическая полезность разработки заключается в повышении качества арматурного проката путем стабилизации механических свойств по длине проката при охлаждении профилей из различных марок стали и возможных изменениях режимов прокатки, в упрощении решения задачи расширения сортамента арматуры по классу прочности и стандартам и сокращении сроков и потерь металла при освоении новой продукции, в оптимизации расхода охлаждающей воды и электроэнергии на насосах порядка на 10–20%.

Аппаратно разработка может быть реализована в нескольких вариантах: как советчик оператора участка термоупрочнения, подсистема в АСУ ТП, цифровой двойник прокатного стана.

Список литературы

1. [Электронный ресурс]: URL: <https://metal-archive.ru/metallurgicheskie-processy/1076-sposoby-i-ustroystva-dlya-ohlazhdeniya-sortovogo-prokata.html>. (дата обращения: 08.07.2025).
2. Губинский В.И., Минаев А.Н., Гончаров Ю.В., Уменьшение окисления при производстве проката - К.:Техніка, 1981.-135 с.
3. Губинский В.И., Рыбалов А.А. Моделирование ускоренного охлаждения арматурного проката в потоке непрерывных мелкосортных станов. «Промышленная теплотехника», 2002, т.24, №6.-С.24-28.
4. Худик В.Т. Формирование температуры самоотпуска в некоторых арматурных сталях/ В.Т.Худик и др. // Металлургическая и горнорудная промышленность. 2004. № 3. С.74 - 77.
5. Пилюшенко В.Л., Ноговицин А.В. Феноменологическая модель механических свойств проката из низкоуглеродистой микролегированной стали / Фундаментальные и прикладные проблемы черной металлургии. Киев: Наукова думка, 1995. 218с.
6. Левченко Л. Н. Производство арматурной стали / Л. Н. Левченко и др.// М.: Металлургия, 1984. 136 С.
7. Курбатов, Ю.Л. Разработка и исследование двухмерной математической модели ускоренного охлаждения проката / Ю.Л. Курбатов, и др. Донец, политехн. ин-т. Донецк, 1986 9 с. Деп. в Черметинформации 10.12.86, № 3724.
8. Солод В.С., Пашинский В.В. Синтез технологии термоупрочнения и автоматизированной системы управления установкой термоупрочнения арматурных профилей с помощью математической модели структурообразования / [Черная металлургия. Бюллетень научно-технической и экономической информации](#). 2009. № 11 (1319). С. 61-65.

9. Кириленко, О. М. Разработка способа эффективного термоупрочнения проката арматурного профиля диаметром 8 мм из стали обыкновенного качества в условиях стана 320 СПЦ-1 ОАО «БМЗ – управляющая компания холдинга «БМК» / О. М. Кириленко, П. А. Бобков, М. В. Петренко, А. В. Русаленко // *Литье и металлургия*. 2020. № 1. С. 21–24. <https://doi.org/10.21122/1683-6065-2020-1-21-24>.
10. Киркин, Д.С. Разработка математической модели процесса термического упрочнения горячекатаной арматурной стали / Д.С. Киркин, Ю.А. Калягин // *Молодые исследователи - регионам: Материалы Всероссийской науч. конф. студ. и аспирантов*. - Вологда, 21-22 апреля 2005 г. - Вологда: ВГТУ, 2005.-Т. 1.-С79-81.
11. Эйсмонтд К.Ю. Автоматизированная система управления устройством термоупрочнения арматуры стана 350 / К.Ю. Эйсмонтд и др. // *Сталь*. 2007. № 6. С. 40 – 42.
12. Математическое моделирование на ЭВМ процессов превращения аустенита низколегированных сталей при прерывном охлаждении арматуры в условиях стана 450 ОАО ЗСМК / В.Д. Сарычев, А.Б. Юрьев, П.М. Новосельцев, В.Е. Громов // *Краевые задачи и математическое моделирование* : сб. тр. 4-й всерос. науч. конф., 1-4 дек. 2001 г. / под ред. О.В. Каледина ; НФИ КемГУ. – Новокузнецк, 2001. – Т. 1. – С. 67-72.
13. А.Б.Юрьев Оптимизация технологии упрочнения стержневой арматуры диаметром 32-40 мм класса А500С / А.Б. Юрьев и др. // *Сталь*. - 2002. - № 2.-С. 68-69.
14. Савьюк, А.Н. Оптимизация химического состава стали, технологии прокатки и термообработки для обеспечения механических свойств проката 162 В узком диапазоне значений / А.Н. Савьюк, и др. // *Тр. 5- го конгресса прокатчиков*. - М., 2004. - С. 200-205.
15. Липунов Ю.И. Разработка и ввод в эксплуатацию участка термоупрочнения арматуры в потоке литейно-прокатного комплекса стана 280 / Ю.И.Липунов и др. // *Сталь* 2010. № 3. С.86-90
16. Липунов Ю.И., Траянов Г.Г., Эйсмонтд К.Ю. Разработка систем регулируемого охлаждения и технологий термоупрочнения. *Сталь* 2010. № 3. С.96-99
17. Липунов Ю.И., Эйсмонтд К.Ю., Некрасова Е.В. Автоматизированные системы управления термоупрочнением проката в потоке стана. *Сталь* 2020. № 5. С.57-60. ISSN 0038—920X
18. Цкитишвили Э.О. Освоение производства высококачественной продукции на самом новом в Украине среднесортно-мелкосортном стане 390/ Э.О.Цкитишвили и др. // *ОАО «Черметинформация». Бюллетень «Черная металлургия»*.–2011.–№7.–С.56–65.
19. Солод В.С. Оценка возможностей использования нейросетевых моделей для оптимизации технологии охлаждения арматурной стали / В.С.Солод, В.М. Зуев, С.Б.Иванова // *Проблемы искусственного интеллекта 2025 № 3 (38)*. С. 65-77.
20. Солод В.С. Физико-математическая модель для разработки экспертной системы ускоренного охлаждения арматурного проката / *Проблемы искусственного интеллекта 2025 № 2 (37)*. С. 44-58. DOI 10.34757/2413-7383.2025-2-37-44-58.

References

1. [Electronic resource]: URL: <https://metal-archive.ru/metallurgicheskie-processy/1076-sposoby-i-ustroystva-dlya-ohlazhdeniya-sortovogo-prokata.html>. (date accessed: 07/08/2025).
2. Gubinsky V.I., Minaev A.N., Goncharov Yu.V., Reducing scale formation in the production of rolled products - K.: Tekhnika, 1981.-135 p.
3. Gubinsky V.I., Rybalov A.A. Modeling of accelerated cooling of reinforcing bars in the flow of continuous small-section mills. "Industrial Heat Engineering", 2002, v.24, no. 6.-P.24-28.
4. Khudik V.T. Formation of self-tempering temperature in some reinforcing steels / V.T. Khudik at al. // *Metallurgical and mining industry*. 2004. No. 3. P. 74-77.
5. Pilyushenko V.L., Nogovitsyn A.V. Phenomenological model of mechanical properties of rolled products made of low-carbon microalloyed steel / *Fundamental and applied problems of ferrous metallurgy*.- Kyiv: Naukova Dumka, 1995. - 218 p.
6. Levchenko L.N. Production of reinforcing steel / L.N. Levchenko at al. // М.: Metallurgy, 1984. 136 p.
7. Kurbatov, Yu.L. Development and study of a two-dimensional mathematical model of accelerated cooling of rolled products / Yu.L. Kurbatov, at al. // . - Donetsk, Polytechnical Institute. - Donetsk, 1986.- 9 p. - Dep. in Chermetinformatiya 12/10/86, No. 3724.
8. Solod V.S., Pashinsky V.V. Synthesis of heat hardening technology and automated control system for a reinforcement section heat hardening unit using a mathematical model of structure formation/ *Ferrous metallurgy. Bulletin of scientific, technical and economic information*. 2009. No. 11 (1319). P. 61-65.
9. Kirilenko, O. M. Development of a method for efficient heat hardening of rolled reinforcement sections with a diameter of 8 mm made of ordinary quality steel under the conditions of the 320 SPC-1 mill of

- OJSC BMZ - BMK Holding Management Company / O. M. Kirilenko et al. // Casting and metallurgy. 2020. No. 1. P. 21–24. <https://doi.org/10.21122/1683-6065-2020-1-21-24>.
10. Kirkin, D.S. Development of a Mathematical Model of the Process of Thermal Hardening of Hot-Rolled Reinforcing Steel / D.S. Kirkin, Yu.A. Kalyagin // Young Researchers - to the Regions: Proceedings of the All-Russian Scientific Conf. of Students and Postgraduates. - Vologda, April 21-22, 2005. - Vologda: VSTU, 2005.-Vol. 1.-P 79-81.
 11. Eysmond K. Automated Control System for the Rebar Yu Thermal Hardening Device of Mill 350 / K.Yu. Eysmond et al. // Steel. 2007. No. 6. Pp. 40 – 42.
 12. Sarychev V.D. Mathematical Modeling on a Computer of the Processes of Austenite Transformation in Low-Alloy Steels with Intermittent Cooling of Rebar in the Conditions of Mill 450 of OAO ZSMK / V.D. Sarychev et al. // Boundary value problems and mathematical modeling: Coll. t. 4th All-Russian scientific conf., December 1-4, 2001 / edited by O.V. Kaledin; NFI KemsU. - Novokuznetsk, 2001. - Vol. 1. - Pp. 67-72.
 13. Yuryev A.B. Optimization of the technology for strengthening rod reinforcement with a diameter of 32-40 mm of class A500C / A.B. Yuryev et al. // Steel. - 2002. - No. 2.-Pp. 68-69.
 14. Savyuk, A.N. Optimization of the chemical composition of steel, rolling technology and heat treatment to ensure the mechanical properties of rolled products 162 In a narrow range of values / A.N. Savyuk, et al. // Proc. 5th congress of rolling mill workers. - M., 2004. - P. 200-205.
 15. Lipunov Yu.I. Development and commissioning of a section for thermal strengthening of reinforcement in the flow of the casting and rolling complex of mill 280 / Yu.I. Lipunov et al.// Steel 2010. No. 3. pp. 86-90
 16. Lipunov Yu. I., Trayanov G.G., Eysmond K.Yu. Development of controlled cooling systems and thermal strengthening technologies. Steel 2010. No. 3. pp. 96-99
 17. Lipunov Yu.I., Eysmond K.Yu, Nekrasova E.V. Automated control systems for thermal strengthening of rolled products in the flow of the mill. Steel 2020. No. 5. P. 57-60. ISSN 0038—920X
 18. Tskitishvili E.O. Development of high-quality production at the newest medium-small-section mill 390 in Ukraine / E.O. Tskitishvili et al. // JSC Chermetinformatiya. Bulletin "Ferrous Metallurgy". - 2011. - No. 7. - P. 56-65.
 19. Solod V.S., Assessment of the possibilities of using neural network models to optimize the cooling technology of reinforcing steel / V.S. Solod, V.M. Zuyev, S.B.Ivanova// Problems of Artificial Intelligence 2025 No. 3 (38). - P. 65-77 .
 20. Solod V.S. Physical and mathematical model for the development of an expert system for accelerated cooling of reinforcing bars / Problems of Artificial Intelligence 2025 No. 2 (37). P. 44-58. DOI 10.34757/2413-7383.2025-2-37-44-58.

RESUME

V. S. Solod

Algorithm and program for optimizing the accelerated cooling operation mode of reinforcement bars in the thermal unit of the "Thermex" design

Introduction: The increasing of requirements for the quality of reinforcement bars, can be satisfied by using an integrated approach to solving the problems of determining the influence of the chemical composition of steel and the main technological parameters of the production process on the final quality indicators.

The following methods were used: mathematical modeling of the process, algorithmizing and programming method.

Results: Based on the physical and mathematical model of accelerated cooling of reinforcement bars made of low-carbon and low-alloy steels in the rolling mill flow, an algorithm and program for optimizing the hardening mode were developed for the first time. The program provides for the calculation of the optimal settings for the heat treatment line in terms of water consumption to obtain strength properties at a level close to the lower limit determined by the standards, and an increase in the plastic properties of the reinforcement bars.

Conclusion: In terms of hardware, the development can be implemented in several versions: advisor to the operator of the heat treatment device, a subsystem built into the APCS, a digital twin of the rolling mill.

РЕЗЮМЕ

В. С. Солод

Алгоритм и программа оптимизации режима ускоренного охлаждения арматурных профилей в термоустановке конструкции «Thermex»

Введение: Ужесточение требований к качеству арматурных профилей возможно удовлетворить путём применения комплексного подхода к решению задач определения влияния химического состава стали и основных технологических параметров производственного процесса на конечные показатели качества.

Использованные методы: математическое моделирование процесса, метод алгоритмизации и программирования.

Результаты: на основе математической модели ускоренного охлаждения арматурных профилей в потоке прокатного стана впервые разработаны алгоритм и программа оптимизации режима упрочнения. Программа обеспечивает расчет оптимальных по расходу воды настроек термоустановки для получения прочностных свойств на уровне близком к нижней границе, определяемой стандартами, и повышение пластических свойств арматуры.

Заключение: Аппаратно разработка может быть реализована в нескольких вариантах, таких как: советчик оператора участка термоупрочнения, подсистема, встроенная в АСУТП; цифровой двойник прокатного стана.

Солод Владимир Сергеевич – к.т.н., научный сотрудник, отдел интеллектуальных робототехнических систем, Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта».

Область научных интересов: компьютерное моделирование технологии прокатного производства, эл. почта vs.solod@mail.ru, адрес: 283048, г. Донецк, ул. Артема, д. 118 б, Россия, телефон +7(949) 459- 02 -42,

Статья поступила в редакцию 24.09.2025.

Раздел 3
Section 3

Автоматизация и управление
технологическими процессами
и производствами

*Automation and Control
of Technological Processes
and Production*

УДК 658.011.56

DOI 10.24412/2413-7383-2025-4-39-125-136

Н. В. Водолазская¹, М.К. Вотинцев¹, Е. В. Копытова¹, А.И. Хиценко²¹ МИРЭА-Российский технологический университет,
119454, г. Москва, проспект Вернадского д.78² Государственный университет просвещения
105005, г. Москва, ул. Радио, 10А, стр. 2

ИНФОРМАЦИОННО-УПРАВЛЯЮЩАЯ СИСТЕМА КОНТРОЛЯ ТЕХНОЛОГИЧЕСКОГО ПРОЦЕССА ИЗГОТОВЛЕНИЯ КОМПЛЕКСНЫХ МИНЕРАЛЬНЫХ УДОБРЕНИЙ

N. V. Vodolazskaya¹, M. K. Votintsev¹, E. V. Kopytova¹, A. I. Khitsenko²¹ MIREA-Russian Technological University,
119454 Moscow, Vernadsky Avenue, 78² State University of Education,
105005 Moscow, Radio str., 10A, Bldg 2

INFORMATION AND CONTROL SYSTEM FOR MONITORING THE TECHNOLOGICAL PROCESS OF MANUFACTURING COMPLEX MINERAL FERTILIZERS

Предлагается информационно-управляющая система для контроля процесса производства комплексных минеральных удобрений с созданием специального SCADA-интерфейса. Анализируются требования к нему, включая необходимость отображения динамических параметров процесса, состояния технологического оборудования и аварийных сигналов. Этот интерфейс обеспечивает отображение текущего состояния оборудования, дозирование компонентов, гранулирование, сушку продукции в реальном времени, и предоставляет оператору инструменты для управления технологическим процессом.

Ключевые слова: SCADA, автоматизация производства, комплексные минеральные удобрения, информационно-управляющая система, технологический контроль.

An information management system is proposed for monitoring the production process of complex mineral fertilizers with the creation of a special SCADA interface. The requirements for it are analyzed, including the need to display dynamic process parameters, the status of technological equipment and emergency signals. This interface provides a display of the current equipment status, component dosing, granulation, product drying in real time, and offers the operator tools for managing the technological process.

Key words: SCADA, production automation, complex mineral fertilizers, information management system, technological control.

Введение

В условиях стремительной цифровизации и глобализации экономики применение современных информационных технологий представляет собой важный аспект развития предприятия [1-5]. Эти технологии включают в себя как программное обеспечение для проектирования, так и разработку соответствующих инструментов для моделирования производственных процессов, что позволяет предприятиям существенно повысить эффективность своей работы, снизить затраты и улучшить качество продукции [6-9]. В связи с этим, актуальным направлением является исследование, разработка и применение инновационных решений, в частности информационно-управляющих систем анализа технологических процессов, например, изготовления комплексных минеральных удобрений, производство которых основано на переработке природного сырья [10-15]. Они компенсируют дефицит микро- и макроэлементов в почве, что особенно важно в условиях интенсивного земледелия [16-19].

Целью работы является повышение эффективности технологического процесса изготовления комплексных минеральных удобрений.

Материалы и методы

Для реализации задач мониторинга и управления технологическим процессом изготовления комплексных минеральных удобрений предложена информационно-управляющая система (ИУС), построенная на принципах системности, отказоустойчивости и адаптивности к изменяющимся производственным условиям. Ключевым элементом такой системы является **SCADA-система**, выполняющая функции верхнего уровня управления. Выбор обусловлен не только стандартными функциями сбора и визуализации данных, но и возможностью реализации сложных алгоритмов адаптивного управления, основанных на математических моделях процессов гранулирования и сушки. Для обеспечения надежной и безопасной связи между уровнем управления (SCADA) и уровнем контроллеров (PLC) и датчиков применяется современный промышленный стандарт – **протокол OPC UA**. Он позволяет интегрировать в единое информационное пространство оборудование различных вендоров, что критически важно для модернизируемых производств, где парк аппаратуры неоднороден, а встроенные механизмы шифрования, аутентификации и контроля целостности данных делают его пригодным для использования в ответственных производствах, каким является химико-технологический комплекс.

Предложенное авторское решение представляет собой новую целостную систему, в которой аппаратная часть, коммуникационные протоколы и программное обеспечение подобраны и спроектированы исходя из единой цели – повышения управляемости, предсказуемости и экономической эффективности технологического процесса производства комплексных минеральных удобрений

Основная часть

Производство азотных удобрений начинается с синтеза аммиака (NH_3) из азота воздуха и водорода под высоким давлением и температурой в присутствии катализатора (процесс Габера-Боша) [20-25]. Показана схема производства NP/NPS/NPK - удобрений по типу аммонизатор-гранулятор (АГ) - сушильный барабан (СБ) в соответствии с рисунком 1. Есть три типа производства комплексных минеральных удобрений: распылительно-кипящая сушиллка-гранулятор (аппарат РКСТГ), барабан гранулятор-сушиллка, аммонизатор-гранулятор (АГ) - сушильный барабан (СБ) [18-20].

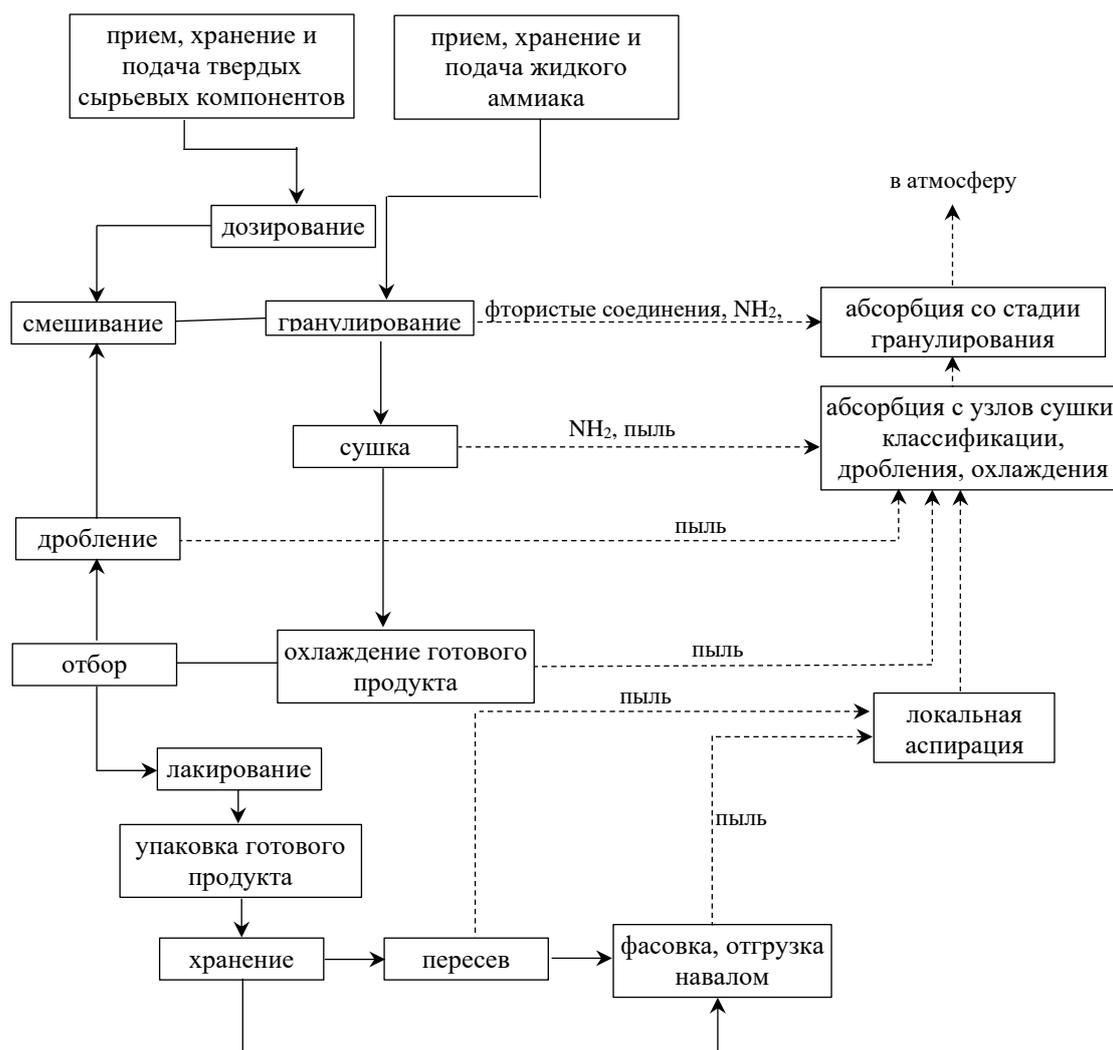


Рисунок 1 – Схема производства NP/NPS/NPK-удобрений по типу АГ-СБ

При использовании аппарата типа РКСТ возникает ряд сложностей из-за особенностей типа производства, т.е., он подходит не для всех типов удобрений, так как некоторые составы могут слипаться или разрушаться в условиях кипящего слоя и по сравнению с барабанными или тарельчатыми грануляторами аппараты РКСТ уступают в объеме выпускаемой продукции. Производство гранулированных удобрений в барабанном грануляторе-сушилке также требует значительного количества электроэнергии, при этом возможен разброс размеров гранул из-за неравномерного распределения материала в барабане, что приводит к необходимости дополнительной сортировки и, в целом, может привести к ухудшению качества продукта.

Использование системы аммонизатор-гранулятор (АГ) - сушильный барабан (СБ) позволяет эффективно вводить азот в состав удобрений за счет равномерного распределения жидкого аммиака в гранулируемой массе, обеспечивает стабильный размер и форму частиц, что улучшает сыпучесть и удобство внесения удобрений и объединяет стадии аммонизации, гранулирования и частичной сушки в одном аппарате, сокращая технологическую цепочку. По сравнению с барабанными и другими типами грануляторов это оборудование требует меньших энергозатрат, а получаемые

гранулы обладают хорошей механической устойчивостью, что снижает их разрушение при транспортировке и хранении. Использование аммонизаторов-грануляторов позволяет оптимизировать технологический процесс и получать продукцию, соответствующую требованиям современного сельского хозяйства. Недостатками данного способа являются затруднения в производстве уравновешенных марок удобрений, вызванное неполным растворением хлорид калия (ХК) и необходимость проведения отдельной нейтрализации азотной кислоты (АК) аммиаком с последующей упаркой пульпы до плава, что не только усложняет технологическую схему процесса, но и повышает ее пожаро- и взрывоопасность. Одним из способов, устранения этих недостатков является моделирование этапов технологического процесса изготовления комплексных минеральных удобрений и разработка информационно-управляющей системы контроля.

Современные производства минеральных удобрений представляют собой сложные технологические комплексы с множеством взаимосвязанных процессов (дозирование, смешивание, гранулирование, сушка и т.д.). Ручное управление такими процессами неэффективно и повышает риск аварийных ситуаций. SCADA обеспечивает централизованный контроль всех технологических параметров в реальном времени [26-29]. Чтобы разработать интерфейс для SACDA системы, надо определить параметры технологического процесса. Представлены значения технологических параметров процесса гранулирования в соответствии с таблицей 1.

Таблица 1 – Значения технологических параметров процесса гранулирования

№	Место нахождения	Измеряемые параметры	Единицы измерения	Номинальные значения	Контроль	Регулирование
1	Барабан	влажность	%	1,5	+	-
2	Барабан	температура	°С	89,85-99,85	+	+
3	Бак	давление	Бар	5	+	+
4	Барабан	частота вращения	об/с	8-12	+	+
5	Бак	жидкость	%	0	+	-

Контроль каждого параметра этапа гранулирования выполняется определенными устройствами:

- контроль влажности гранул в резервуаре на выходе — датчик влажности;
- контроль температуры в барабане — датчик температуры, термопара; регулирование температуры в барабане — подача ретур;
- контроль давления — датчик давления; регулирование давления аммиака — насос;
- контроль частоты вращения барабана — датчик скорости, тахометр;
- регулирование частоты вращения барабана — частотный преобразователь;
- контроль утечки аммиака — датчик аммиака.

Приведены значения технологических параметров процесса сушки в соответствии с таблицей 2. На этапе сушки используются такие устройства для контроля параметров: контроль частоты вращения цилиндра — датчик скорости, тахометр; регулирование частоты вращения цилиндра — частотный преобразователь; контроль влажности материала на выходе — датчик влажности; контроль подачи теплоносителя — датчик давления.

Таблица 2 – Значения технологических параметров процесса сушки

№	Место нахождения	Измеряемые параметры	Единицы измерения	Номинальные значения	Контроль	Регулирование
1	Барабан	частота вращения	об/с	1,8	+	+
2	Барабан	влажность	%	1	+	-
3	Топка	температура	°С	144,85 - 174,85	+	-
4	Барабан	температура	°С	99,85 - 109,85	+	+
5	Топка	Газ	%	0	+	-
6	Топка	давление	Бар	3	+	+
7	Барабан	температура	°С	94,85 - 99,85	+	-

А также задействованы: контроль температуры теплоносителя в сушилке — термopара; регулирование температуры в сушилке — подача кислоты; контроль температуры теплоносителя до выхода — термopара; контроль утечки газа — датчик газа; контроль температуры гранул после сушки — датчик температуры; контроль влажность пара в барабане — датчик влажности.

Обработка и анализ результатов

С учетом проведенного анализа процесса изготовления комплексных минеральных удобрений и технологических параметров был разработан интерфейс на SCADA системе для этапов гранулирования и сушки для информационно – управляющей системы. Были предложены уникальные схемы мониторинга ключевых параметров процесса гранулирования, включая такие показатели как температура массы, влажность гранул, частота вращения барабана. Аналогичный интерфейс контроля с соответствующими технологическими параметрами реализован для других критически важных узлов производства: для бака с аммиаком (уровень заполнения, давление, температура хранения), сушильного барабана (температура сушки, скорость вращения, влажность отходящих газов), топки (температура горения, расход топлива, состав выхлопных газов), насосного оборудования (давление на выходе, производительность, вибрационные характеристики) и шлюзовых затворов (частота срабатывания, герметичность, состояние уплотнений). Для примера представлены конвейер, который подает гранулы на этап сушки после этапа гранулирования, сушильный барабан с параметрами, кнопка для открытия/закрытия шлюза и конвейер для отправки готового продукта на следующий технологический этап в соответствии с рисунком 2. Подобные новые разработки выполнены для аммонизатора – гранулятора и бака с ретуром, бака с аммиаком и бака, предназначенного для хранения кислоты, с подключенным к нему насосом. Рядом с каждым баком расположены кнопка управления, позволяющие вручную включать и выключать насос при необходимости. Причем автоматическая работа насоса активируется, когда температура в процессе сушки удобрений выходит за установленные границы допустимых значений. Такая система обеспечивает контроль за технологическим процессом, предотвращая перегрев или недостаточную сушку удобрений.

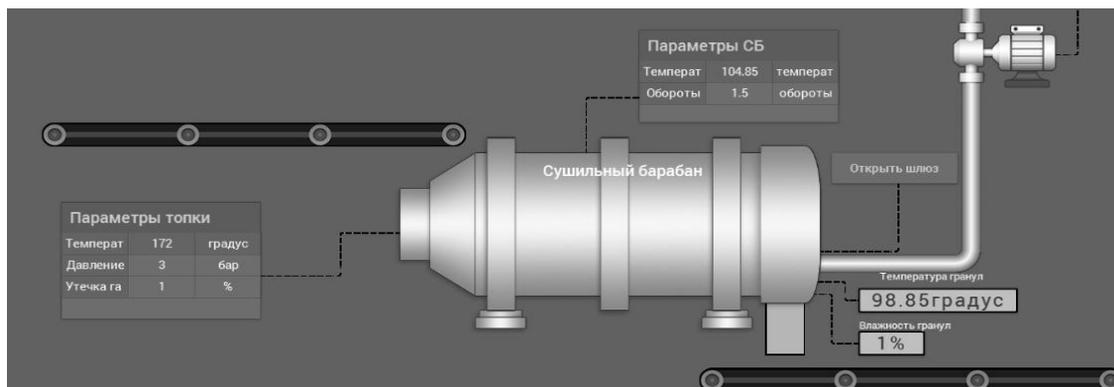


Рисунок 2 – Сушильный барабан с параметрами

В соответствии с рисунком 3 для настройки системных переменных разработан скрипт для изменения параметров.

```

1 begin
2     Datchik_Chastoty_Vrashenia_AG.Value := 7 + Random(6);
3     Datchik_Temperaturey_AG.Value := 88 + Random(13);
4     Datchik_Vlajnosty_AG.Value := 1.5 + Random(2);
5     Datchik_Davleniy.Value := 4 + Random(2);
6     Datchik_Ammiaka.Value := Random(2);
7     Datchik_Davleniy_Topka.Value := 2 + Random(2);
8     Datchik_Gaza.Value := Random(2);
9     Datchik_Temperaturey_Topka.Value := 143 + Random(33);
10    Datchik_Chastoty_Vrasheniy_SB.Value := 1.5 + Random(2);
11    Datchik_Temperaturey_SB.Value := 99.85 + Random(15);
12    Datchik_Vlajnosty_SB.Value := 1 + Random(2);
13    Datchik_Temperaturey_Granul.Value := 94.85 + Random(7);
14 end.

```

Рисунок 3 – Скрипт с параметрами

Скрипт написан на языке Python и включает в себя несколько ключевых функций, а также реализована проверка вводимых данных, чтобы избежать ошибок при выполнении. После запуска скрипта происходит автоматическое обновление конфигурации системы в соответствии с заданными значениями. Следовательно, представленный код позволяет гибко настраивать параметры, экономя время пользователя. Создание соответствующих трендов представлено блоком интерфейса в соответствии с рисунком 4. Обозначаются параметры, которые будут меняться и изображаться на графиках для отслеживания на этапе гранулирования. Выбраны параметры частоты вращения АГ и температуры гранулирования. Также для бака АГ указаны параметры давления, топки СБ параметры давления и температуры, процесса сушки параметры частоты вращения и температуры.

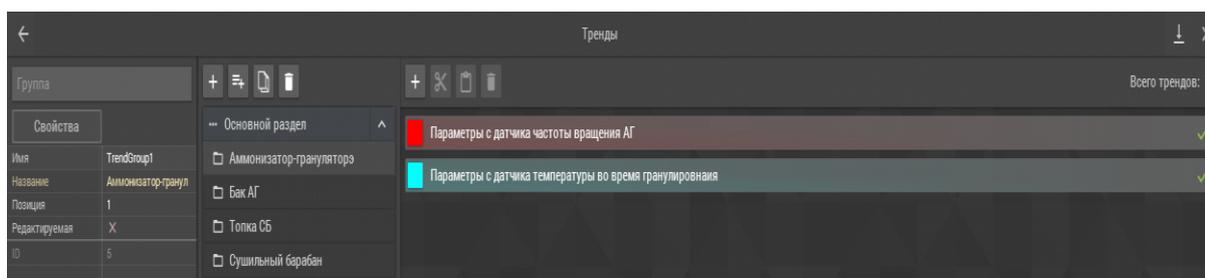
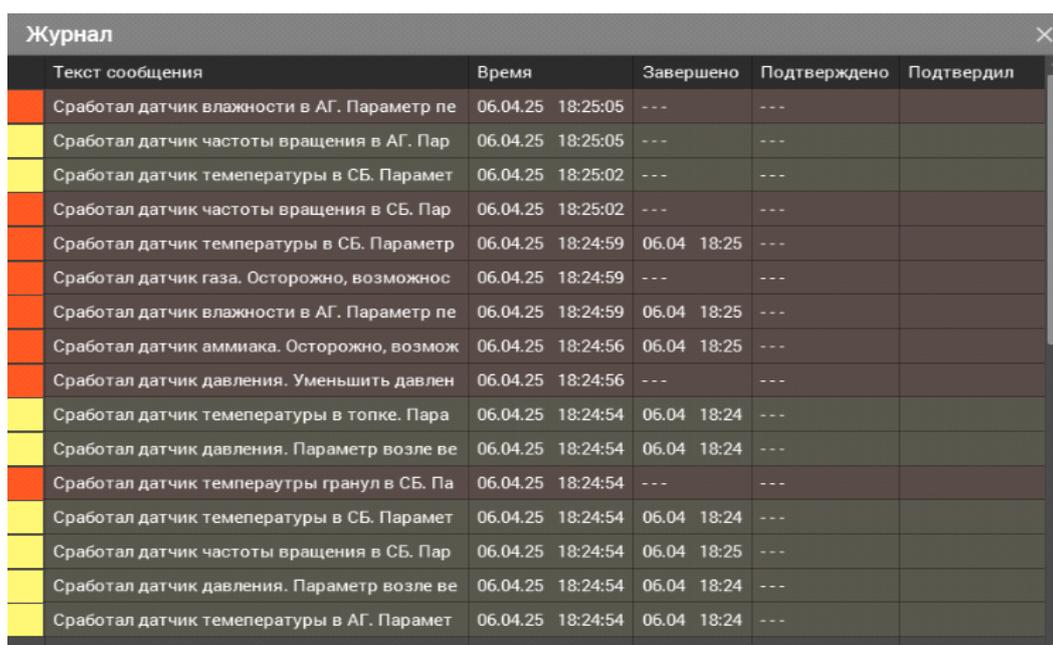


Рисунок 4 – Параметры для процесса гранулирования

Сформирован редактор сообщений о процессе производства комплексных минеральных удобрений для аммонизатора-гранулятора, в этом окне можно редактировать сообщения для бака с ретуром, процесса сушки и топки. На рисунке 5 представлен системный журнал, в который автоматически записываются все сообщения, формируемые в ходе процессов гранулирования и сушки. Данный журнал служит для фиксации ключевых параметров работы оборудования, включая температуру, давление и время выполнения операций. Все события, такие как запуск или остановка аппаратуры, ошибки и предупреждения, также отображаются в этом журнале для последующего анализа. Записи в журнале помогают операторам контролировать ход технологического процесса и своевременно реагировать на отклонения от нормы.



Текст сообщения	Время	Завершено	Подтверждено	Подтвердил
Сработал датчик влажности в АГ. Параметр пе	06.04.25 18:25:05	---	---	
Сработал датчик частоты вращения в АГ. Пар	06.04.25 18:25:05	---	---	
Сработал датчик температуры в СБ. Парамет	06.04.25 18:25:02	---	---	
Сработал датчик частоты вращения в СБ. Пар	06.04.25 18:25:02	---	---	
Сработал датчик температуры в СБ. Параметр	06.04.25 18:24:59	06.04 18:25	---	
Сработал датчик газа. Осторожно, возможнос	06.04.25 18:24:59	---	---	
Сработал датчик влажности в АГ. Параметр пе	06.04.25 18:24:59	06.04 18:25	---	
Сработал датчик аммиака. Осторожно, возмож	06.04.25 18:24:56	06.04 18:25	---	
Сработал датчик давления. Уменьшить давлен	06.04.25 18:24:56	---	---	
Сработал датчик температуры в толке. Пара	06.04.25 18:24:54	06.04 18:24	---	
Сработал датчик давления. Параметр возле ве	06.04.25 18:24:54	06.04 18:24	---	
Сработал датчик температура гранул в СБ. Па	06.04.25 18:24:54	---	---	
Сработал датчик температуры в СБ. Парамет	06.04.25 18:24:54	06.04 18:24	---	
Сработал датчик частоты вращения в СБ. Пар	06.04.25 18:24:54	06.04 18:25	---	
Сработал датчик давления. Параметр возле ве	06.04.25 18:24:54	06.04 18:24	---	
Сработал датчик температуры в АГ. Парамет	06.04.25 18:24:54	06.04 18:24	---	

Рисунок 5 – Журнал сообщений

Далее строится тренд параметров в аммонизаторе-грануляторе, отражающий изменение ключевых показателей за определенный период. Анализ тренда позволяет сделать вывод о стабильности работы аммонизатора-гранулятора в заданном режиме. Сравнение с предыдущими данными показывает, что текущие значения находятся в пределах нормы. Таким образом, разработанный тренд помогает операторам контролировать процесс и своевременно вносить корректировки.

Выводы

Выполненная новая унификация интерфейсов мониторинга позволяет оператору быстро ориентироваться в системе управления, обеспечивая оперативный контроль за всеми технологическими процессами. Каждый элемент системы оснащен датчиками, передающими данные в централизованную SCADA-систему в режиме реального времени. Это обеспечивает возможность оперативного реагирования на любые отклонения от заданных технологических параметров, поддерживая стабильность производственного процесса. Единый принцип визуализации данных значительно сокращает время на обучение персонала и снижает вероятность ошибок при управлении оборудованием.

Кроме того, подобная стандартизация интерфейсов упрощает интеграцию новых модулей в случае модернизации производства, а также позволяет проводить сравнительный анализ работы различных узлов технологической линии. Все параметры архивируются в единой базе данных, что обеспечивает возможность последующего детального анализа эффективности работы оборудования и оптимизации производственных процессов.

Следовательно, применение SCADA-системы в производстве комплексных минеральных удобрений является технологически и экономически обоснованным решением, обладающим новизной. Система позволяет существенно повысить эффективность производства, обеспечить стабильное качество продукции, улучшить показатели безопасности и создать основу для цифровой трансформации предприятия. Внедрение SCADA соответствует современным тенденциям Industry 4.0 и обеспечивает конкурентные преимущества на рынке.

Список литературы

1. Анцыферов С. С. Интеллектуальные системы управления технологическими процессами / С. С. Анцыферов, К. Н. Фазилова, К. Е. Русанов // Проблемы искусственного интеллекта. 2024. №2(33). С. 37-44. DOI 10.24412/2413-7383-2024-2-37-44. EDN: WLSZIT
2. Vodolazskaya N. Process Modeling of Increasing Reliability of Assembly Responsible Connections// Advanced Engineering Letters. 2024. №3(3). P. 91-99. DOI:10.46793/adeletters.2024.3.3.1. EDN: DBBCYW.
3. Нейросетевая оценка эффективности использования фульвокислот в сочетании с минеральными удобрениями на продуктивность и качество листового салата сорта Афицион, выращенного в условиях вертикальной гидропоники/ Н.И. Воробьев, Я.В. Пухальский, С.И. Лоскутов [и др.] // Аграрная наука. 2025. №(8). С. 104-114. <https://doi.org/10.32634/0869-8155-2025-397-08-104-114>. EDN: HUQMUH.
4. Кривоудубский О. А. Методология, определяющая правила разработки моделей и алгоритмов систем управления в условиях информационной и технологической трансформации / О. А. Кривоудубский // Проблемы искусственного интеллекта. 2020. № 1 (16). С. 47-50. EDN: FHGGXJ .
5. Свидетельство о государственной регистрации программы для ЭВМ № 2025618197 РФ. Разработка программного обеспечения для автоматизированного расчёта теплообменных процессов в биореакторе: заявл. 19.03.2025; опубл. 02.04.2025 / Н. В. Водолазская, Д. Н. Клесов, А. А. Ореховская, В. Э. Ващилин; заявитель ФГБОУ ВО «МИРЭА – Российский технологический университет». EDN: VDCGZB.
6. Enhancing food security in central asia and the caucasus: a SWOT analysis of agro-industrial potential/ D. Zhilyakov, O. Petrushina, K. Meshcheryakov, [et al.] International Journal of Safety and Security Engineering. 2025. Т. 15. № 7. 1461-1470 p. DOI 10.18280/ijss.150713. EDN: NOGODG.
7. Фам К.Б. Нечеткая модель диагностики технологических процессов / Фам К.Б., Мурашев П.М., Богатилов В.Н. // Проблемы искусственного интеллекта. 2023. №3 (30). С. 76-86. DOI 10.34757/2413-7383.2023.30.3.007. EDN: ANQJVJ.
8. Математическое моделирование и алгоритмизация функций мониторинга технологических процессов на основе многоточечных измерительных систем / И.В. Ковалев, Д.И. Ковалев, В.В. Лосев, [и др.] // Современные наукоемкие технологии. 2021 № 6-1. С. 29-38. DOI:10.17513/snt.38693. EDN: ERYLYC.
9. Свидетельство о государственной регистрации программы для ЭВМ № 2023685439 РФ. Калькулятор общей численности почвенных микроорганизмов: № 2023684091; заявл. 14.11.2023; опубл. 27.11.2023 / В.Э.Ващилин, А.А.Ореховская, Д.Н.Клесов, Н.В.Водолазская; заявитель ФГБОУ ВО «Белгородский ГАУ». EDN: MLGJEC.
10. Влияние степени нейтрализации фосфорной кислоты на свойства нитратсодержащих азотно-фосфорно-калийных удобрений / В. М. Колпаков, А. М. Норов, Д. А. Пагалешкин [и др.] // Известия высших учебных заведений. Серия: Химия и химическая технология. 2021. Т. 64, № 3. С. 52-58. DOI 10.6060/ivkkt.20216403.6289. EDN JBMFQX.
11. The Importance of State Support and Regulation in the Agro-Industrial Complex / D. I. Zhilyakov, D. I. Ryakhovskiy, N. G. Bondarenko [et al.] // Res Militaris. 2022. Vol. 12, No. 2. P. 2549-2560. EDN BGCRRS.
12. Водолазская Н.В. Модели повышения надежности некоторых технических систем // Совершенствование инженерно-технического обеспечения производственных процессов и технологических систем: Материалы нац. научно-практ. конф. с международным участием. Оренбург: ООО "Агентство "Пресса". 2022. С.304-308. EDN: BQWADD.
13. Antsyferov S.S. Intelligent Quality Control Systems for High-Tech Manufacturing Processes and Equipment // S.S. Antsyferov, K.N. Fazilova, D.S. Muravyov // Problems of Artificial Intelligence. 2025. №2(37). P. 60-65. DOI: 10.24412/2413-7383-2025-2-37-60-65. EDN: CWZEJN.

14. Перинская Е. В. Математическое моделирование процессов функционирования специализированных аппаратов конвективного типа / Е. В. Перинская // Проблемы искусственного интеллекта. - Донецк: ГУ ИПИИ. 2022. № 2(25). С. 55-63. EDN: CDLKER
15. Математическое моделирование процесса измельчения материалов / Э.М. Кольцова, М.А. Бабкин, Н.А. Попова, А.В. Женса // Теоретические основы химической технологии. 2024. Т 58. № 1. С. 115-121. DOI:10.31857/S0040357124010141. EDN: YVTKPL.
16. Pittman W. Lessons to be learned from an analysis of ammonium nitrate disasters in the last 100 years / W. Pittman // Journal of Hazardous Materials. 2014. Vol. 280. P. 472-477. DOI:10.1016/j.jhazmat.2014.08.037
17. Жилияков Д.И. Сравнительная оценка экономической эффективности использования двухбарабанного разбрасывателя гранулированных удобрений/ Д.И. Жилияков, В.Н. Трубников, И.В. Коротков // Вестник Курской государственной сельскохозяйственной академии. 2022. № 9. С. 244-250. EDN: JXWVFJ.
18. Интенсивная технология NPK-удобрений на основе мытого сушёного концентрата центральных Кызылкумов // Р.М. Назирова, С.М. Таджиев, С.Р. Мирсалимова, Д. Худаярова // Проблемы Науки. 2019. №2, С. 6-11. DOI 10.20861/2304-2338-2019-135-001. EDN: YXSHJR.
19. Бабкин В.В. Фосфорные удобрения России / В.В. Бабкин, А.А. Бродский М.: ТОО «Агрохим-принт». 1995. 464 с.
20. Исследование кинетики термического разложения NPK-удобрений на основе нитрата аммония / К. Г. Горбовский, А. И. Казаков, А. М. Норов [и др.] // Труды Кольского научного центра РАН. – 2019. Т. 10, № 1-3. С. 64-71. DOI 10.25702/KSC.2307-5252.2019.10.1.64-71. EDN NRFNAY.
21. Vyazovkin S. ICTAC Kinetics Committee recommendations for collecting experimental thermal analysis data for kinetic computations / S. Vyazovkin // Thermochimica Acta. 2014. Vol. 590. P. 1-23. DOI 10.1016/j.tca.2014.05.036. EDN UWKEXP.
22. The reliability analysis of combined harvesters in the usual conditions of operation / R. V. Pavlyuk, A. T. Lebedev, Y. I. Zhevorga, E. V. Zubenko // IOP Conference Series: Earth and Environmental Science, Stavropol, 21–22 октября 2019 года. Vol. 488. Stavropol, 2020. P. 012033. DOI 10.1088/1755-1315/488/1/012033. EDN: WEPOXL.
23. Izato, Y. I. Thermal decomposition mechanism of ammonium nitrate and potassium chloride mixtures / Y. I. Izato, A. Miyake // Journal of Thermal Analysis and Calorimetry. 2015. Vol. 121, No. 1. P. 287-294. DOI 10.1007/s10973-015-4739-1. EDN: USVFUP.
24. Сорокин К.Н. Математическая модель для разработки технологической линии по производству гуминовых удобрений на основе модульного оборудования / К.Н. Сорокин // Агроинженерия. 2022. Т 24. № 3. С. 19-26. DOI: 10.26897/2687-1 149-2022-3-19-26. EDN: XSFYTX.
25. Кошелева М.К. Математическое моделирование тепло- и массопереноса при сушке гранул носителя для никелевого катализатора / М.К. Кошелева, В.П. Мешалкин, О.Р. Дорняк // Теоретические основы химической технологии. 2021. Т 55. № 3. С. 339-346. DOI: 10.31857/S0040357121030258. EDN: VBFPMU.
26. Мозаидзе Е.С. Тематическое моделирование в потоке коротких сообщений на русском языке/ Е.С. Мозаидзе // Russian Technological Journal. 2025. №13(1) С. 38-48. <https://doi.org/10.32362/2500-316X-2025-13-1-38-48>. EDN: HJHQTR.
27. Водолазская Н. В. Информационная система контроля доступа и отслеживания этапов производственного процесса / Н. В. Водолазская, Е. В. Копытова, К. В. Лаврентьева, В. Д. Черненко // Проблемы искусственного интеллекта. 2025. Т. 37. №.2. С. 116-128. doi:10.24412/2413-7383-2025-2-37-116-128.
28. Information monitoring system design for industrial enterprises / Т. О. Bochkareva, I. E. Kirillov, V. N. Bogatkov [et al.] // Software Journal: Theory and Applications. 2020. No. 2. P. 20-28. DOI 10.15827/2311-6749.20.2.3. EDN QZRRMP.
29. Мосияш А.Э. Моделирование процессов разработки и внедрения информационных систем в организации / А.Э. Мосияш, А.В. Каширский // Петербургский экономический журнал. 2022. № 1-2. С. 167-177. DOI: 10.24412/2307-5368-2022-1-2-167-177. EDN: GNRBCB.

References

1. Antsyferov S. S., Fazilova K. N., Rusanov K. E. Intelligent control systems for technological processes // Problems of Artificial Intelligence, №2(33). – 2024. – P. 37-44. DOI 10.24412 / 2413-7383-2024-2-37-44. EDN: WLSZIT.
2. Vodolazskaya N. Process Modeling of Increasing Reliability of Assembly Responsible Connections // Advanced Engineering Letters, №3(3). – 2024. – P. 91-99. DOI: 10.46793/adeletters.2024.3.3.1. EDN: DBBCYW.
3. Vorobyov N.I., Pukhalsky Y.V., Loskutov S.I., [et al.]. Neural network assessment of the effectiveness of using fulvic acids in combination with mineral fertilizers on the productivity and quality of Athenian lettuce grown under vertical hydroponic conditions //Agrarian Science, №(8). – 2025. – P. 104-114. <https://doi.org/10.32634/0869-8155-2025-397-08-104-114>. EDN: HUQMUH.
4. Krivodubsky, O. A. Methodology that Defines the Rules for Developing Models and Algorithms for Control Systems in the Context of Information and Technological Transformation // Problems of Artificial Intelligence, №1(16). – 2020. – P. 47-50. EDN: FHGGXJ.
5. Certificate of state registration of computer program № 2025618197 of the Russian Federation.

- Development of software for automated calculation of heat exchange processes in a bioreactor / N.V. Vodolazskaya, D.N. Klesov, A.A. Orekhovskaya, V.E. Vashchilin: declared. 19.03.2025: published. 02.04.2025 declared. RTU-MIREA. EDN: VDCGZB.
6. Zhilyakov D., Petrushina O., Meshcheryakov K., [et al.] Enhancing food security in central asia and the caucasus: a swot analysis of agro-industrial potential // *International Journal of Safety and Security Engineering*, №7, T. 15. – 2025. – 1461-1470 p.
 7. Fam K.B., Murashev P.M., Bogatikov V.N. Fuzzy model for diagnostics of technological processes // *Problems of Artificial Intelligence*, №3(30). – 2023. – P. 76-86. EDN: AHQJVJ.
 8. Kovalev I.V., Kovalev D.I., Losev V.V., [et al.] Mathematical modeling and algorithmization of monitoring functions of technological processes based on multi-point measuring systems // *Modern science-intensive technologies*, №6-1. – 2021 – P. 29-38. DOI: 10.17513/snt.38693. EDN: ERYLYC
 9. Certificate of state registration of the computer program № 2023685439 of the Russian Federation. Calculator of the total number of soil microorganisms: № 2023684091: declared. 14.11.2023: published. 27.11.2023 / V.E. Vashchilin, A.A. Orekhovskaya, D.N. Klesov, N.V. Vodolazskaya; applicant Belgorod State Agrarian University. EDN: MLGJEC
 10. Kolpakov V.M., Norov A.M., Pagaleshkin D.A., [et al.] Influence of the degree of neutralization of phosphoric acid on the properties of nitrate-containing nitrogen-phosphorus-potassium fertilizers – *News of higher educational institutions. Series "Chemistry and chemical technology"*, №3, T. 64. – 2021. – P. 52-58. – DOI 10.6060/ivkkt.20216403.6289. EDN JBMFQX.
 11. Zhilyakov D.I., Ryakhovsky D.I., Bondarenko N.G., [et al.] The importance of state support and regulation in the agro-industrial complex // *Res Militaris*, №2, T. 12 – 2022. – P. 2549-2560. EDN BGCRRS
 12. Vodolazskaya N.V. Models for improving the reliability of some technical systems // *Improving the engineering and technical support of production processes and technological systems: Proceedings of the national scientific and practical conference with international participation.* – Orenburg: House of Press Agency LLC. – 2022. – P. 304-308. EDN BQWADD.
 13. Antsyferov S.S., Fazilova K.N., Muravyov D.S. Intelligent Quality Control Systems for High-Tech Manufacturing Processes and Equipment // *Problems of Artificial Intelligence*, №2(37). – 2025. – P. 60-65. DOI: 10.24412/2413-7383-2025-2-37-60-65. EDN: CWZEJN.
 14. Perinskaya, E. V. Mathematical Modeling of the Processes of Functioning of Specialized Convective-Type Devices // *Problems of Artificial Intelligence*, № 2(25). – 2022. – P. 55-63. EDN: CDLKER
 15. Koltsova E.M., Babkin M.A., Popova N.A., Zhensa A.V. Mathematical modeling of the process of grinding materials // *Theoretical Foundations of Chemical Technology*, №1, T 58. – 2024. – P. 115-121. DOI: 10.31857/S0040357124010141. EDN: YVTKPL.
 16. Pittman W. Lessons to be learned from an analysis of ammonium nitrate disasters in the last 100 years – *Journal of Hazardous Materials*, Vol. 280. – 2014. – P. 472-477. DOI: 10.1016/j.jhazmat.2014.08.037
 17. Zhilyakov D.I., Trubnikov V.N., Korotkov I.V. Comparative assessment of the economic efficiency of using a two-drum spreader of granular fertilizers // *Bulletin of the Kursk State Agricultural Academy*, №9. 2022. P. 244-250. EDN: JXWVFJ.
 18. Nazirova R.M., Tadzhiyev S.M., Mirsalimova S.R., Khudayarova D. Intensive technology of NPK fertilizers based on washed dried concentrate of the central Kyzylkum // *Problems of Science*, №2. – 2019. – P. 6-11. – DOI 10.20861/2304-2338-2019-135-001. EDN: YXSHJR.
 19. Babkin V.V., Brodsky A.A. Phosphorus fertilizers in Russia – M.: Agrokhim-print LLP. – 1995. – 464 p.
 20. Gorbovsky K. G., Kazakov A. I., Norov A. I. [et al.] Investigation of the kinetics of thermal decomposition of NPK fertilizers based on ammonium nitrate // *Proceedings of the Kola Scientific Center of the Russian Academy of Sciences*. №1-3, T. 10. – 2019. – P. 64-71. DOI 10.25702/KSC.2307-5252.2019.10.1.64-71. EDN NRFNAY.
 21. Vyazovkin S. ICTAC Kinetics Committee recommendations for collecting experimental thermal analysis data for kinetic computations // *Thermochimica Acta*, Vol. 590. – 2014. – P. 1-23. DOI 10.1016/j.tca.2014.05.036. EDN UWKEXP.
 22. Pavlyuk R. V., Lebedev A. T., Zhevora Y. I., Zubenko E. V. The reliability analysis of combined harvesters in the usual conditions of operation // *IOP Conference Series: Earth and Environmental Science*, Vol. 488. – 2020. – P. 012033. DOI 10.1088/1755-1315/488/1/012033. EDN: WEPOXL.
 23. Izato, Y. I. Miyake A. Thermal decomposition mechanism of ammonium nitrate and potassium chloride mixtures // *Journal of Thermal Analysis and Calorimetry*, №1. Vol. 121. – 2015. – P. 287-294. DOI 10.1007/s10973-015-4739-1. EDN: USVFUP.
 24. Sorokin, K.N. Mathematical model for developing a technological line for the production of humic fertilizers based on modular equipment // *Agroengineering*, №3, T. 24. – 2022. – P. 19-26. DOI: 10.26897/2687-1149-2022-3-19-26. EDN: XSFYTX.
 25. Kosheleva M.K., Meshalkin V.P., Dorniyak O.R. Mathematical Modeling of Heat and Mass Transfer during Drying of Carrier Granules for Nickel Catalyst // *Theoretical Foundations of Chemical Technology*, №3, T 55. – 2021. – P. 339-346. DOI: 10.31857/S0040357121030258. EDN: VBFPMU.
 26. Mozaidze E.S. Thematic Modeling in the Flow of Short Messages in Russian // *Russian Technological Journal*, №13(1). – 2025. – P. 38-48. <https://doi.org/10.32362/2500-316X-2025-13-1-38-48>. EDN: HJHQTR

27. Vodolazskaya N. V., Kopytova E. V., Lavrentyeva K. V., Chernenko V. D. Information System for Access Control and Tracking of Production Process Stages // Problems of Artificial Intelligence, №2, Т. 37 – 2025. P. 116-128. doi:10.24412/2413-7383-2025-2-37-116-128.
28. Information monitoring system design for industrial enterprises / T. O. Bochkareva, I. E. Kirillov, V. N. Bogatikov [et al.] // Software Journal: Theory and Applications. – 2020. – No. 2. – P. 20-28. – DOI 10.15827/2311-6749.20.2.3. – EDN QZRRMP.
29. Mosiyash A.E., Kashirskiy A.V. Modeling the Processes of Developing and Implementing Information Systems in an Organization // Petersburg Economic Journal, №1-2. – 2022. – P. 167-177. DOI: 10.24412/2307-5368-2022-1-2-167-177. EDN: GNRBCB.

RESUME

*N. V. Vodolazskaya, M.K. Votintsev, E. V. Kopytova, A.I.Khitsenko
Information And Control System For Monitoring The Technological Process
Of Manufacturing Complex Mineral Fertilizers*

This work is aimed at improving the technology for the production of complex mineral fertilizers. Fertilizers such as nitrogen, phosphorus and potash are vital for modern farming, as they provide plants with nutrients and prevent soil depletion. Improving the efficiency of their manufacture will optimize this complex but critical process.

An adaptive information management system was proposed to monitor and control the fertilizer production process. The key element is the SCADA system, which implements complex control algorithms and integrates with the equipment through the OPC UA standard. This end-to-end solution improves the manageability, predictability and cost-effectiveness of the entire process.

The technological process of manufacturing complex mineral fertilizers includes dosing, mixing of raw materials and subsequent granulation in an ammoniator-granulator, where, due to ammonization and optimal moistening, a fertilizer balanced in composition with high agronomic value is formed. The necessary process parameters for developing the interface for the SCADA system are highlighted.

Based on the analysis of the technological process and key parameters of the production of complex mineral fertilizers, a specialized interface was developed in the SCADA system. This interface covers the critical stages of granulation and drying, ensuring their monitoring and control. Thus, a central element of the information and control system has been created, which allows the operator to control the production cycle in real time.

A single visualization and monitoring standard speeds up the work of the operator and simplifies the integration of new equipment. The real-time data collection and archiving implemented in the SCADA system allows both operational control of the process and deep analysis to optimize it. Thus, the implementation of the system is technologically and economically justified, increasing the efficiency, safety and competitiveness of production.

РЕЗЮМЕ

*Н. В. Водолазская, М.К. Вотинцев, Е. В. Копытова, А.И. Хиценко
Информационно-управляющая система контроля технологического
процесса изготовления комплексных минеральных удобрений*

Данная работа направлена на совершенствование технологии производства комплексных минеральных удобрений. Удобрения, такие как азотные, фосфорные и калийные, жизненно необходимы для современного земледелия, так как обеспечивают растения питательными элементами и предотвращают истощение почв. Повышение эффективности их изготовления позволит оптимизировать этот сложный, но критически важный процесс.

Для мониторинга и управления процессом производства удобрений была предложена адаптивная информационно-управляющая система. Ключевой элемент — SCADA-

система, которая реализует сложные алгоритмы управления и интегрируется с оборудованием через стандарт OPC UA. Данное комплексное решение повышает управляемость, предсказуемость и экономическую эффективность всего технологического процесса.

Рассмотрен технологический процесс изготовления комплексных минеральных удобрений, который включает дозировку, перемешивание сырья и последующее гранулирование в аммонизаторе-грануляторе, где за счет аммонизации и оптимального увлажнения формируется сбалансированное по составу удобрение с высокой агрономической ценностью. Выделены необходимые параметры технологического процесса для разработки интерфейса для SCADA системы.

На основе проведенного анализа технологического процесса и ключевых параметров изготовления комплексных минеральных удобрений был разработан специализированный интерфейс в SCADA-системе. Данный интерфейс охватывает критические этапы гранулирования и сушки, обеспечивая их мониторинг и управление. Таким образом, создан центральный элемент информационно-управляющей системы, позволяющий оператору контролировать производственный цикл в реальном времени.

Единый стандарт визуализации и мониторинга ускоряет работу оператора и упрощает интеграцию нового оборудования. Реализованный в SCADA-системе сбор данных в реальном времени, и их архивация позволяют как оперативно управлять процессом, так и проводить глубокий анализ для его оптимизации. Таким образом, внедрение системы является технологически и экономически обоснованным, повышая эффективность, безопасность и конкурентоспособность производства.

Водолазская Н. В. – к.т.н., доцент, МИРЭА — Российский технологический университет, институт искусственного интеллекта, кафедра промышленной информатики, 119454, Россия, ЦФО, г. Москва, Проспект Вернадского, д. 78, +7-909-202-51-27, vpv26@bk.ru. *Область научных интересов:* цифровые технологии для автоматизации производственных процессов, моделирование технических систем, обеспечение надежности и безопасности технических систем.

Вотинцев М. К. – инженер, МИРЭА — Российский технологический университет, институт искусственного интеллекта, кафедра промышленной информатики, 119454, Россия, ЦФО, г. Москва, Проспект Вернадского, д. 78, votintsev.mikhail.k@mail.ru. *Область научных интересов:* интеграция информационной системы с MES- и SCADA-платформами.

Копытова Е. В. – ст. преподаватель, МИРЭА — Российский технологический университет, институт искусственного интеллекта, кафедра промышленной информатики, 119454, Россия, ЦФО, г. Москва, Проспект Вернадского, д. 78, +7 916 473 5973, kopytova@mirea.ru. *Область научных интересов:* комплексные системы управления производством.

Хиценко А. И. – к.т.н., доцент, Государственный университет просвещения, кафедра профессионального и технологического образования, 105005, Россия, г. Москва, [ул. Радио, 10А, стр. 2](#), +7 926 780-58-00, ganna_x@mail.ru. *Область научных интересов:* вопросы анализа конструкторских параметров деталей различного оборудования.

Статья поступила в редакцию 22.09.2025.

Д.М. Логунова, А.М. Володина, Е.В. Курнасов, С.К. Кириллов, Н.Д. Тонконоженко
МИРЭА – Российский технологический университет,
119454, Проспект Вернадского, д. 78, Москва, Россия

ПРИМЕНЕНИЕ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ НА ПРОИЗВОДСТВЕ ГИТАР ДЛЯ ВЫЯВЛЕНИЯ ДЕФЕКТОВ ПРОДУКЦИИ И АВТОМАТИЗИРОВАННОГО ПОДБОРА ПАР ЗАГОТОВОК

D.M. Logunova, A.M. Volodina, E.V. Kurnasov, S.K. Kirillov, N.D. Tonkonozhenko
MIREA – Russian Technological University,
119454, Vernadsky ave., bld. 78, Moscow, Russia

THE USE OF CONVOLUTIONAL NEURAL NETWORKS IN THE MANUFACTURE OF GUITARS TO IDENTIFY PRODUCT DEFECTS AND AUTOMATED SELECTION OF PAIRS OF BLANKS

В статье представлена информационно-аналитическая система, обеспечивающая автоматизированный контроль качества деревянных заготовок на производстве гитар с помощью технологий машинного зрения и нейронных сетей. Рассмотрены ключевые компоненты системы, ориентированные на обнаружение брака продукции и автоматизированный подбор пар заготовок с совпадающим рисунком древесины. Приведены результаты работы модуля анализа изображений, использующего для выявления дефектов продукции нейронные сети с архитектурами YOLO и EfficientNet. Показано, что предложенный способ контроля качества продукции позволит в значительной мере исключить субъективный человеческий фактор при выявлении дефектов и подборе заготовок гитар. Представлены алгоритмы и фрагменты программной реализации системы, которые могут быть интегрированы с действующими на производстве SCADA- и MRP-системами.

Ключевые слова: нейронные сети, информационно-аналитические системы, автоматизация производства, деревообрабатывающая промышленность, машинное зрение, контроль качества продукции.

The article presents an information and analytical system that provides automated quality control of wooden blanks in the manufacture of guitars using machine vision technologies and neural networks. The key components of the system focused on the detection of defective products and the automated selection of pairs of blanks with a matching wood pattern are considered. The results of the image analysis module, which uses neural networks with YOLO and EfficientNet architectures to detect product defects, are presented. It is shown that the proposed method of product quality control will significantly eliminate the subjective human factor in identifying defects and selecting guitar blanks. Algorithms and fragments of the software implementation of the system are presented, which can be integrated with SCADA and MRP systems operating in production.

Keywords: neural networks, information and analytical systems, production automation, woodworking industry, machine vision, product quality control.

Введение

На сегодняшний день степень автоматизации на производствах музыкальных инструментов, таких как гитары, может сильно различаться. На предприятиях, где изготавливают инструменты премиального качества, специалисты выполняют почти все технологические операции вручную, в то время как фабрики массового и крупносерийного производства используют транспортные конвейерные линии и современные деревообрабатывающие станки с числовым программным управлением. В основном средства автоматизации, внедряемые на производства, направлены на сокращение ручного труда на этапах механической обработки древесины и транспортировки заготовок между единицами технологического оборудования. Контроль качества исходных материалов на данный момент остаётся на уровне визуальной оценки или измерения необходимых параметров при помощи специализированных инструментов оператором вручную.

Несмотря на то, что автоматизированный контроль качества уже успешно применяется на других деревообрабатывающих производствах, на фабриках по производству гитар автоматизация на данном этапе отсутствует. Такой подход повышает риск накопления дефектов, что приводит к изготовлению инструментов из некачественных материалов и несёт в себе финансовые и репутационные риски. Цель работы: для решения данной проблемы предлагается автоматизировать процесс контроля качества деревянных заготовок на производстве гитар с помощью технологий машинного зрения и нейронных сетей. С учётом всё большего числа областей успешного применения нейронных сетей [1-9], такой подход позволит осуществлять своевременное выявление дефектов деревянных заготовок на начальных этапах производства и снизит вероятность появления брака на последующих.

1. Постановка задачи автоматизации процесса контроля качества изготовления корпусов гитар на различных стадиях производства

Корпус гитары состоит из двух основных частей – дека (верхняя и нижняя) и обечайки. Дека представляет собой плоскую деревянную пластину, формирующую верхнюю или нижнюю часть корпуса и влияющую на акустические характеристики инструмента. Обечайками называют изогнутые боковые стенки корпуса, соединяющие верхнюю и нижнюю деки и задающие общую форму инструмента [10].

Каждый этап производства гитар, начиная с выбора материалов и заканчивая финальной отделкой, оказывает непосредственное влияние на конечный результат – качество звучания и внешний вид инструмента.

На рис. 1, в качестве примера, приведена последовательность выполнения технологических операций, которую целесообразно рассматривать для автоматизации процессов контроля качества изготовления корпусов гитар.

На каждом этапе выполнения технологической операции существует вероятность обнаружения брака, связанного с природными свойствами древесины и с особенностями механического и термического воздействия.

К числу наиболее распространенных дефектов древесины относятся:

- сучки – участки, образующиеся в местах выхода ветвей;
- трещины – разрывы структуры древесины, возникающие при сушке или механическом воздействии;
- вмятины – локальные деформации поверхности вследствие давления или удара.

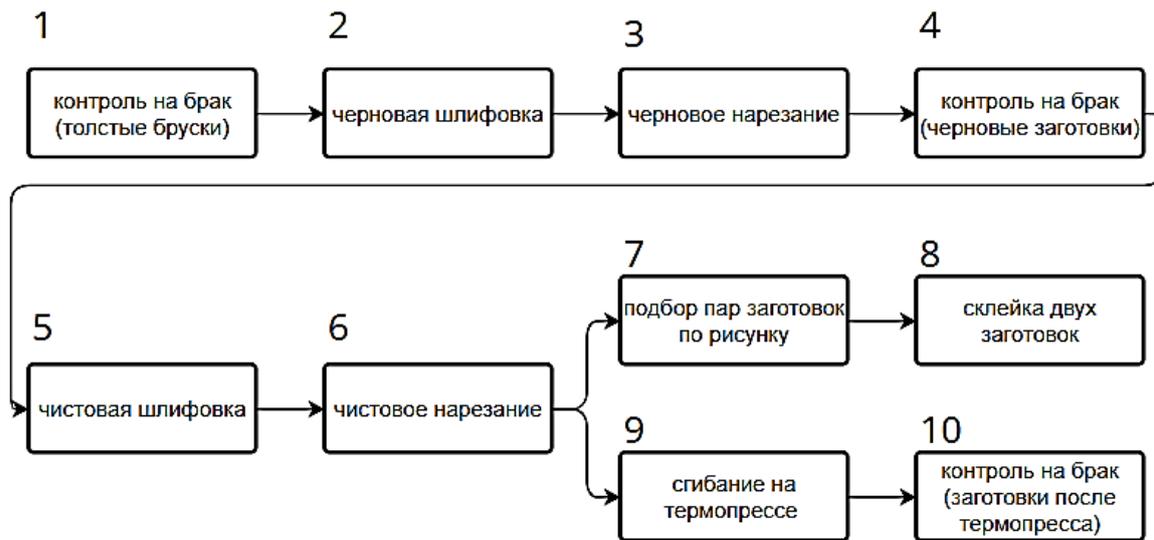


Рисунок 1 – Последовательность технологических операций для изготовления корпусов гитар

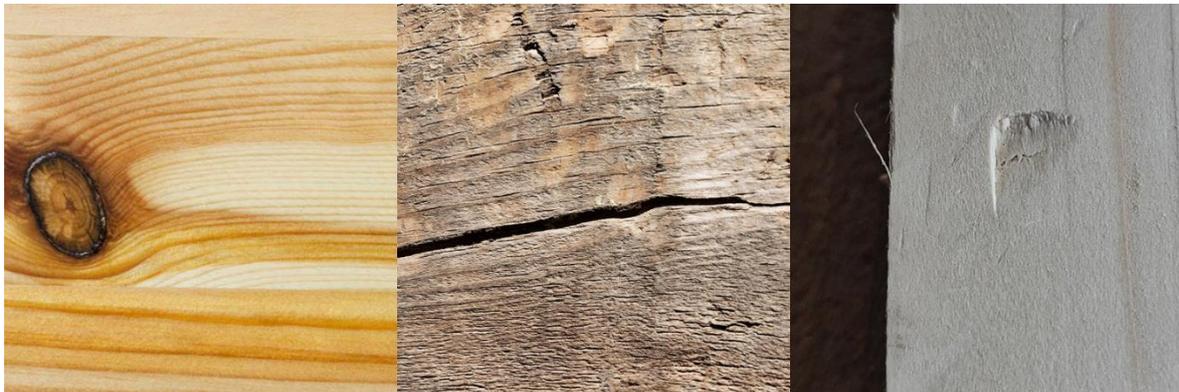


Рисунок 2 – Наиболее распространенные дефекты древесины

Существующие системы на производстве в настоящее время не содержат функционала поиска дефектов, и отбраковка происходит вручную даже на больших фабриках. Именно поэтому актуальной задачей является автоматизация прежде всего процесса обнаружения заготовок гитар, несоответствующих стандартам качества, с помощью нейронной сети, обученной для выявления дефектов.

Кроме этого, необходимо также автоматизировать процесс подбора пар заготовок с совпадающим рисунком древесины, поскольку при изготовлении каждой деки, состоящей из двух совмещаемых вместе деревянных заготовок, требуется обеспечить не только геометрическую симметрию, но и соответствие природного рисунка древесины. Совпадение текстуры поверхностей определяет эстетическую составляющую корпуса и воспринимаемое качество готового инструмента. На производствах гитар среднего и высокого ценового сегмента данная операция выполняется мастером на основе субъективной визуальной оценки [11].

Одним из наиболее эффективных подходов к решению таких задач является внедрение предлагаемой нами информационно-аналитической системы, способной обеспечить анализ изображений заготовок на этапах производства 1, 4 и 10 (в соответствии с рисунком 1) и автоматизировать подбор пар с максимально схожим рисунком и текстурой древесины на седьмом этапе 10 (в соответствии с рисунком 1).

2. Реализация информационно-аналитической системы

Архитектура информационно-аналитической системы (в соответствии с рисунком 3) построена по модульному принципу, основные компоненты которой обеспечивают сбор, обработку и анализ данных, а также взаимодействие с производственными системами и оператором. Модуль приёма изображений информационно-аналитической системы реализует захват кадров с камер машинного зрения и их передачу по протоколу GigE Vision и формирует буфер изображений для последующего анализа. Модуль анализа использует нейронные сети, реализованные с помощью фреймворка PyTorch, а также инструменты библиотеки OpenCV для детектирования и классификации дефектов деревянных заготовок. Результаты анализа передаются в модуль сбора и обработки данных, где сохраняются в базе данных на SQL сервере. Для визуализации и отчётности используется модуль отображения информации, написанный на JavaScript, и модуль формирования отчётов, реализованный на языке Python при помощи специализированных библиотек. Обмен сообщениями между компонентами системы осуществляется через брокер RabbitMQ. Для интеграции с производственными системами (SCADA, MRP) реализован модуль взаимодействия, использующий OPC UA клиент и шаблон Data Mapper, обеспечивающий передачу данных о производственных показателях.

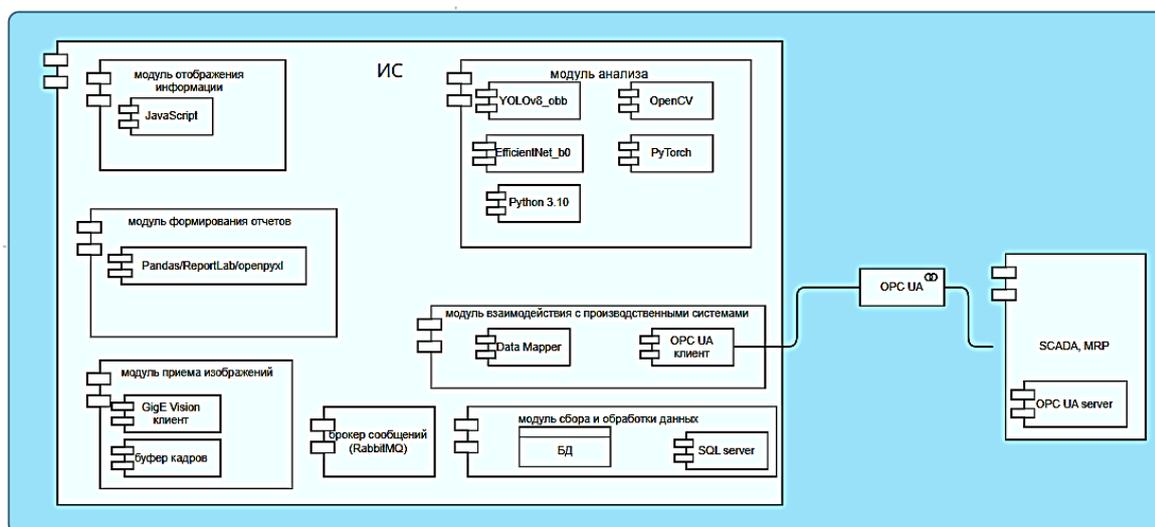


Рисунок 3 – Архитектура системы – программный слой

Для реализации модуля анализа, отвечающего за автоматизированное выявление дефектов деревянных заготовок, была выбрана нейросеть архитектуры YOLO. YOLO относится к семейству одноэтапных детекторов объектов – выполняет распознавание в едином свёрточном проходе [12]. Модели YOLO сочетают высокую скорость обработки изображений, возможности работы в реальном времени и относительно низкую вычислительную нагрузку, что делает их подходящими для внедрения в производственные системы.

Сбор данных для обучения модели осуществлялся вручную. Полученный набор изображений содержал в себе как деревянные панели с дефектами, так и без них для разнообразия данных и для устойчивости модели к вариациям внешних условий.

Для разметки изображений использовался специализированный инструмент CVAT (Computer Vision Annotation Tool), развёрнутый с применением технологии

Docker Compose. CVAT представляет собой веб-ориентированный инструмент для аннотирования данных, который поддерживает различные форматы экспорта, включая специфические структуры для обучения нейросетевых моделей. Интерфейс CVAT позволил создавать ограничивающие рамки (bounding boxes) вокруг участков с дефектами и назначать им соответствующие классы.

Для обучения YOLO требуется определённый формат данных. Датасет разделён на три подмножества: train, validation и test. Подмножество train используется непосредственно для обучения модели, validation – для контроля процесса обучения и подбора гиперпараметров, test – для независимой итоговой оценки точности. Каждое изображение в датасете сопровождается текстовым файлом, содержащим координаты ограничивающих рамок и метки классов.

В качестве модели использовалась YOLO-OBV (Oriented Bounding Box). В отличие от классической YOLO, которая ограничивается прямоугольными рамками, YOLO-OBV позволяет предсказывать повернутые ограничивающие рамки. Это критически важно для анализа деревянных заготовок, поскольку многие дефекты (например, трещины) имеют выраженную наклонную ориентацию. Обучение модели выполнялось в среде Python с использованием библиотеки Ultralytics, которая предоставляет удобные средства для конфигурации параметров обучения, мониторинга метрик и оценки результатов.

На рис. 4 представлены результаты обучения модели YOLOv8n-OBV для задачи обнаружения дефектов деревянных заготовок. Основными показателями качества являются Precision (P), Recall (R) и интегральные метрики mAP@50 и mAP@50-95. Значение Precision характеризует долю корректных предсказаний среди всех найденных объектов, Recall отражает полноту детекции, а mAP (mean Average Precision) оценивает совокупную точность модели при различных порогах пересечения областей (IoU).

Для класса knot (сучки) наблюдается высокая точность (P=0,794), полнота (R=0,9) и метрика mAP@50 (0,904), что свидетельствует о надёжной работе модели на данном типе дефекта. Для классов crack (трещины) и dent (вмятины) показатели ниже: Recall=0,5 для трещин и 0,69 для вмятин, что указывает на трудности при выявлении этих дефектов из-за их природных особенностей (нечёткие контуры, отсутствие контрастности) и возможного недостатка изображений для датасета.

YOLOv8n-obv summary (fused): 81 layers, 3,077,804 parameters, 0 gradients, 8.3 GFLOPs							
Class	Images	Instances	Box(P	R	mAP50	mAP50-95):	
all	29	152	0.759	0.697	0.755	0.513	
knot	20	30	0.794	0.9	0.904	0.734	
crack	16	80	0.703	0.5	0.595	0.203	
dent	18	42	0.781	0.69	0.765	0.603	

Рисунок 4 – Результаты обучения модели YOLOv8n-OBV

Изображения на рисунках 5 и 6 отображают результат задачи обнаружения дефектов. Дополнительно система решает задачу автоматизированного подбора пар деревянных заготовок для изготовления дек. Для этого применяются свёрточные нейросети, используемые в качестве экстракторов признаков. В частности, использовалась модель EfficientNet-B0, в которой применяется метод compound scaling, обеспечивающий одновременное масштабирование глубины, ширины и разрешения входных изображений.



Рисунок 5 – Обнаружение дефекта класса «сучок»

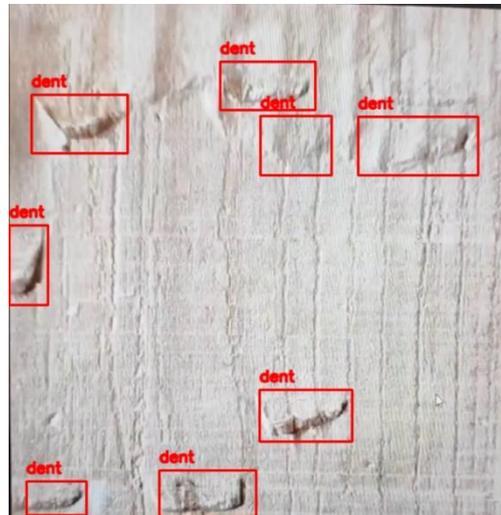


Рисунок 6 – Обнаружение дефекта класса «вмятина»

Архитектура EfficientNet строится на блоках MBConv, включает слои типа inverted residual и модули Squeeze-and-Excitation, что позволяет достигать высокой точности при сравнительно небольшом числе параметров и вычислительных затрат. Модель использовалась для преобразования изображения в вектор признаков высокой размерности, отражающих текстурные, цветовые и структурные особенности древесины.

На этапе предобработки изображения нормализуются и приводятся к единому размеру с использованием стандартных методов библиотеки torchvision. Полученные векторы признаков нормализуются и сравниваются с применением метрик близости (например, косинусного расстояния). На основе матрицы попарных сходств автоматически формируются пары изображений с наибольшей степенью текстурного соответствия. Алгоритм работы компонента информационно-аналитической системы для автоматизированного подбора пар деревянных заготовок представлен на рис. 7.

Результатом работы алгоритма является список пар в формате txt с наиболее схожими заготовками и их визуализированные комбинированные изображения (10 в соответствии с рисунком 8, 9), что обеспечивает объективность и воспроизводимость процесса подбора древесины.

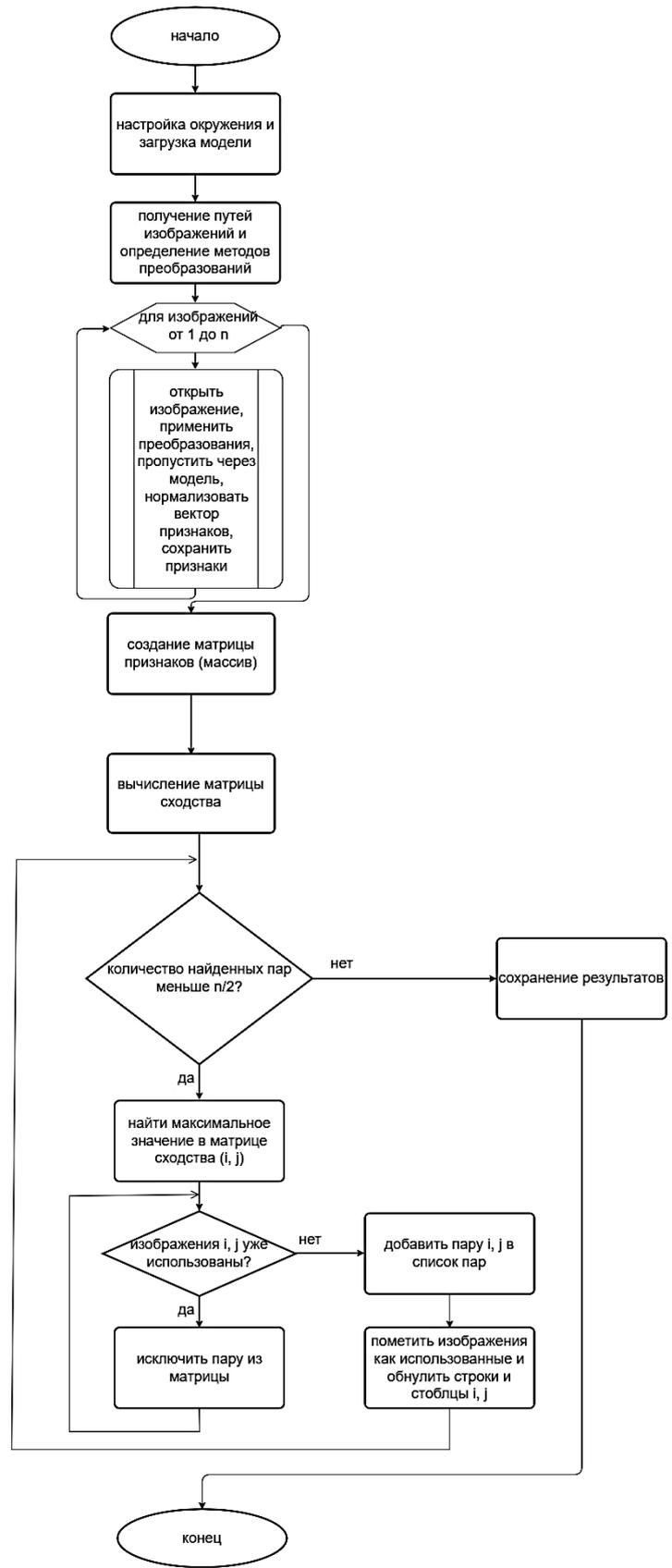


Рисунок 7 – Алгоритм работы компонента информационно-аналитической системы для автоматизированного подбора пар заготовок



Рисунок 8 – Результат работы экстрактора признаков, пара № 1



Рисунок 9 – Результат работы экстрактора признаков, пара № 2

3. Результаты исследования

Ключевые модули информационно-аналитической системы продемонстрировали возможности применения двух архитектур свёрточных нейронных сетей для решения задач обнаружения дефектов деревянных заготовок и автоматизированного поиска пар заготовок со схожими рисунками. Для задачи обнаружения дефектов применялась модель YOLOv8n-OBV, показавшая высокую точность на классе knot ($P=0,794$, $R=0,9$, $mAP@50=0,904$), что свидетельствует о надёжной работе модели на данном типе дефекта. Для классов crack и dent полнота оказалась ниже ($R=0,5$ и $R=0,69$ соответственно), что связано с более сложными условиями выявления этих дефектов и ограниченным объемом данных для обучения. Для задачи подбора пар деревянных заготовок использовалась модель EfficientNet-B0 в качестве экстрактора признаков. На основе векторов признаков высокой размерности, отражающих текстуру и цветовую структуру поверхности древесины, выполнялось сравнение изображений с использованием метрик близости. Результатом работы модуля является список пар с максимальным сходством и их визуализированные изображения.

Заключение

Применение информационно-аналитической системы для автоматизации контроля качества изготовления деревянных заготовок гитар решает актуальную задачу повышения качества выпускаемой продукции. Внедрение компонента обнаружения дефектов в модуль анализа позволяет выявлять дефекты на начальных стадиях выполнения технологических операций с помощью технологий машинного зрения и нейронных сетей и снизить вероятность появления брака на последующих. Использование модели нейронной сети YOLO-OBV демонстрирует целесообразность применения ориентированных ограничивающих рамок при анализе древесины, где дефекты могут иметь наклонную форму. Реализация компонента подбора пар заготовок на основе архитектуры EfficientNet в модуле анализа обеспечивает объективность и высокую скорость выбора и исключает субъективный фактор гитарного мастера. В перспективе данная система может быть интегрирована в производственные линии и взаимодействовать со SCADA, MRP и другими системами промышленной автоматизации. Это позволит повысить эффективность производственного процесса, минимизировать влияние человеческого фактора и обеспечить стабильное качество продукции.

Список литературы

1. Pansombut T., Wikaisuksakul S., Khongkraphan K., Aniruth P. Convolutional Neural Networks for Recognition of Lymphoblast Cell Images // *Computational Intelligence and Neuroscience*. 2019. Vol. 20. P. 1-12. DOI: 10.1155/2019/7519603.
2. Беловодский В.Н. Об использовании нейронных сетей к построению областей притяжения периодических режимов нелинейных динамических систем // *Проблемы искусственного интеллекта*. 2023. № 4 (31). С. 56-67. DOI: 10.34757/2413-7383.2023.31.4.006
3. Пикалёв Я.С., Ермоленко Т.В. О нейронных архитектурах извлечения признаков для задачи распознавания объектов на устройствах с ограниченной вычислительной мощностью // *Проблемы искусственного интеллекта*. 2023. № 3 (30). С. 44-54. DOI: 10.34757/2413-7383.2023.30.3.004
4. Зуев В. М. Сравнение обнаружения объектов средствами искусственного интеллекта в сравнении с классическими методами // *Проблемы искусственного интеллекта*. 2024. № 3(34). С. 30-35. DOI: 10.24412/2413-7383-2024-3-30-35
5. Бабичева М.В., Третьяков И.А. Автоматизация процедуры распознавания фальшивых изображений посредством нейронных сетей // *Проблемы искусственного интеллекта*. 2025. № 1 (36). С. 94-105. DOI: 10.24412/2413-7383-94-105
6. Андриянов Н.А., Дементьев В.Е., Ташлинский А.Г. Обнаружение объектов на изображении: от критериев Байеса и Неймана-Пирсона к детекторам на базе нейронных сетей EfficientDet // *Компьютерная оптика*. 2022. Т. 46. № 1. С. 139-159. DOI: 10.18287/2412-6179-CO-92
7. Володин И.В., Путятю М.М., Макарян А.С., Евглевский В.Ю. Классификация механизмов атак и исследование методов защиты систем с использованием алгоритмов машинного обучения и искусственного интеллекта // *Прикаспийский журнал: управление и высокие технологии*. 2021. № 2(54). С. 91-98. DOI: 10.21672/2074-1707.2021.53.1.090-098
8. Беловодский В.Н. Об интерполяции и аппроксимации функций с использованием нейронных сетей // *Проблемы искусственного интеллекта*. 2024. № 3 (34). С. 6-19. DOI: 10.24412/2413-7383-2024-3-4-19
9. Харламов А., Самаев Е., Кузнецов Д., и др. Семантический анализ текста с использованием искусственных нейронных сетей на основе нейроподобных элементов с временным суммированием сигналов // *Проблемы искусственного интеллекта*. 2023. № 3 (30). С. 4-27. DOI: 10.34757/2413-7383.2023.30.3.001.
10. Guitar for Nerds / Neale Morison [Электронный ресурс] // Нотный архив христианской музыки: [сайт]. Режим доступа: https://noty.propovednik.com/Public/_УЧЕБНИКИ/ГИТАРА/Guitar%20For%20Nerds.pdf, свободный. Загл. с экрана. Дата обращения: 10-11-2025.
11. Как делают гитары Maton. Тур по фабрике | gitaraclub.ru / Гитарный клуб [Электронный ресурс] // YouTube.com: [видеохостинг]. Режим доступа: <https://www.youtube.com/live/IHReR19aEzI>, свободный. – Загл. с экрана. Дата обращения: 15-01-2023.

12. Redmon J., Divvala S.K., Girshick R.B., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 779-788. DOI:10.1109/CVPR.2016.91. – URL: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf

References

1. Pansombut T., Wikaisuksakul S., Khongkraphan K., Aniruth P. Convolutional Neural Networks for Recognition of Lymphoblast Cell Images // Computational Intelligence and Neuroscience. 2019. Vol. 20. P. 1-12. DOI: 10.1155/2019/7519603.
2. Belovodsky V.N. On the use of neural networks to construct areas of attraction for periodic modes of nonlinear dynamic systems // Problems of artificial intelligence. 2023. No. 4 (31). P. 56-67. DOI: 10.34757/2413-7383.2023.31.4.006
3. Pikalev Ya.S., Ermolenko T.V. On neural architectures for feature extraction for the task of object recognition on devices with limited computing power // Problems of artificial Intelligence. 2023. No. 3 (30). P. 44-54. DOI: 10.34757/2413-7383.2023.30.3.004.
4. Zuev V. M. Comparison of object detection by means of artificial intelligence in comparison with classical methods // Problems of artificial intelligence. 2024. No. 3(34). P. 30-35. DOI: 10.24412/2413-7383-2024-3-30-35.
5. Babicheva M.V., Tretyakov I.A. Automation of the procedure for recognizing fake images using neural networks // Problems of artificial intelligence. 2025. No. 1 (36). P. 94-105. DOI: 10.24412/2413-7383-94-105.
6. Andrianov N.A., Dementyev V.E., Tashlinsky A.G. Object detection in an image: from Bayes and Neiman-Pearson criteria to detectors based on EfficientDet neural networks // Computer Optics. 2022. Vol. 46. No. 1. P. 139-159. DOI: 10.18287/2412-6179-CO-92.
7. Volodin I.V., Putyato M.M., Makaryan A.S., Evgenevsky V.Yu. Classification of attack mechanisms and research of system protection methods using machine learning and artificial intelligence algorithms // Caspian Journal: Management and High Technologies. 2021. No. 2(54). P. 91-98. DOI: 10.21672/2074-1707.2021.53.1.090-098.
8. Belovodskiy V.N. On interpolation and approximation of functions using neural networks // Problems of artificial intelligence. 2024. No. 3 (34). P. 6-19. DOI: 10.24412/2413-7383-2024-3-4-19.
9. Kharlamov A., Samaev E., Kuznetsov D., and others. Semantic text analysis using artificial neural networks based on neural-like elements with temporal summation of signals // Problems of artificial intelligence. 2023. No. 3 (30). P. 4-27. DOI: 10.34757/2413-7383.2023.30.3.001.
10. Guitar for Nerds / Neale Morison – [Electronic resource] // Music archive of Christian music: [website]. – Access mode: https://noty.propovednik.com/Public/_УЧЕБНИКИ/ГИТАРА/Guitar%20For%20Nerds.pdf, free. – Caption from the screen. Date of request: 10-11-2025.
11. How Maton guitars are made. Factory Tour | gitaraclub.ru / Guitar Club – [Electronic resource] // YouTube.com: [video hosting]. – Access mode: <https://www.youtube.com/live/IHRrR19aEzI>, free. – Caption from the screen. Date of request: 15-01-2023.
12. Redmon J., Divvala S.K., Girshick R.B., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 779-788. DOI:10.1109/CVPR.2016.91. – URL: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf

RESUME

*Logunova D.M., Volodina A.M., Kurnasov E.V., Kirillov S.K., Tonkonozhenko N.D.
The use of convolutional neural networks in the manufacture of guitars to identify product defects and automated selection of pairs of blanks*

The study is devoted to issues that arise in the guitar manufacturing process and are related to the lack of effective methods for detecting product defects. It is shown that to date, the detection of any defects in guitar blanks in production is carried out by the operator through visual assessment, which reduces the stability of the quality of these products and increases the risk of defects. The architecture of an information and analytical system with an image analysis module based on the use of neural networks of the YOLO and EfficientNet architectures is proposed. For the task of detecting wood defects (knots, cracks, dents), the

YOLOv8n-OBb model was used, which allows taking into account the orientation of defects due to rotated bounding boxes. The model demonstrated high accuracy on the class of knots ($P=0.794$, $R=0.9$, $mAP@50=0.904$). The EfficientNet-B0 model was used as a feature extractor for the automated selection of pairs of blanks with a similar wood pattern. The images were compared using similarity metrics of vector representations, which made it possible to objectively form pairs with the greatest textural similarity. The algorithm of the system for automated selection of pairs of workpieces is presented. The proposed system can be integrated with industrial SCADA and MRP systems, ensuring timely detection of defects, standardization of the selection of blanks for guitar decks and reducing the influence of the human factor. The work demonstrates the prospects of using machine vision technologies and neural networks in the woodworking industry and their applicability in the tasks of quality control of musical instruments.

РЕЗЮМЕ

*Логунова Д.М., Володина А.М., Курнасов Е.В., Кириллов С.К., Тонконоженко Н.Д.
Применение свёрточных нейронных сетей на производстве гитар
для выявления дефектов продукции и автоматизированного подбора пар
заготовок*

Исследование посвящено вопросам, возникающим в процессе изготовления гитар и связанным с отсутствием эффективных методов выявления дефектов продукции. Показано, что на сегодняшний день обнаружение каких-либо дефектов заготовок гитар на производстве выполняется оператором посредством визуальной оценки, что снижает стабильность качества данной продукции и повышает риск брака.

Предложена архитектура информационно-аналитической системы с модулем анализа изображений, основанном на применении нейросетей архитектур YOLO и EfficientNet. Для задачи обнаружения дефектов древесины (сучков, трещин, вмятин) использовалась модель YOLOv8n-OBb, позволяющая учитывать ориентацию дефектов за счёт повернутых ограничивающих рамок. Модель продемонстрировала высокую точность на классе сучков ($P=0,794$, $R=0,9$, $mAP@50=0,904$). Для автоматизированного подбора пар заготовок со схожим рисунком древесины применялась модель EfficientNet-B0 в качестве экстрактора признаков. Сравнение изображений осуществлялось по метрикам сходства векторных представлений, что позволило объективно формировать пары с наибольшей текстурной схожестью. Представлен алгоритм работы системы для автоматизированного подбора пар заготовок.

Предложенная система может быть интегрирована с промышленными SCADA и MRP системами, обеспечивая своевременное выявление брака, стандартизацию подбора заготовок для деки гитар и снижение влияния человеческого фактора. Работа демонстрирует перспективность использования технологий машинного зрения и нейронных сетей в деревообрабатывающей промышленности и их применимость в задачах контроля качества музыкальных инструментов.

Логунова Дарья Михайловна, студент Института искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, dm.logunova@gmail.com

Область научных интересов: искусственный интеллект, экспертные системы, умные производственные системы, обработка информации, распознавание образов, кибернетика.

Володина Анна Михайловна, старший преподаватель кафедры промышленной информатики Института искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, volodina@mirea.ru

Область научных интересов: искусственный интеллект, экспертные системы, базы данных, онтологические модели, обработка естественного языка (NLP), рекомендательные системы, большие данные, умные производственные системы, автоматизированные системы управления.

Курнасов Евгений Вячеславович, кандидат технических наук, доцент кафедры промышленной информатики Института искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, kurnasov@mirea.ru

Область научных интересов: искусственный интеллект, предиктивный анализ данных, системная интеграция в области промышленной автоматизации и управления, разработка интеллектуальных систем управления технологическими объектами и процессами, умные IT-решения для промышленности.

Кириллов Сергей Константинович, студент Института искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, kirillov.s342@gmail.com

Область научных интересов: искусственный интеллект, экспертные системы, умные производственные системы, обработка информации, распознавание образов, кибернетика.

Тонконоженко Никита Денисович, студент Института искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, tonkononikita@gmail.com

Область научных интересов: искусственный интеллект, экспертные системы, умные производственные системы, обработка информации, распознавание образов, кибернетика.

Статья поступила в редакцию 07.10.2025.

Раздел 4
Section 4

Методы и системы
защиты информации,
информационная безопасность

*Methods and systems
information protection,
information security*

УДК 681.5:621.37:004.056

DOI 10.24412/2413-7383-2025-4-39-150-161

Я. И. Рушечников, И. А. Третьяков, В. В. Данилов
Федеральное государственное бюджетное образовательное учреждение
высшего образования «Донецкий государственный университет», г. Донецк
283001, г. Донецк, ул. Университетская, 24

АДАПТИВНАЯ АВТОМАТИЗИРОВАННАЯ СИСТЕМА ПОДАВЛЕНИЯ ПОБОЧНЫХ ЭЛЕКТРОМАГНИТНЫХ ИЗЛУЧЕНИЙ И НАВОДОК *

Ya. I. Rushechnikov, I. A. Tretiakov, V. V. Danilov
Federal State Budgetary Educational Institution of Higher Education
"Donetsk State University", Donetsk
283001, Donetsk, st. Universitetskaia, 24

ADAPTIVE AUTOMATED SYSTEM FOR SUPPRESSION OF ELECTROMAGNETIC SIDE EFFECTS AND INTERFERENCE

В настоящей работе описана комплексная методика поиска и последующего подавления обнаруженных сигналов побочных электромагнитных излучений и наводок в рамках единого программно-аппаратного комплекса подавления с использованием обратной связи. Предложенная методика позволит использовать самые доступные программно-определяемые платформы для сканирования радиообстановки, что снизит конечную стоимость адаптивной автоматизированной системы подавления таких сигналов.

Ключевые слова: АСНИ, ПЭМИН, радиомониторинг, модельный сигнал, SDR, подавление сигналов, генератор SI5351, коэффициент совпадения, корреляция

This paper describes a comprehensive technique for searching for and subsequent suppression of detected electromagnetic side effects and interference signals within a single feedback suppression software and hardware package. The proposed technique will make it possible to use the most accessible software-defined platforms for scanning radio placement, which will reduce the final cost of an adaptive automated system for suppressing such signals.

Key words: ASRS, TEMPEST, radio monitoring, model signal, SDR, signal suppression, SI5351 generator, coincidence coefficient, correlation

* Исследование проводилось в рамках НИР «Исследование природы каналов побочных электромагнитных излучений и наводок элементов и устройств офисной вычислительной техники» (регистрационный номер 124012400347-2).

Введение

Остро стоящие в данный момент проблемы автоматизированного радиомониторинга и радиопротиводействия имеют непосредственное влияние и на информационно-измерительные системы, основная цель которых не передача, а непосредственная обработка информации [1-5]. Такие системы в виде побочных электромагнитных излучений и наводок (ПЭМИН) непреднамеренно распространяют информацию как о своём поведении, так и о своих внутренних процессах. Подобного рода информационное воздействие фиксируется при помощи методов, описанных в [6-10], а противодействие или подавление реализуется механизмами, рассмотренными в [11-17], что в комплексе позволяет решать задачу детектирования и подавления ПЭМИН, но за счёт экспериментальной сложности реализует этот процесс условно не оптимальным способом. Из-за множества и сложности вычислений решение этой задачи невозможно без использования интеллектуальных инструментов идентификации и создания специализированных автоматизированных систем, таких как автоматизированные системы научных исследований (АСНИ) радиоэлектронной обстановки [1-5]. Наиболее перспективной оптимизацией предложенных в [6-8], [16], [17] решений является синтез адаптивной системы с обратной связью, которая в непрерывном режиме будет заниматься периодическим тестированием наличия ПЭМИН и подавления наиболее выраженных спектральных составляющих с последующей оценкой качества подавления.

Принцип детектирования сигнала ПЭМИН

В случаях пренебрежения фактором задач электромагнитной совместимости при разработке разнообразных устройств, побочные излучения, в той или иной мере будут проявлять себя на определённых частях частотного спектра. Иногда эти частоты явно коррелированы программно-аппаратными технологиями (например, с технологией HDMI), но очень часто эти частоты могут быть смещены в большую или меньшую сторону, что связано с несовершенством как аппаратных частей, так и кабельной продукции, реализующей информационную технологию. Задачи по выявлению частот ПЭМИН активно решаются как в научно-инженерной практике в виде разнообразных исследований, так и в конкретно прикладной области, при помощи специальных программно-аппаратных систем, примером которой является комплекс «Навигатор-ПЗМ» [9], [13-18].

Описанные выше способы в некоторых случаях качественно, в некоторых случаях с инструментальной точностью определяют разнообразные характеристики сигнала, что для задач детектирования и подавления может быть несколько излишне, как по производительности, так и по объёму метаинформации о сигнале.

Оптимизация процесса автоматизированного поиска может заключаться в запуске определённого модельного сигнала по типу телевизионной настроечной таблицы и сканирование диапазона частот, коррелированного с технологией, через которую происходит эманация побочного излучения. Этот диапазон не сильно широкий, а также относительно его можно найти и гармоники сигналов. Предложенная методика позволит использовать самые доступные программно-определяемые платформы (SDR) для сканирования, что, несомненно, снизит конечную стоимость адаптивной автоматизированной системы подавления ПЭМИН.

На рисунке 1 представлена блок-схема алгоритма, реализующего поиск частот ПЭМИН методом сканирования эфира и сравнения звуковых отпечатков модельного сигнала с сигналом, записанным на просканированных частотах. Приём и демодуляция сигнала позволит по представленному на рисунке 1 алгоритму отнести сигнал к группе ПЭМИН даже в условиях электромагнитной обстановки средней жёсткости.



Рисунок 1 – Блок-схема алгоритма поиска частот ПЭМИН

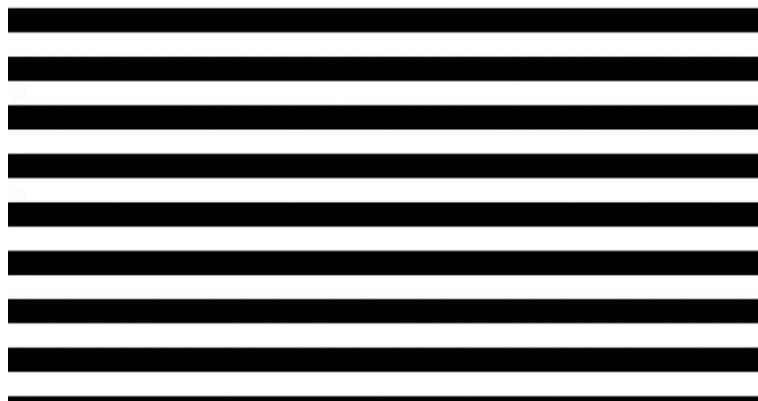


Рисунок 2 – Графическое представление модельного сигнала

В качестве модельного сигнала использовалась специальная видеопоследовательность, которая запускалась на персональном компьютере, и представляла из себя набор чередующихся горизонтальных линий (рисунок 2), которые после демодуляции и записи на SDR системе интерпретируются оператором как звук. Записанный и оцифрованный звук является тем самым модельным сигналом.

По сути наборы из таких комбинаций модельных сигналов, схожих с настроечной таблицей кинескопных телевизоров, можно сформировать в пакет, аналоги которой в отрасли информационной безопасности называются «Сигурд» [8], [18], [19].

Спектр модельного сигнала воспринимаемый SDR системой и интерпретированный представленным выше алгоритмом как ПЭМИН имеет форму, представленную на рисунке 3. Из представленного снимка спектра видно не только максимальную по уровню несущую ПЭМИН, но также много групповых составляющих, которые в динамике изменяют свои энергетические характеристики. Кроме того, можно заметить широкополосность сигнала, которая не вмещается в обозреваемую полосу приёмника.

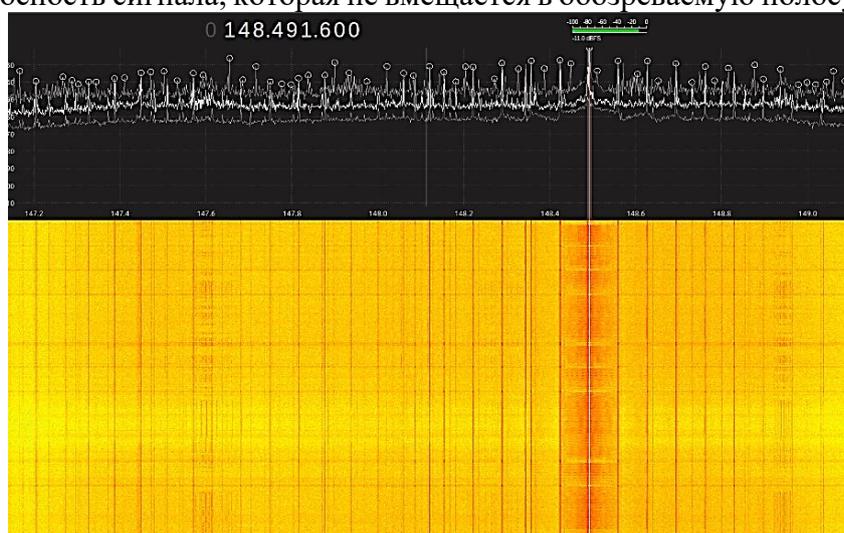


Рисунок 3 – Спектр ПЭМИН, полученный в результате воспроизведения модельного сигнала

Сопоставление модельного сигнала с его демодулированным образом происходит при помощи быстрого преобразования Фурье и алгоритмической задачи поиска подпоследовательности в последовательности. Факт наличия подпоследовательности в последовательности однозначно определяет наличие ПЭМИН на частоте, с которой был записан образец. Предложенный алгоритм позволяет решать задачу определения частот ПЭМИН на абсолютно любой SDR платформе, начиная от самых доступных, например RTL SDR.

Методика подавления ПЭМИН

Задачи подавления радиосигналов, в том числе и ПЭМИН, решены многократно при помощи большого разнообразия средств и методик. Основополагающим фактором являются следующие параметры:

1. Наличие конфигурируемого передатчика, способного перестраиваться в широких пределах по частоте.
2. Уровень мощности передатчика должен быть сопоставим с уровнем подавляемого сигнала.
3. Передатчик должен обладать стабильностью частоты, или её флуктуации не должны негативным образом сказываться на подавлении.

В качестве подобной системы гибкой генерации сигналов целесообразно использовать синтезатор частоты на базе микросхемы SI5351 (рис. 4), который путём несложных манипуляций позволяет опираясь на опорный сигнал кварцевого резонатора получить один или более выходных сигналов с частотой до 160 МГц, чего вполне достаточно для подавления ПЭМИН на частотах около 150 МГц. Кроме того, программируемый генератор позволяет менять параметры выходного сигнала в широких пределах во время работы.

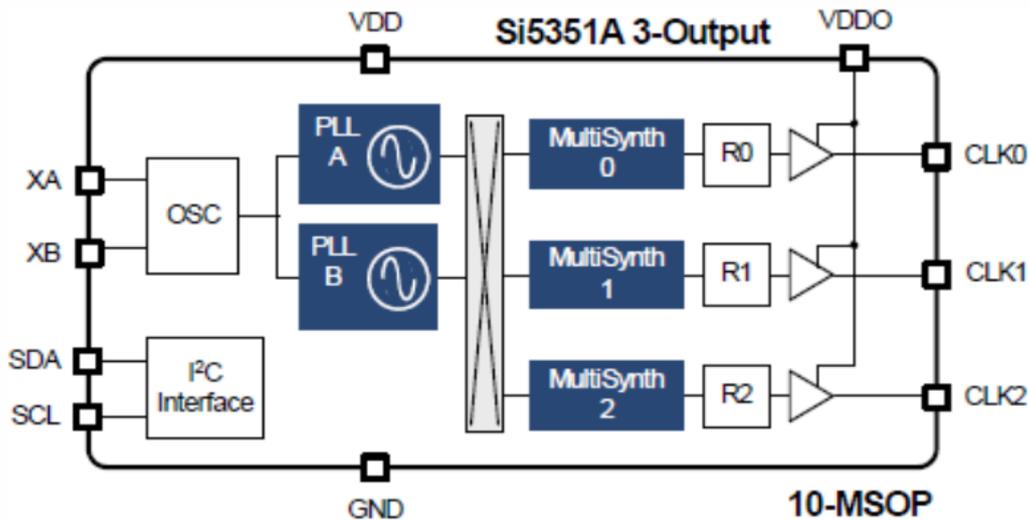


Рисунок 4 – Структурная схема конфигурируемого генератора SI5351

Принцип работы схемы, представленной на рисунке 4 следующий:

- В качестве источника тактирования используется внешний кварцевый осциллятор.
- Тактовый сигнал попадает на цепи фазовой автоподстройки частоты, на выходе которой получается высокая промежуточная частота, порядка 900 МГц.
- Следующая стадия – дробление сигнала при помощи специальных целочисленных делителей, благодаря которым есть возможность настроиться с высокой точностью на генерацию в пределах от 2,5 кГц до 160 МГц.
- Последним функциональным блоком схемы является маршрутизирующая матрица, которая позволяет распределить сформированные тактовые сигналы по нескольким выходам.

Отличительной способностью данного генератора является наличие функции расширения спектра, что позволит генерировать условно широкополосный стабильный сигнал без использования свипирования по времени. Формат расширения спектра представлен на рисунке 5.

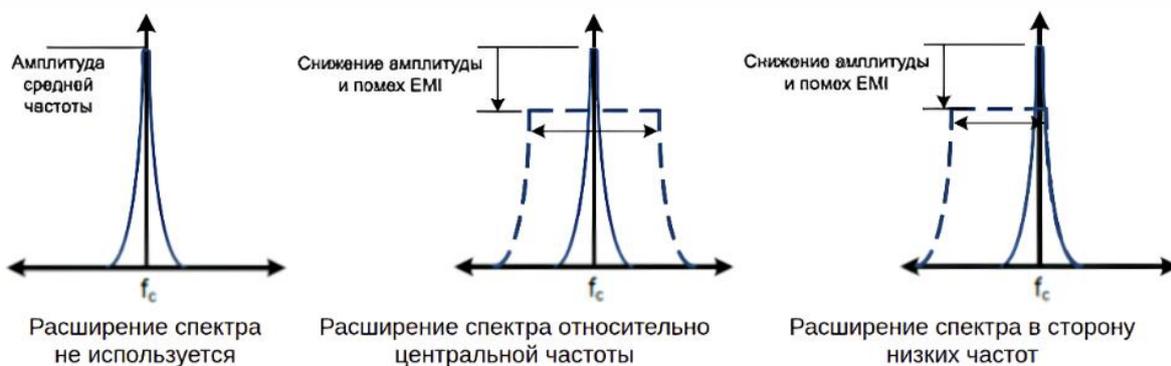


Рисунок 5 – Схемы расширения спектра SI5351

Представленный на рисунке 6 модуль за счёт реализованного согласования на нагрузку в 50 Ом позволяет получить на одном из выходов генератора мощность в эквиваленте не более 15мВт составляющую порядка 12 дБм, что значительно больше, нежели сигнал ПЭМИН с порядком мощности -60 дБм.

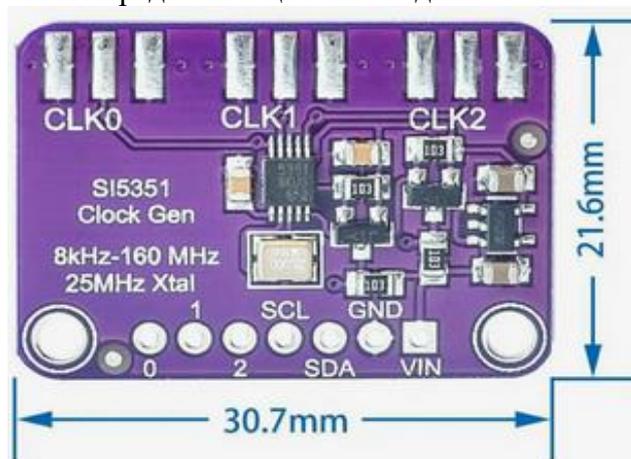


Рисунок 6 – Модуль генератора SI5351

Используемый в качестве генератора модуль SI5351 управляется по интерфейсу I2C проходя следующие конфигурационные стадии:

1. Тест доступности генератора.
2. Выбор канала для генерации (мощность вещания пропорционально делится между каналами).
3. Назначение коэффициентов коррекции и предделителей для получения промежуточной и результирующей частоты.
4. Передача тактирования на выход, согласованный с 50 Ом.

При реализации сигнала подавления, явно влияющего на ПЭМИН, можно получить следующий спектр, представленный на рисунке 7.

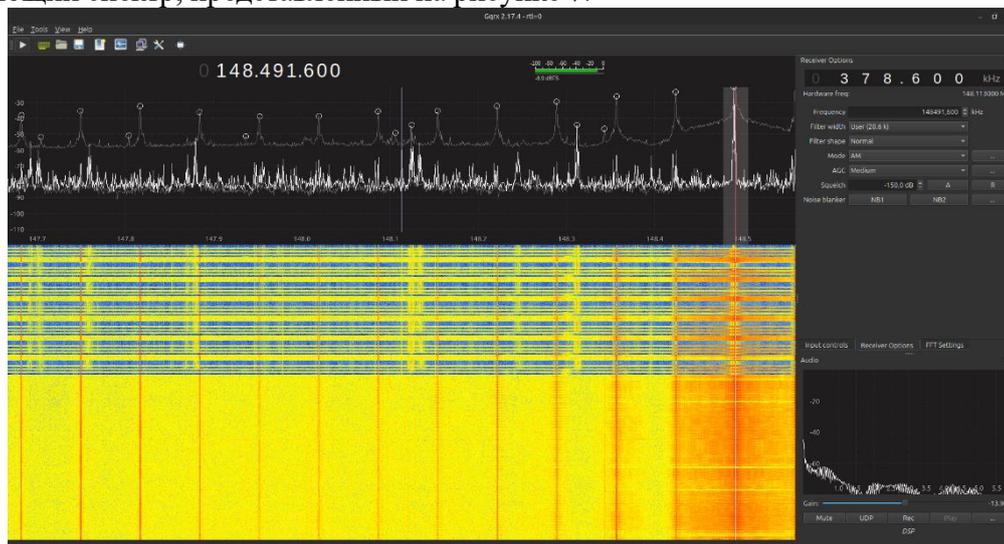


Рисунок 7 – Спектр работы подавителя сигнала ПЭМИН

Как видно из спектрограммы в соответствии с рисунком 7, мешающий сигнал с определённым периодическим паттерном перебивает сигнал ПЭМИН, что явным образом снижает вероятность декодирования и распознавания информационных составляющих сигнала при помощи современных SDR систем.

Систематизация обратной связи и оценка эффективности подавления

Описанные выше концепции, такие как поиск ПЭМИН при помощи алгоритма сопоставления сигнала с модельным и также средства подавления сигнала можно увязать в единую информационно-управляемую систему с обратной связью, состоящую из следующих элементов:

- Объект управления – процесс поиска и подавления ПЭМИН
- Датчик - перестраиваемый генератор.
- Сравнивающее устройство - компьютер, определивший или классифицировавший сигнал как ПЭМИН.
- Регулятор - микроконтроллер и генератор, реализующие концепцию адаптивного подавления.
- Обратная связь - канал, по которому информация о результате возвращается к регулятору.

Концептуальный вид системы представлен на рисунке 8, содержащем схему, работающую по следующему алгоритму:

1. На исследуемо-управляемом компьютере запускается воспроизведение модельного «Сигурд» сигнала.
2. SDR система ищет частоты методом сканирования, записи и сопоставления сигнала с первообразного модельного сигнала.
3. В случае классификации сигнала как ПЭМИН в момент сравнения с моделью происходит отправка команды генератору, чтобы тот инициировал процедуру подавления на определённой частоте.

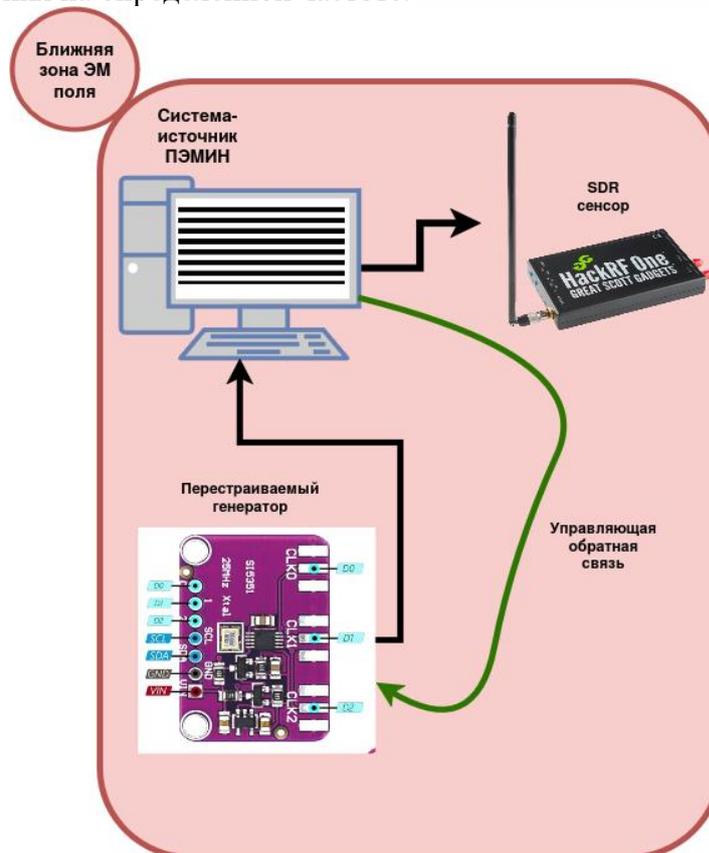


Рисунок 8 – Функциональная схема адаптивной автоматизированной системы подавления ПЭМИН

Немаловажным этапом процесса подавления сигнала ПЭМИН является выработка механизма оценки характеристик качества подавления. Несмотря на то, что в результате трансляции сигнала подавления в ближнюю зону действия ПЭМИН программное средство «TempestSDR» не в состоянии синхронизировать и воспроизвести картинку, происходящую на мониторе (то есть, по сути, сигнал подавлен), необходимо выработка механизма количественной оценки в виде коэффициента совпадения.

В качестве такого коэффициента совпадения можно использовать результат корреляции записи модельного сигнала ПЭМИН с записью формата модельный сигнал + сигнал подавления. Такой коэффициент покажет, насколько оригинальный сигнал отличается от группового, и если такой коэффициент малоразмерный, то эффективность мероприятия по подавлению была высокая, и наоборот. Описанная задача уже многократно решена и имеет своё название «Получение и детерминация отпечатков звуковых файлов» [20], [21], а также имеет качественную, доступную и переносимую реализацию на языке Python.

Представленные на рисунке 9а и 9б диаграммы совпадения, из которых можно взять глобальный максимум могут являться относительным индикатором качества мероприятий по подавлению, а сам выраженный коэффициент имеет инверсный порядок (чем меньше размерность коэффициента, тем выше качество подавления).

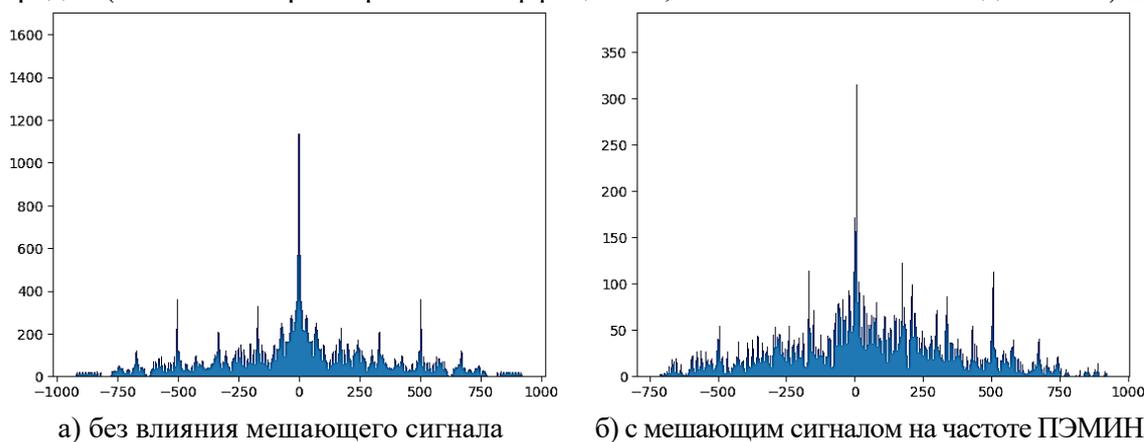


Рисунок 9 – Диаграммы совпадения

Как видно из рисунка 9, без влияния мешающего сигнала размерность максимума диаграммы значительно выше, нежели с его влиянием, что явным образом коррелирует с возможностью получения информационных составляющих анализируемого ПЭМИН сигнала.

Заключение

Разработка комплексных, доступных, автоматизированных систем подавления сигналов непреднамеренных каналов утечки информации, к которым относят ПЭМИН, является востребованной научно-технической задачей, результаты решения которой могут быть внедрены в системы, предназначенные для компьютеризированной обработки конфиденциальной информации. Данный класс адаптивных автоматизированных систем предназначен и актуален исключительно для компьютерных систем, для интерфейсной номенклатуры которых ещё не решена задача полной электромагнитной совместимости, ведь в случае её решения функции подавления станут невостребованными.

Таким образом, описана методика описана комплексная методика поиска и последующего подавления обнаруженных сигналов ПЭМИН, позволяющая применять самые доступные программно-определяемые платформы для сканирования радиообстановки, что существенно снижает стоимость адаптивной автоматизированной системы подавления таких сигналов.

Список литературы

1. Ашихмин, А. В. Автоматизированная система радиомониторинга / А. В. Ашихмин, В. А. Козьмин, А. М. Рембовский. *Спецтехника и связь*. 2012. № 1. С. 43-50. EDN OXUPNJ.
2. Соловьев, А. М. О результатах разработки автоматизированной системы научных исследований средств мониторинга радиоэлектронной обстановки в АСУТП / А. М. Соловьев, А. А. Головин, Р. К. Кинденев. *Технические науки - от теории к практике*. 2015. № 53. С. 80-88. EDN VDWRPV.
3. Рушечников, Я. И. Информационная технология радиомониторинга на основе программно-определяемой радиосистемы / Я. И. Рушечников, В. В. Данилов. *Вестник Донецкого национального университета. Серия Г: Технические науки*. 2020. № 1. С. 31-36. EDN PXYYTX.
4. Третьяков, И. А. Функциональные блоки для АСНИ радиосигналов / И. А. Третьяков, В. В. Данилов. *Донецкие чтения 2024: образование, наука, инновации, культура и вызовы современности: материалы IX Международной научной конференции (Донецк, 15–17 октября 2024 г.)*. Т. 2. Ч. 2. Донецк: Изд-во ДонГУ, 2024. С. 88-91. EDN ZEFZIV.
5. Элементы автоматизированных систем научных исследований радиоэлектронной обстановки / И. А. Третьяков, М. В. Бабичева, Я. И. Рушечников, В. В. Данилов. *Проблемы искусственного интеллекта*. 2025. № 3(38). С. 172-181. DOI 10.24412/2413-7383-2025-3-38-172-181. EDN HKDOZN.
6. Хорев, А. А. Оценка возможности обнаружения побочных электромагнитных излучений видеосистемы компьютер / А. А. Хорев. *Доклады Томского государственного университета систем управления и радиоэлектроники*. 2014. № 2(32). С. 207-213. EDN SEBGWX.
7. Железняк, В. К. Методика оценки защищенности видеoinформации ШИМ-преобразователя средств вычислительной техники / В. К. Железняк, С. В. Харченко. *Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки*. 2020. № 4. С. 23-30. EDN JRRBRM.
8. Шпилевой, А. А. О возможности аналитического обнаружения сигнала ПЭМИН в видеоинтерфейсах стандарта HDMI / А. А. Шпилевой, А. А. Персичкин. *Вестник Балтийского федерального университета им. И. Канта. Серия: Физ.-мат. и техн. науки*. 2022. № 1. С. 34-39. EDN AWTJYC.
9. Методика автоматизированной оценки объёма побочных электромагнитных излучений / Я. И. Рушечников, В. В. Данилов, И. А. Третьяков, В. А. Ступак. *Вестник Донецкого национального университета. Серия Г: Технические науки*. 2024. № 2. С. 51-57. DOI 10.5281/zenodo.12548548. EDN UNOZKU.
10. Свидетельство о государственной регистрации программы для ЭВМ № 2025688246 Российская Федерация. Программа для определения качественных характеристик побочных электромагнитных излучений и наводок элементов и устройств офисной вычислительной техники : заявл. 26.05.2025 : опубл. 17.10.2025 / Я. И. Рушечников, И. А. Третьяков; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий государственный университет». EDN HSARZF.
11. Кулешов, И. А. Анализ способов некриптографической защиты информации от побочных электромагнитных излучений и наводок на объектах связи / И. А. Кулешов, А. И. Спивак, А. Е. Львов. *Техника средств связи*. 2021. № 3(155). С. 67-73. EDN SXQOTD.
12. Паршуткин, А. В. Повышение защищенности информации от утечки через побочные электромагнитные излучения / А. В. Паршуткин, М. Р. Неаскина. *Вопросы кибербезопасности*. 2022. № 3(49). С. 82-89. DOI 10.21681/2311-3456-2022-3-82-89. EDN ARQCJO.
13. Язов, Ю. К. Пути построения многоагентной системы защиты информации от утечки по техническим каналам / Ю. К. Язов, А. О. Авсентьев. *Вопросы кибербезопасности*. 2022. № 5(51). С. 2-13. DOI 10.21681/2311-3456-2022-5-2-13. EDN DGYOEF.
14. Паршуткин, А. В. Экспериментальные исследования возможности применения программно-реализуемых мер снижения информативности побочных электромагнитных излучений видеосистемы стандарта DVI / А. В. Паршуткин, М. Р. Неаскина, И. П. Степанов. *СПБНТОРЭС: труды ежегодной НТК*. 2023. № 1(78). С. 126-129. EDN LAZWZX.

15. Язов, Ю. К. Проблемные вопросы управления защитой информации от утечки по техническим каналам с применением многоагентных систем / Ю. К. Язов, А. О. Авсентьев. *Вопросы кибербезопасности*. 2024. № 6(64). С. 85-97. DOI 10.21681/2311-3456-2024-6-85-97. EDN DDTRVS.
16. Алгоритм генерации высокочастотной ШИМ-помехи / Я. И. Рушечников, И. А. Третьяков, В. А. Ступак, Д. Р. Зеленченко. *69-я Международная научная конференция Астраханского государственного технического университета (Астрахань, 26–30 мая 2025 г.): материалы*. Астрахань: Изд-во АГТУ, 2025. С. 188-190. EDN VUWJJV.
17. Третьяков, И. А. Подавление ПЭМИН в видеоинтерфейсах посредством высокочастотной ШИМ помехи / И. А. Третьяков, Я. И. Рушечников, В. В. Данилов. *Информационные системы и технологии: материалы XI международного научного конгресса по информатике. В 2 ч. (Минск, 29–31 октября 2025 г.)*. Ч. 1. Минск: БГУ, 2025. С. 111-114.
18. Бедердинова, О. И. Оценивание защищенности конфиденциальной информации от утечки по техническим каналам / О. И. Бедердинова, И. В. Жукова. *Вестник Северного (Арктического) федерального университета. Серия: Естественные науки*. 2015. № 4. С. 119-130. DOI 10.17238/issn2227-6572.2015.4.119. EDN VRWNZV.
19. Зайцев, А. П. Автоматизация расчетов показателей защищенности защищаемого помещения от утечки информации / А. П. Зайцев, Д. А. Толстунов. *Доклады Томского государственного университета систем управления и радиоэлектроники*. 2010. № 1-1(21). С. 74-77. EDN MPWCRB.
20. A review of audio fingerprinting / Cano P. et [al.]. *Journal of VLSI signal processing systems for signal, image and video technology*. 2005. V. 41. № 3. P. 271-284. DOI:10.1007/s11265-005-4151-3.
21. Веялкин, И. А. Обзор структур алгоритмов поиска на основе отпечатков аудиоданных / И. А. Веялкин, Д. В. Шиман. *Труды БГТУ. №6. Физико-математические науки и информатика*. 2015. № 6(179). С. 195-199. EDN XCSNLH.

References

1. Ashikhmin, A. V. Avtomatizirovannaiia sistema radiomonitoringa / A. V. Ashikhmin, V. A. Kozmin, A. M. Rembovskii. *Spetstekhnika i sviaz*. 2012. № 1. S. 43-50. EDN OXUPNJ.
2. Solovev, A. M. O rezultatakh razrabotki avtomatizirovannoi sistemy nauchnykh issledovaniy sredstv monitoringa radioelektronnoi obstanovki v ASUTP / A. M. Solovev, A. A. Golovin, R. K. Kindenov. *Tekhnicheskie nauki - ot teorii k praktike*. 2015. № 53. S. 80-88. EDN VDWRPV.
3. Rushechnikov, Ia. I. Informatsionnaia tekhnologiia radiomonitoringa na osnove programmno-opredeliaemoi radiosistemy / Ia. I. Rushechnikov, V. V. Danilov. *Vestnik Donetskogo natsionalnogo universiteta. Serii G: Tekhnicheskie nauki*. 2020. № 1. S. 31-36. EDN PXYTYX.
4. Tretiakov, I. A. Funktsionalnye bloki dlia ANSI radiosignalov / I. A. Tretiakov, V. V. Danilov. *Donetskie chteniia 2024: obrazovanie, nauka, innovatsii, kultura i vyzovy sovremennosti: materialy IX Mezhdunarodnoi nauchnoi konferentsii (Donetsk, 15–17 oktiabria 2024 g.)*. Т. 2. Ch. 2. Donetsk: Izd-vo DonGU, 2024. S. 88-91. EDN ZEFZIV.
5. Elementy avtomatizirovannykh sistem nauchnykh issledovaniy radioelektronnoi obstanovki / I. A. Tretiakov, M. V. Babicheva, Ia. I. Rushechnikov, V. V. Danilov. *Problemy iskusstvennogo intellekta*. 2025. № 3(38). S. 172-181. DOI 10.24412/2413-7383-2025-3-38-172-181. EDN HKDOZN.
6. Khorev, A. A. Otsenka vozmozhnosti obnaruzheniia pobochnykh elektromagnitnykh izlucheniia videosistemy kompiuter / A. A. Khorev. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniia i radioelektroniki*. 2014. № 2(32). S. 207-213. EDN SEBGWX.
7. Zhelezniak, V. K. Metodika otsenki zashchishchennosti videoinformatsii ShIM-preobrazovatel'ia sredstv vychislitelnoi tekhniki / V. K. Zhelezniak, S. V. Kharchenko. *Vestnik Polotskogo gosudarstvennogo universiteta. Serii S. Fundamentalnye nauki*. 2020. № 4. S. 23-30. EDN JRRBRM.
8. Shpilevoi, A. A. O vozmozhnosti analiticheskogo obnaruzheniia signala PEMIN v videointerfeisakh standarta HDMI / A. A. Shpilevoi, A. A. Persichkin. *Vestnik Baltiiskogo federalnogo universiteta im. I. Kanta. Serii: Fiz.-mat. i tekhn. nauki*. 2022. № 1. S. 34-39. EDN AWTJYC.
9. Metodika avtomatizirovannoi otsenki obieema pobochnykh elektromagnitnykh izlucheniia / Ia. I. Rushechnikov, V. V. Danilov, I. A. Tretiakov, V. A. Stupak. *Vestnik Donetskogo natsionalnogo universiteta. Serii G: Tekhnicheskie nauki*. 2024. № 2. S. 51-57. DOI 10.5281/zenodo.12548548. EDN UHOZKU.
10. Svidetelstvo o gosudarstvennoi registratsii programmy dlia EVM № 2025688246 Rossiiskaia Federatsiia. Programma dlia opredeleniia kachestvennykh kharakteristik pobochnykh elektromagnitnykh izlucheniia i navodok elementov i ustroystv ofisnoi vychislitelnoi tekhniki : zaiavl. 26.05.2025 : opubl. 17.10.2025 / Ia.

- I. Rushechnikov, I. A. Tretiakov ; zaiavitel Federalnoe gosudarstvennoe biudzhethnoe obrazovatelnoe uchrezhdenie vysshego obrazovaniia «Donetskii gosudarstvennyi universitet». EDN HSARZF.
11. Kuleshov, I. A. Analiz sposobov nekriptograficheskoi zashchity informatsii ot pobochnykh elektromagnitnykh izlucheniï i navodok na obiektakh svyazi / I. A. Kuleshov, A. I. Spivak, A. E. Lvov. *Tekhnika sredstv svyazi*. 2021. № 3(155). S. 67-73. EDN SXQOTD.
 12. Parshutkin, A. V. Povyshenie zashchishchennosti informatsii ot utechki cherez pobochnye elektromagnitnye izlucheniia / A. V. Parshutkin, M. R. Neaskina. *Voprosy kiberbezopasnosti*. 2022. № 3(49). S. 82-89. DOI 10.21681/2311-3456-2022-3-82-89. EDN ARQCJO.
 13. Iazov, Iu. K. Puti postroeniia mnogoagentnoi sistemy zashchity informatsii ot utechki po tekhnicheskim kanalám / Iu. K. Iazov, A. O. Avsentev. *Voprosy kiberbezopasnosti*. 2022. № 5(51). S. 2-13. DOI 10.21681/2311-3456-2022-5-2-13. EDN DGYOEF.
 14. Parshutkin, A. V. Eksperimentalnye issledovaniia vozmozhnosti primeneniia programmno-realizuemykh mer snizheniia informativnosti pobochnykh elektromagnitnykh izlucheniï videosistemy standarta DVI / A. V. Parshutkin, M. R. Neaskina, I. P. Stepanov. *SPbNTORES: trudy ezhegodnoi NTK*. 2023. № 1(78). S. 126-129. EDN LAZWZX.
 15. Iazov, Iu. K. Problemnye voprosy upravleniia zashchitoi informatsii ot utechki po tekhnicheskim kanalám s primeneniem mnogoagentnykh sistem / Iu. K. Iazov, A. O. Avsentev. *Voprosy kiberbezopasnosti*. 2024. № 6(64). S. 85-97. DOI 10.21681/2311-3456-2024-6-85-97. EDN DDTRVS.
 16. Algoritm generatsii vysokochastotnoi ShIM-pomekhi / Ia. I. Rushechnikov, I. A. Tretiakov, V. A. Stupak, D. R. Zelenchenko. *69-ia Mezhdunarodnaia nauchnaia konferentsiia Astrakhanskogo gosudarstvennogo tekhnicheskogo universiteta (Astrakhan, 26–30 maia 2025 g.): materialy*. Astrakhan: Izd-vo AGTU, 2025. S. 188-190. EDN VUWJJV.
 17. Tretiakov, I. A. Podavlenie PEMIN v videointerfeisakh posredstvom vysokochastotnoi ShIM pomekhi / I. A. Tretiakov, Ia. I. Rushechnikov, V. V. Danilov. *Informatsionnye sistemy i tekhnologii: materialy XI mezhdunarodnogo nauchnogo kongressa po informatike. V 2 ch. (Minsk, 29–31 oktiabria 2025 g.)*. Ch. 1. Minsk: BGU, 2025. S. 111-114.
 18. Bederdinova, O. I. Otsnivanie zashchishchennosti konfidentsialnoi informatsii ot utechki po tekhnicheskim kanalám / O. I. Bederdinova, I. V. Zhukova. *Vestnik Severnogo (Arkticheskogo) federalnogo universiteta. Seriya: Estestvennye nauki*. 2015. № 4. S. 119-130. DOI 10.17238/issn2227-6572.2015.4.119. EDN VRWNZV.
 19. Zaitsev, A. P. Avtomatizatsiia raschetov pokazatelei zashchishchennosti zashchishchaemogo pomeshcheniia ot utechki informatsii / A. P. Zaitsev, D. A. Tolstunov. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniia i radioelektroniki*. 2010. № 1-1(21). S. 74-77. EDN MPWCRB.
 20. A review of audio fingerprinting / Cano P. et [al.]. *Journal of VLSI signal processing systems for signal, image and video technology*. 2005. V. 41. №. 3. P. 271-284. DOI:10.1007/s11265-005-4151-3.
 21. Veialkin, I. A. Obzor struktur algoritmov poiska na osnove otpechatkov audiodannykh / I. A. Veialkin, D. V. Shiman. *Trudy BGTU. №6. Fiziko-matematicheskie nauki i informatika*. 2015. № 6(179). S. 195-199. EDN XCSNLH.

RESUME

Ya. I. Rushechnikov, I. A. Tretiakov., V. V. Danilov

Adaptive automated system for suppression of electromagnetic side effects and interference

The development of complex, accessible, automated systems for suppressing signals from unintended information leakage channels is a sought-after scientific and technical task, the results of which can be implemented into systems designed for the computerized processing of confidential information. Due to the multitude and complexity of calculations, the solution of this problem is impossible without the use of intelligent identification tools and the creation of specialized automated systems, such as automated systems for scientific research of the electronic environment. This paper describes a comprehensive technique for searching for and subsequent suppression of detected electromagnetic side effects and interference signals within a single feedback suppression software and hardware package. The proposed technique will make it possible to use the most accessible software-defined platforms for scanning radio settings, which will reduce the final cost of an adaptive automated system for suppressing such signals.

РЕЗЮМЕ

Я. И. Рушечников, И. А. Третьяков, В. В. Данилов
Адаптивная автоматизированная система подавления побочных электромагнитных излучений и наводок

Разработка комплексных, доступных, автоматизированных систем подавления сигналов непреднамеренных каналов утечки информации, является востребованной научно-технической задачей, результаты решения которой могут быть внедрены в системы, предназначенные для компьютеризированной обработки конфиденциальной информации. Из-за множества и сложности вычислений решение этой задачи невозможно без использования интеллектуальных инструментов идентификации и создания специализированных автоматизированных систем, таких как автоматизированные системы научных исследований радиоэлектронной обстановки. В настоящей работе описана комплексная методика поиска и последующего подавления обнаруженных сигналов побочных электромагнитных излучений и наводок в рамках единого программно-аппаратного комплекса подавления с использованием обратной связи. Предложенная методика позволит использовать самые доступные программно-определяемые платформы для сканирования радиообстановки, что снизит конечную стоимость адаптивной автоматизированной системы подавления таких сигналов.

Рушечников Ярослав Иванович – старший преподаватель, кафедра радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», г. Донецк, Российская Федерация. *Область научных интересов:* автоматизированные системы радиомониторинга, нейронные сети, методы и системы защиты информации, информационная безопасность. Эл. почта ya.rushechnikov@donnu.ru, адрес 283001, РФ, г. Донецк, ул. Университетская, 24.

Третьяков Игорь Александрович – кандидат технических наук, доцент, заведующий кафедрой радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», г. Донецк, Российская Федерация. *Область научных интересов:* автоматизация научных исследований и автоматизированные системы; оптические информационные технологии; методы и системы защиты информации, информационная безопасность. Эл. почта i.tretiakov@mail.ru, адрес 283001, РФ, г. Донецк, ул. Университетская, 24.

Данилов Владимир Васильевич – доктор технических наук, профессор, профессор кафедры радиофизики и инфокоммуникационных технологий ФГБОУ ВО «Донецкий государственный университет», г. Донецк, Российская Федерация. *Область научных интересов:* оптические информационные технологии; автоматизация научных исследований и автоматизированные системы; информационно-измерительные системы. Эл. почта ut5iv@mail.ru, адрес 283001, РФ, г. Донецк, ул. Университетская, 24.

Статья поступила в редакцию 06.10.2025.

Раздел 5
Section 5

Роботы, мехатроника.
Робототехнические системы

Robots, mechatronics.
Robotic systems

М. Л. Абдулвадудов¹, К.Л. Вахидова², С.М. Ногамирзаев²

¹ Научно-производственный центр «Электронные вычислительно-информационные системы», Россия, г. Зеленоград

² Федеральное государственное бюджетное образовательное учреждение высшего образования "Грозненский государственный нефтяной технический университет имени академика М.Д. Миллионщикова"

364051, Чеченская Республика, г. Грозный, пр-т им. Х.А. Исаева, 100

ПРИМЕНЕНИЕ МЕТОДОВ ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ НЕЙРОСЕТЕВОГО АНАЛИЗА ГЕТЕРОГЕННЫХ СИСТЕМ НА КРИСТАЛЛЕ В ЗАДАЧАХ РОБОТОТЕХНИКИ

M. L. Abdulvadudov¹, K. L. Vakhidova², S. M. Nogamirzaev²

¹ Scientific and Production Center "Electronic Computing and Information Systems", Russia, Zelenograd

² Federal State Budgetary Educational Institution of Higher Education "Grozny State Petroleum Technological University named after Academician M.D. Millionshchikov"

364051, Chechen Republic, Grozny, Ave. named after Kh.A. Isaev, 100

APPLICATION OF METHODS FOR INCREASING THE EFFICIENCY OF NEURAL NETWORK ANALYSIS OF HETEROGENEOUS SYSTEMS-ON-CHIP IN ROBOTICS TASKS

В статье представлено исследование методов повышения эффективности выполнения нейросетевых алгоритмов на гетерогенных системах на кристалле (СнК) RoboDeus, ориентированные на применение в робототехнических системах. Приводится анализ архитектуры СнК RoboDeus, включая DSP-ядро Elcore50, и демонстрируется оптимизация работы сверточных нейронных сетей на примере детектора объектов YOLO V5s. Предложенные подходы к распараллеливанию вычислений, эффективному использованию памяти и аккумуляторных регистров позволяют достичь значительного прироста производительности при выполнении задач компьютерного зрения в реальном времени. Результаты экспериментов показывают, что использование многоядерной обработки на базе RoboDeus обеспечивает частоту кадров до 33.3 FPS для задачи детектирования объектов на изображениях разрешения 640×640. Обсуждаются перспективы интеграции оптимизированных нейросетевых алгоритмов в автономные робототехнические комплексы для навигации, распознавания окружения и взаимодействия с объектами

Ключевые слова: гетерогенные вычислительные системы, RoboDeus, YOLO V5s, сверточные нейронные сети, DSP-ядро, архитектура набора инструкций (АНК), система на кристалле (СнК).

The aim of this work is to study methods for improving the efficiency of neural network algorithms on heterogeneous RoboDeus systems-on-a-chip (SoCs), aimed at application in robotic systems. An analysis of the RoboDeus SoC architecture, including the Elcore50 DSP core, is presented, and optimization of convolutional neural networks is demonstrated using the YOLO V5s object detector as an example. The proposed approaches to parallelizing computations, efficient use of memory, and accumulator registers enable significant performance gains when executing computer vision tasks in real time. Experimental results show that the use of multi-core processing based on RoboDeus provides a frame rate of up to 33.3 FPS for the task of detecting objects in 640×640 images. Prospects for integrating optimized neural network algorithms into autonomous robotic systems for navigation, environment recognition, and interaction with objects are discussed.

Keywords: heterogeneous computing systems, RoboDeus, YOLO V5s, convolutional neural networks, DSP core, instruction set architecture (ISA), system on a chip (SoC).

Современная робототехника предъявляет высокие требования к вычислительным платформам: необходимость обработки больших объёмов сенсорных данных (изображений, лидаров, аудио) в реальном времени, энергоэффективность, компактность и надёжность. Ключевую роль в таких системах играют алгоритмы искусственного интеллекта, в частности, свёрточные нейронные сети (CNN), используемые для распознавания образов, детектирования объектов, сегментации сцены. Однако их выполнение на традиционных CPU часто оказывается недостаточно производительным, что стимулирует разработку специализированных гетерогенных архитектур, сочетающих различные типы процессорных ядер.

Искусственный интеллект возник как способ имитации естественного интеллекта, базирующегося на нейронных сетях человеческого мозга и ставшего основой человеческой культуры, науки и искусства. Нейрокибернетика, возникшая во второй половине сороковых годов прошлого века, дала начало исследованиям по технической имитации решения интеллектуальных задач человеком. В шестидесятые годы эта волна исследований уступила место имитации элементов человеческого мышления на основе формальной логики. В начале восьмидесятых (после работ Дж. Хопфилда), в связи с появлением алгоритмов, которые позволили обучать «скрытые» (не связанные со входами и выходами нейросети) слои нейронов, возникла вторая волна интереса к нейронным сетям. Эти алгоритмы значительно расширили возможности нейронных сетей, породив нейроинформатику – науку об обработке информации и управлении с помощью нейронных сетей.

Несмотря на то, что нейронные сети как научно-технический объект существуют более полувека, основные способы их практической реализации ограничиваются областью программного моделирования в различных средах визуального (и не визуального) проектирования. Программные реализации преобладают над аппаратными. Кроме того, аппаратные реализации вследствие своей относительной дороговизны до сих пор не имеют повсеместного распространения. Однако их удельный вес в общем количестве мировых технологических разработок неуклонно растёт. Впереди планеты всей в данном направлении, как всегда, шагает Япония, где разработки на основе нейросетевых технологий успешно внедряются во множестве образцов бытовой техники, таких как фотоаппараты, микроволновые печи, видеокамеры и т.д. Не отстают и другие развитые страны [22].

Нейросетевые аппаратные решения внедряют в свои продукты такие известные фирмы, как Siemens, Intel, Phillips Research (Нидерланды), 3M Laboratories и многие другие. При этом разработано на удивление много разнообразных аппаратных нейросетевых архитектур. Отечественные фирмы, такие как МЦСТ, НТЦ Модуль и НПЦ ЭЛВИС [1] предлагают свои решения.

Цель работы является повышение эффективности программ нейросетевого анализа для гетерогенных вычислительных систем на основе отечественного процессора 1892BM248 RoboDeus путем внедрения ускоренных алгоритмов нейронных сетей. Для достижения цели в работе необходимо решить следующие задачи: исследовать особенности архитектуры гетерогенных СнК; предложить способы повышения эффективности нейросетевого анализа на гетерогенных СнК; разработка методики реализации нейросетевого алгоритма на гетерогенной СнК; результаты применения методики повышения эффективности нейронной сети YOLO V5s для распознавания и детектирования объектов.

Для решения поставленных задач будут рассмотрены архитектура нейронной сети YOLO V5s и схема вычислителя СнК Robodeus, а также будет составлена схема принципа работы свёрточного нейронного слоя в DSP-ядре СнК RoboDeus.

Гетерогенная топология ЦП

Обычно неоднородность в контексте вычислений упоминается к разным архитектурам набора инструкций (далее АНК), где основной процессор имеет одну, а другие процессоры имеют другую архитектуру (возможно более одной), а не просто другую микроархитектуру (обработка чисел с плавающей точкой — это особый случай, такой процессор не называют гетерогенным).

В прошлом гетерогенные вычисления означали, что разные АНК должны обрабатываться по-разному, в то время как в современном примере системы с гетерогенной системной архитектурой устраняют разницу (для пользователя) при использовании нескольких типов процессоров (обычно CPU и GPU). На одной и той же интегральной схеме, чтобы обеспечить лучшее из 2-х главных задач: общую обработку графического процессора (помимо хорошо известных возможностей графического рендеринга 3D-графики, он также может выполнять математически интенсивные вычисления с очень большими наборами данных), в то время как другие процессоры могут работать с операционной системой и выполнять традиционные последовательные задачи. Недавние результаты показывают [2], что многоядерный процессор с гетерогенной АНК, который использует разнообразие, предлагаемое несколькими АНК, может превзойти лучшую однородную архитектуру с той же АНК на целых 21% с экономией энергии на 23% и сокращением продукта задержки энергии (EDP) на 32%. Пример гетерогенного СнК является RoboDeus, который включает в себя CPU, GPU и DSP.

Архитектура DSP - ядер СнК RoboDeus

В состав DSP-ядра Elcore50 входят следующие основные блоки: процессорное ядро DSP; устройство управления памятью VMMU; L1 кэш инструкций L1 I\$ объемом 4 КВ; L1 кэш данных L1 D\$ объемом 16 КВ; память программ PRAM объемом 64 КВ; память данных XYRAM объемом 512 КВ, которая может быть частично или полностью реконфигурирована в L2 кэш L2\$; контроллер прямого доступа к внешней памяти DMA; коммутатор шин данных MUX; устройство вывода трассы Trace; контроллеры внешних интерфейсов.

DSP-ядро Elcore50 работает в виртуальном адресном пространстве. Трансляция виртуальных адресов в физические выполняется с помощью устройства управления памятью VMMU. В составе DSP-ядра имеется внутренняя статическая память - память программ PRAM и объединенная память данных и программ XYRAM. Обращения к памяти программ выполняются по 128-разрядной шине P, а к памяти данных – по двум 512-разрядным шинам X и Y (в соответствии с рисунком 1).

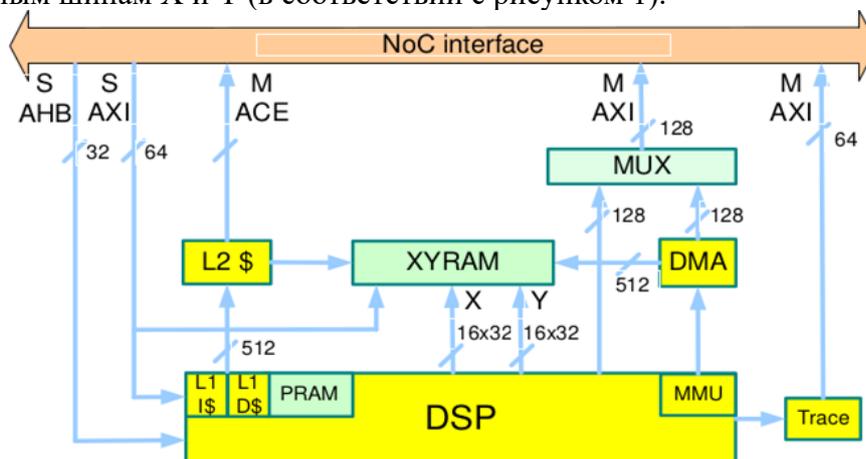


Рисунок 1 – структурная схема DSP – ядра Elcore50

Большинство инструкций выполняются над векторными регистрами в пределах вычислительной секции. Вычислительная секция имеет ширину обработки 64 бита. Данные обрабатываются как упакованный SIMD. В Elcore50 предусмотрено 8 одинаковых вычислительных SIMD секций (в соответствии с рисунком 2). Векторный вычислитель (сопроцессор) EVX оперирует 512-разрядными векторами данных.

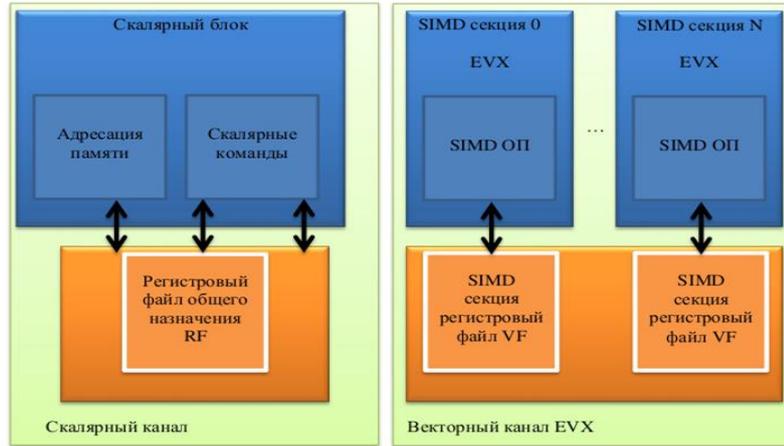


Рисунок 2 – Блоки и регистры одноступенчатого детектора YOLO v5

Архитектура модели YOLO V5s

Поскольку YOLO v5 является одноступенчатым детектором объектов, он состоит из трех важных частей, как и любой другой одноступенчатый детектор объектов: Model Backbone, Model Neck, Model Head (в соответствии с рисунком 3). Модель Backbone в основном используется для извлечения важных функций из заданного входного изображения. В YOLO V5s CSP — частичные сети с перекрестными этапами используются в качестве основы для извлечения богатых информативных функций из входного изображения. Нейросетевой алгоритм CSPNet продемонстрировал значительное улучшение времени обработки в более глубоких сетях. Модель Neck в основном используется для создания пирамид признаков. Пирамиды признаков помогают моделям хорошо обобщаться при масштабировании объектов. Это помогает идентифицировать один и тот же объект с разными размерами и масштабами. Пирамиды функций очень полезны и помогают моделям хорошо работать со скрытыми слоями.

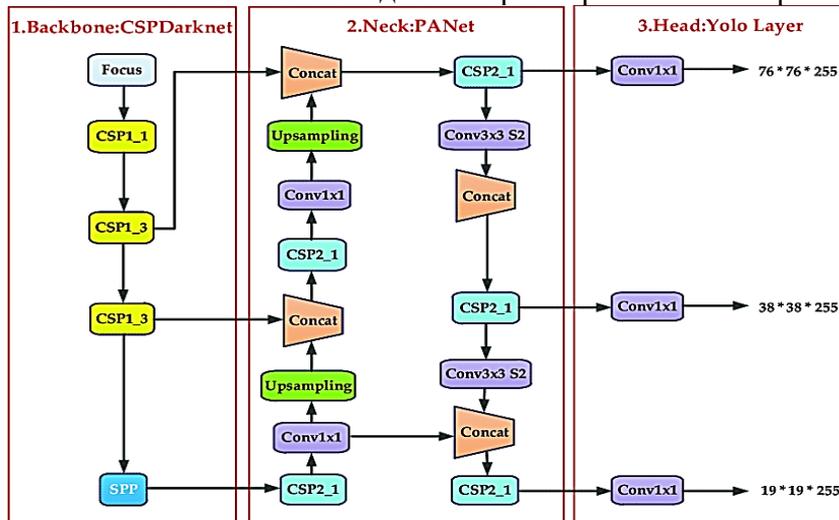


Рисунок 3 – Архитектура нейронной сети YOLO V5s

Модель Head в основном используется для выполнения заключительной части обнаружения. Он применяет рамки привязки к функциям и генерирует окончательные выходные векторы с вероятностями классов, оценками объектности и ограничивающими рамками. В модели YOLO V5s Model Head такая же, как и в предыдущих версиях YOLO V3 и V4.

Выбор функций активации является наиболее важным в любой глубокой нейронной сети. Недавно было введено множество функций активации, таких как: Leaky ReLU, swish и т.д. В YOLO V5s функция активации Leaky ReLU используется в средних/скрытых слоях, а сигмовидная функция активации используется в последнем слое обнаружения.

Алгоритм сверточной нейронной сети

Свёрточные нейронные сети (СНС) составляют вычислительное ядро большинства современных систем машинного зрения, являясь ключевым инструментом для решения задач распознавания образов, детектирования объектов и семантической сегментации в реальном времени. В робототехнике, где требуется высокая скорость реакции на изменяющееся окружение, энергоэффективность и надёжность алгоритмов, оптимизация выполнения операций свёртки приобретает особую значимость. Однако традиционные подходы к реализации СНС на универсальных процессорах зачастую не позволяют достичь требуемой производительности при ограниченных аппаратных ресурсах, что стимулирует активный поиск специализированных методов вычислений, адаптированных под архитектуру гетерогенных систем на кристалле.

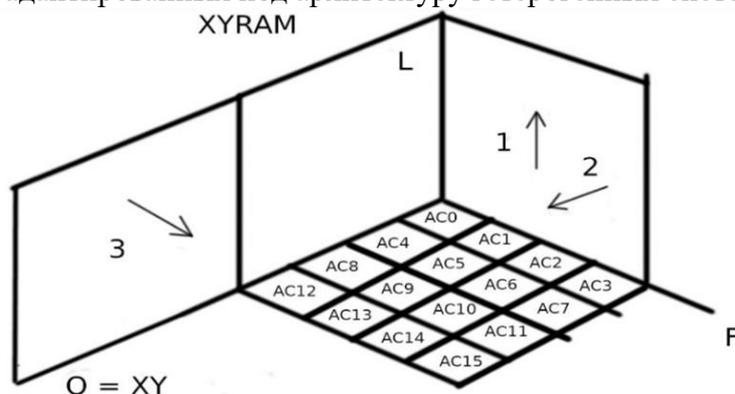


Рисунок 4 – 3-х мерное представление вычислений свертки и расположения аккумуляторов

В соответствии с рисунком 4 изображено 3-х мерное представление сверточного слоя. Пространство LF — данные весов нейронного слоя; пространство QL — входные данные, где $Q = X*Y$; пространство QF — выходные данные. В DSP – ядре SnK RoboDeus присутствуют 16 аккумуляторов, которые накапливают с каждой итерацией скалярно-векторное произведение матриц. Преимущество этой схемы в том, что при расположении аккумуляторов в пространстве QF в виде 4x4, получается оптимальный параллельный код вычислений и загрузок данных. Т.е. весь код загрузок работает параллельно с инструкциями вычислений с помощью VLIW пакетов. Данные в аккумуляторах локализуются по оси QF. Скалярно-векторные операции с накоплением в аккумуляторах проходят первоначально в сторону оси L, затем локализуются данные по оси LF в памяти XYRAM. Со следующей итерацией в пространстве QF в сторону Q можно работать с новыми выходными данными (2-е направление из рисунка) эффективно, т.к. данные весов уже хранятся в памяти XYRAM и времени на загрузку в

память XYRAM не потребуется. И наконец локализируются данные по оси QL в L3 кэше, и работа со следующими данными в пространстве QF в сторону F выполняется эффективно, т.к. входные данные в пространстве QL уже хранятся в L3 кэше и времени на загрузку в L3 кэш не потребуется.

Таким образом для повышения эффективности работы нейросетевых алгоритмов на СнК RoboDeus, а именно на DSP-ядрах ELscore применены следующие подходы: максимальное использование аккумуляторных регистров, оптимальное расположение аккумуляторов (в данном случае по схеме 4x4) для объединения кода загрузок с вычислительными инструкциями в VLIW пакеты, размещение данных фильтров в память XYRAM для повторного использования в итерациях вычислений по направлению Q, размещение входных данных в L3 кэш для повторного использования в итерациях вычислений по направлению F, они позволили значительно повысить производительность алгоритмов на аппаратной платформе. Для улучшения эффективности можно предложить также в начале вычислений инициализировать аккумуляторы нейронами смещения (biases), а в конце вычислений выполнить функцию активации для аккумуляторов.

Результаты вычислений нейронной сети YOLO V5s

Custom Dataset, 4 класса:

Ошибка: диапазон значений (0 - 1): 0.01

1 ядро, типы данных — half(fp16), размеры картинок = 640x640x3: FPS = 3.07

16 ядер, типы данных — half(fp16), размеры картинок = 640x640x3: FPS = 33.3

Coco Dataset, 80 классов

Ошибка: диапазон значений (0 - 1): 0.0016

1 ядро, типы данных — half(fp16), размеры картинок = 640x640x3: FPS = 2.16

16 ядер, типы данных — half(fp16), размеры картинок = 640x640x3: FPS = 21.7

Таблица – Производительность нейросетевых алгоритмов на СнК RoboDeus

Задача	Тип нейросети	1892BM248 «RoboDeus», fps		Jetson AGX Xavier, fps	Jetson Nano, fps
		single stream	пакетная обработка		
Классификация	ResNet 18	31,1	432,5	743 (fp16)/338(fp32)	90,4(fp16)/53
	ResNet 34	21,3	276,6	419(fp16)/191,5(fp32)	50,4(fp16)/28,5
	ResNet 50	17,3	204,5	325(fp16)/115	33,5(fp16)/18,4
	MobileNetv1	43,26	583,2	513 (fp32)	72 (fp32)
Обнаружение	SSD-ResNet34	1,15	11,68	149(fp16)/51,47(fp32)	12(fp16)
	SSD-MobileNetv2	1,28	12,03	232,3/118,2	

Заключение

Стоит отметить, что использование гетерогенных СнК, таких как RoboDeus, в сочетании с оптимизированными нейросетевыми алгоритмами позволяет достичь высокой производительности в задачах компьютерного зрения, актуальных для робототехники. Ключевыми преимуществами являются энергоэффективность, возможность параллельной обработки нескольких потоков данных и адаптируемость архитектуры под различные типы нейронных сетей. Перспективным исследованием является использование режима single-stream обработки, когда несколько DSP-ядер работают над

одним кадром, что может дополнительно повысить эффективность для задач с высоким разрешением. Также перспективным направлением является разработка фреймворка для обучения нейронных сетей непосредственно на целевом аппаратном обеспечении, что позволит сократить цикл разработки робототехнических систем с интеллектуальными функциями.

В работе были представлены архитектура нейронной сети YOLO V5s, архитектура гетерогенного СнК RoboDeus, пример оптимизации сверточной нейронной сети на DSP-ядре СнК RoboDeus. Нейронная сеть YOLO V5s была оптимизирована в режиме multi stream, т.е. каждое DSP - ядро обрабатывало по одному фрейму. Такие внутренние нейронные слои в YOLO V5s, как транспонирование матрицы, умножение и суммирование тензора на тензор стали эффективными с помощью векторных инструкций и приемами программирования кэша XYRAM. В формуле sigmoid экспонента была представлена в виде $e^x = e^{k \ln 2 + r} = e^{k \ln 2} e^r = 2^k e^r$, где $k = \text{round}\left(\frac{x}{\ln 2}\right) = \left\lfloor \frac{x}{\ln 2} + 0.5 \right\rfloor$, $|r| \leq 0.5 \ln 2$ и $r = x - k \ln 2$. Число k было найдено с помощью ассемблерных команд DSP-ядра СнК RoboDeus, а число r было найдено с помощью аппроксимации Лагранжа.

В дальнейшем планируется оптимизация нейронных сетей в режиме single stream, т.е. каждое DSP - ядро будет обрабатывать один и тот же frame по кусочкам. С представленным опытом эффективности работы нейронных слоев можно применить в таких нейронных сетях, которые применяются в NLP, т.е. распознавание речи и обработка текстовых документов, а также в разработке фреймворка обучения нейронных сетей.

Список литературы

1. АО НПЦ «ЭЛВИС». Электронный ресурс: <https://elvees.ru> (дата обращения: 24.10.2025).
2. Галушкин А. И. Теория нейронных сетей: учеб. пособие для вузов / под общ. ред. А. И. Галушкина. М.: ИПРЖР, 2000. Кн. 1. 416 с.
3. Осовский С. Нейронные сети для обработки информации. // Пер. с польского И.Д. Рудинского. - М.: Финансы и статистика, 2002. 344 с.
4. Ростовцев В. С. Искусственные нейронные сети [Электронный ресурс]: учеб. для вузов. 2-е изд., стер. СПб: Лань, 2021. 216 с. Режим доступа: <https://e.lanbook.com/book/160142> (дата обращения: 20.08.2022).
5. Цуриков А. Н. Моделирование и обучение искусственных нейронных сетей [Электронный ресурс]: учеб. пособие. Ростов-н/Д.: РГУПС, 112 с. Режим доступа: <https://e.lanbook.com/book/140610> (дата обращения: 20.08.2022).
6. A Heterogeneous Multi-Core System-on-Chip for Energy Efficient Brain Inspired Computing / Antonio Pullini, Francesco Conti, David Rossi, Michael Gautschi / ResearchGate, 2017. – 6 p.
7. Towards AI. Официальный сайт. Электронный ресурс: <https://towardsai.net> (дата обращения: 24.10.2022).
8. Kunzman, D.M. (2011). Programming Heterogeneous Systems. International Symposium on Parallel and Distributed Processing Workshops. doi:10.1109/IPDPS.2011.377
9. Федоров, М. В. (член-корр. РАН) Новые методы машинного обучения: использование фундаментального знания для решения практических задач: [тезисный онлайн-доклад] [Электронный ресурс] // Программа II Всероссийской школы-семинара по искусственному интеллекту и большим данным в технических, промышленных, природных и социальных системах. – Национальный центр физики и математики (НЦФМ). Дата проведения: 25-29 ноября 2024 г. (дата обращения: 04.04.2025).
10. Чудина, О.В. Теория и практика термической обработки металлов: учебно-методическое пособие к мультимедийному изданию / О.В. Чудина, Г.В. Гладова, А.В. Остроух. М.: МАДИ, 2013. 64 с.
11. Сергеев, А. П. Введение в нейросетевое моделирование: учеб. пособие / А. П. Сергеев, Д. А. Тарасов ; под общ. ред. А. П. Сергеева. Екатеринбург : Изд-во Урал. ун-та, 2017. 128 с. ISBN 978-5-7996-2124-7
12. Хливненко Л. В. Практика нейросетевого моделирования : учебное пособие для вузов / Л. В. Хливненко, Ф. А. Пятакович. 2е изд., стер. Санкт-Петербург : Лань, 2021. 200 с.: ил. Текст : непосредственный. ISBN 978.5.8114.8264.1
13. Вичугов В.Н Модифицированный градиентный алгоритм обучения радиально-базисных нейронных сетей. Известия Томского политехнического университета. Управление, вычислительная техника и информатика. 2009. Т. 315. № 5 с150-152
14. Сидоров Д.И., Новикова О.И. Применение методов машинного обучения для предиктивной диагностики оборудования. Проблемы искусственного интеллекта, 2023, №1(28), С.23-34.

15. Козлов С.В., Фролова М.П. Оптимизация энергозатрат промышленных установок с использованием нейронных сетей. Проблемы искусственного интеллекта, 2021, №4(27), С.67-78.
16. Белов И.А., Орлова Т.К. Адаптивные системы автоматического регулирования с использованием гибридных алгоритмов. Проблемы искусственного интеллекта, 2022, №2(29), С.18-29.
17. Никитина Л.В., Павлов В.Г. Моделирование и анализ сложных технических систем в условиях неопределенности. Проблемы искусственного интеллекта, 2023, №4(31), С.78-90.
18. Федоров К.А., Семенова А.А. Цифровые двойники как основа повышения эффективности производственных линий. Проблемы искусственного интеллекта, 2021, №1(24), С.55-66.
19. Григорьев П.С., Медведева Е.В. Интеллектуальный анализ данных в задачах мониторинга технологических параметров. Проблемы искусственного интеллекта, 2022, №4(31), С.34-45.
20. Титов В.М., Ветрова Н.С. Разработка алгоритмов управления на основе обучения с подкреплением для промышленных роботов. Проблемы искусственного интеллекта, 2023, №2(29), С.61-72.
21. Орлова С.К., Крылов Д.В. Методы повышения устойчивости систем автоматического управления в условиях возмущений. Проблемы искусственного интеллекта, 2021, №3(26), С.41-52.
22. Нурулин Ю. Р., Скворцова И. В., Нурулин М. Ю., Наумов А. С. Разработка человеко-машинных интерфейсов для систем спектрального анализа радиосигналов в реальном времени // Современные наукоемкие технологии. – 2025. – № 3. – С. 40-46. – DOI 10.17513/snt.40321.

References

1. JSC Scientific and Production Center "ELVEES." Electronic resource: <https://elvees.ru> (date accessed: 24.10.2025).
2. Galushkin, A. I. Neural Network Theory: a textbook for universities / edited by A. I. Galushkin. Moscow: IPZhR, 2000. Book 1. 416 p.
3. Osovsky, S. Neural Networks for Information Processing. // Translated from Polish by I.D. Rudinsky. Moscow: Finance and Statistics, 2002. 344 p.
4. Rostovtsev, V. S. Artificial Neural Networks [Electronic resource]: a textbook for universities. 2nd ed., reprinted. St. Petersburg: Lan, 2021. 216 p. – Access mode: <https://e.lanbook.com/book/160142> (date accessed: 20.08.2022).
5. Tsurikov A. N. Modeling and training of artificial neural networks [Electronic resource]: textbook. manual. Rostov-on-Don: RGUPS, 112 p. Access mode: <https://e.lanbook.com/book/140610> (date accessed: 20.08.2022).
6. A Heterogeneous Multi-Core System-on-Chip for Energy Efficient Brain Inspired Computing / Antonio Pullini, Francesco Conti, David Rossi, Michael Gautschi / ResearchGate, 2017. – 6 p.
7. Towards AI. Official website. Electronic resource: <https://towardsai.net> (date accessed: 24.10.2022).
8. Kunzman, D.M. (2011). Programming Heterogeneous Systems. International Symposium on Parallel and Distributed Processing Workshops. doi:10.1109/IPDPS.2011.377
9. Fedorov, M. V. (Corresponding Member of the Russian Academy of Sciences) New Methods of Machine Learning: Using Fundamental Knowledge to Solve Practical Problems: [online abstract report] [Electronic resource] / M. V. Fedorov // Program of the II All-Russian School-Seminar on Artificial Intelligence and Big Data in Technical, Industrial, Natural, and Social Systems. – National Center for Physics and Mathematics (NCPM). – Date: November 25-29, 2024. – (date of access: 04.04.2025).
10. Chudina, O.V. Theory and Practice of Heat Treatment of Metals: a teaching aid for a multimedia publication / O. V. Chudina, G. V. Gladova, A. V. Ostroukh. Moscow: MADi, 2013. 64 p.
11. Sergeev, A. P. Introduction to Neural Network Modeling: a tutorial / A. P. Sergeev, D. A. Tarasov; edited by A. P. Sergeev. Yekaterinburg: Ural University Publishing House, 2017. 128 p. ISBN 978-5-7996-2124-7
12. Khlivnenko, L. V. Practice of Neural Network Modeling: a tutorial for universities / L. V. Khlivnenko, F. A. Pyatakovich. 2nd ed., reprinted. Saint Petersburg: Lan, 2021. 200 p.: ill. Text: direct. ISBN 978.5.8114.8264.1
13. Vichugov V.N. Modified Gradient Algorithm for Training Radial Basis Neural Networks. Bulletin of Tomsk Polytechnic University. Management, Computer Science, and Information Technology. 2009. Vol. 315. No. 5, pp. 150-152
14. Sidorov D.I., Novikova O.I. Application of Machine Learning Methods for Predictive Equipment Diagnostics. Problems of Artificial Intelligence, 2023, No. 1(28), pp. 23-34.
15. Kozlov S.V., Frolova M.P. Optimization of Energy Consumption of Industrial Installations Using Neural Networks. Problems of Artificial Intelligence, 2021, No. 4 (27), pp. 67-78.
16. Belov I. A., Orlova T. K. Adaptive automatic control systems using hybrid algorithms. Problems of Artificial Intelligence, 2022, No. 2 (29), pp. 18-29.
17. Nikitina L. V., Pavlov V. G. Modeling and analysis of complex technical systems under uncertainty. Problems of Artificial Intelligence, 2023, No. 4 (31), pp. 78-90.
18. Fedorov K. A., Semenova A. A. Digital twins as a basis for increasing the efficiency of production lines. Problems of Artificial Intelligence, 2021, No. 1 (24), pp. 55-66.
19. Grigoriev P. S., Medvedeva E. V. Intelligent data analysis in process parameter monitoring tasks. Problems of Artificial Intelligence, 2022, No. 4 (31), pp. 34-45.
20. Titov V.M., Vetrova N.S. Development of control algorithms based on reinforcement learning for industrial robots. Problems of Artificial Intelligence, 2023, No. 2 (29), pp. 61-72.

21. Orlova S.K., Krylov D.V. Methods for increasing the stability of automatic control systems under disturbances. *Problems of Artificial Intelligence*, 2021, No. 3 (26), pp. 41-52.
22. Nurulin Yu. R., Skvortsova I. V., Nurulin M. Yu., Naumov A. S. Development of human-machine interfaces for real-time radio signal spectral analysis systems // *Modern High Technologies*. – 2025. – No. 3. – P. 40-46. – DOI 10.17513/snt.40321.

RESUME

M. L. Abdulvadudov, K. L. Vakhidova, S. M. Nogamirzaev

Application of methods for increasing the efficiency of neural network analysis of heterogeneous systems-on-chip in robotics tasks

Modern robotic systems place high demands on the performance and energy efficiency of computing platforms, especially when processing computer vision data in real time. Traditional processor architectures often fail to handle complex neural network algorithms, such as convolutional neural networks, limiting the navigation and interaction capabilities of autonomous robots. Therefore, a pressing issue is the development and optimization of methods for performing neural network computations on specialized energy-efficient heterogeneous SoCs combining various types of processor cores. Objective, Methodology, and Novelty

The objective of this study is to improve the efficiency of neural network algorithms on the domestically produced 1892BM248 "RoboDeus" heterogeneous system-on-chip (SoC) for robotics applications. The scientific novelty lies in the development and application of a set of optimization methods, including efficient parallelization of computations, optimal memory management, and the use of accumulator registers adapted to the architecture of the Elcore50 DSP core of the RoboDeus SoC.

The methodology of this study included an analysis of the architecture of the heterogeneous RoboDeus SoC and the Elcore50 DSP core, a study of convolutional neural network algorithms (using the YOLO V5s object detector as an example), and the development and practical implementation of optimization methods for this hardware. Experiments were conducted using various datasets to evaluate performance.

The optimizations resulted in a significant increase in performance. For the YOLO V5s neural network, running on 16 RoboDeus SoC cores with half-precision (fp16) data and a 640x640 image resolution, the processing frequency was 33.3 frames per second for the custom dataset and 21.7 frames per second for the COCO dataset. A comparative analysis with other platforms (NVIDIA Jetson AGX Xavier and Jetson Nano) demonstrated the solution's competitiveness in object classification and detection tasks.

Thus, applying the proposed optimization methods on the heterogeneous RoboDeus SoC enables high performance in computer vision tasks relevant to robotics. Key advantages include energy efficiency, parallel processing, and architectural adaptability. Promising areas for further research include single-stream optimization (distributing the processing of a single frame across multiple cores) and adapting the methods to neural network algorithms for natural language processing (NLP). The developed approaches facilitate the creation of high-performance and energy-efficient computing systems for autonomous robotic systems.

РЕЗЮМЕ

М. Л. Абдулвадулов, К.Л. Вахидова, С.М. Ногамирзаев

Применение методов повышения эффективности нейросетевого анализа гетерогенных систем на кристалле в задачах робототехники

Современные робототехнические системы предъявляют высокие требования к производительности и энергоэффективности вычислительных платформ, особенно при обработке данных компьютерного зрения в реальном времени. Традиционные процессорные архитектуры зачастую не справляются с выполнением сложных нейросетевых алгоритмов, таких как свёрточные нейронные сети, что ограничивает возможности

автономных роботов по навигации и взаимодействию с окружением. В связи с этим актуальной задачей является разработка и оптимизация методов выполнения нейросетевых вычислений на специализированных энергоэффективных гетерогенных системах на кристалле, сочетающих различные типы процессорных ядер.

Целью работы является повышение эффективности выполнения нейросетевых алгоритмов на отечественной гетерогенной системе на кристалле (СнК) 1892ВМ248 «RoboDeus» для задач робототехники. Научная новизна заключается в разработке и применении комплекса методов оптимизации, включающих эффективное распараллеливание вычислений, оптимальное управление памятью и использование аккумуляторных регистров, адаптированных под архитектуру DSP-ядра Elcore50 СнК RoboDeus.

Методология работы включала анализ архитектуры гетерогенного СнК RoboDeus и DSP-ядра Elcore50, исследование алгоритмов свёрточных нейронных сетей (на примере детектора объектов YOLO V5s), а также разработку и практическую реализацию методик оптимизации для данного аппаратного обеспечения. Эксперименты проводились с использованием различных наборов данных для оценки производительности.

В результате проведённых оптимизаций достигнуто значительное повышение производительности. Для нейронной сети YOLO V5s при работе на 16 ядрах СнК RoboDeus с данными в формате половинной точности (fp16) и разрешении изображений 640×640 частота обработки составила 33.3 кадра в секунду для кастомного набора данных и 21.7 кадра в секунду для набора данных COCO. Сравнительный анализ с другими платформами (NVIDIA Jetson AGX Xavier и Jetson Nano) показал конкурентоспособность решения в задачах классификации и обнаружения объектов.

Таким образом, применение предложенных методов оптимизации на гетерогенной СнК RoboDeus позволяет достичь высокой производительности в задачах компьютерного зрения, актуальных для робототехники. Ключевыми преимуществами являются энергоэффективность, возможность параллельной обработки и адаптируемость архитектуры. Перспективными направлениями дальнейших исследований являются оптимизация в режиме single-stream (распределение обработки одного кадра между несколькими ядрами), а также адаптация методик для нейросетевых алгоритмов в области обработки естественного языка (NLP). Разработанные подходы способствуют созданию высокопроизводительных и энергоэффективных вычислительных систем для автономных робототехнических комплексов.

Абдулвадуов Магомед Лечиевич, ведущий инженер-программист по встраиваемым системам, АО НПЦ «ЭЛВИС», эл. почта abdulvadudovmagomed@mail.ru тел 8(915)178-67-77 адрес: 124460, г Москва, Зеленоград, ул Конструктора Лукина, 14, стр. 14

Область научных интересов: искусственный интеллект, компьютерное зрение, встраиваемые системы

Вахидова Карина Лечиевна, доцент кафедры «Автоматизация технологических процессов и производств», Грозненский государственный нефтяной технический университет имени академика М.Д. Миллионщикова, 364051, Чеченская Республика, г. Грозный, пр-т им. Х.А. Исаева, 100, тел 8(938)894-22-72 *Область научных интересов:* моделирование технических систем

Ногамирзаев Сулиман Мовлаевич – старший преподаватель кафедры «Технологии машиностроения и транспортных процессов» Грозненский государственный нефтяной технический университет имени академика М.Д. Миллионщикова, эл. почта kafedra2020@mail.ru, адрес: 364051, Чеченская Республика, г. Грозный, пр-т им. Х.А. Исаева, 100, *Область научных интересов:* компьютерное моделирование систем и процессов

Статья поступила в редакцию 23.09.2025

УДК 004.93

DOI 10.24412/2413-7383-2025-4-39-173-181

В. М. Зуев

Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта», г. Донецк
283048, г. Донецк, ул. Артема, 118 б

ПОГРЕШНОСТЬ ПРОГНОЗИРОВАНИЯ ПОЛОЖЕНИЯ СЛЕДУЮЩЕГО КАДРА СЪЁМКИ

V.M. Zuev

Federal State Budgetary Scientific Institution «Institute of Artificial Intelligence Problems»
283048, Donetsk, Artem str., 118 b

ERROR IN PREDICTING THE POSITION OF THE NEXT FILM FRAME

V. M. Zuev

Budgetary Foederati Publica Institutione Scientific Institutum Intelligentiae Artificialis Problems,
Donetsk 118 b Artyom str., Donetsk, 283048

ERROR IN PRAEDICTIONE SITUS PROXIMAE IMAGINIS PELLICULAE

В статье рассматривается погрешность прогнозирования следующего кадра съёмки, основанного на решении динамической и кинематической модели перемещения камеры, при условии, что камера жёстко закреплена на борту аппарата. Полученные математические зависимости погрешности от различных условий, в отличие от часто использующихся для этой цели нейронных сетей, дают математические зависимости, из которых легко получить необходимые численные характеристики.

Ключевые слова: детекция, прогнозирование, погрешность.

This article examines the error in predicting the next frame based on a dynamic and kinematic model of camera movement, assuming the camera is rigidly mounted on board the vehicle. The resulting mathematical dependencies of the error on various conditions, unlike neural networks often used for this purpose, yield mathematical dependencies from which the necessary numerical characteristics can be easily derived.

Keywords: detection, forecasting, error

Hic articulus errorem in praedictione proximae imaginis examinat, fretus exemplo dinamico et cinematico motus camerae, assumendo cameram rigide in vehiculo affixam esse. Dependendiae mathematicae erroris a variis condicionibus resultantes, dissimiles retibus neuralibus saepe ad hunc finem adhibitibus, dependentias mathematicas gignunt, ex quibus notae numericae necessariae facile derivari possunt.

Affatus clavis: detectio, praedictio, error

Введение

Прогнозирование следующего кадра съемки позволяет использоваться в алгоритмах обнаружения и детекции объектов, в оптической одометрии, оптической навигации и других приложениях. Эффективность этих алгоритмов в значительной мере зависит от погрешности прогнозирования кадра съемки.

Ниже описан способ вычисления погрешности прогнозирования кадра съёмки, основанного на решении динамической модели камеры, при условии, что камера жёстко закреплена на борту аппаратуры. Из полученных прямых математических зависимостей легко получить необходимые численные характеристики погрешности при различных условиях. В отличие от часто используемого для этой цели анализа нейронных сетей [1], [2], которые дают параметры только конкретной реализации, полученные математические зависимости дают универсальный способ получения необходимых численных характеристик и даже дать рекомендации по выбору лучшей конструкции аппарата.

Общая постановка задачи. Пусть в момент времени t положение и ориентация камеры задаются [3], [4] вектором состояния $\mathbf{x}(t) \in \mathbb{R}^n$, где n – размерность модели (обычно $n=12$: 3 координаты положения, 3 скорости, 3 угла Эйлера или кватерниона, 3 угловые скорости).

Для простоты считаем, что камера жёстко закреплена на корпусе ЛА, поэтому её положение и ориентация в глобальной (инерциальной) системе координат полностью определяются $\mathbf{x}(t)$ и известной калибровкой [5], [6] (внешними параметрами камеры) $\mathbf{T}_{cb} \in SE(3)$ – преобразованием из системы координат корпуса в систему координат камеры. Случай, когда камера имеет поворотное устройство не трудно получить из предлагаемых выкладок.

Цель прогнозирования: по текущему состоянию $\mathbf{x}(t_0)$ и динамической модели $\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}, t)$ (где $\mathbf{u}(t)$ – управление) предсказать изображение сцены в момент $t_0 + \Delta t$. Это эквивалентно [7] прогнозированию проекции точек сцены на изображение, что зависит от положения камеры $\mathbf{P}(t_0 + \Delta t) \in SE(3)$.

Источники погрешности

Погрешность прогнозирования кадра δI обусловлена следующими факторами:

- 1) погрешность модели динамики $\delta \mathbf{x}_{mdl}$: модель f несовершенна (упрощения, неучтённые возмущения и т. д.).
- 2) погрешность начальных условий $\delta \mathbf{x}_0$: неточное знание $\mathbf{x}(t_0)$.
- 3) погрешность управления $\delta \mathbf{u}$: неточное знание или прогноз относительно будущего управления.
- 4) погрешность интегрирования $\delta \mathbf{x}_{num}$: численные ошибки при решении ОДУ.
- 5) погрешность внешней калибровки $\delta \mathbf{T}_{cb}$: неточное определение положения/ориентации камеры относительно корпуса.
- 6) погрешность модели сцены: если прогноз основан на 3D-модели сцены, то её неточности также вносят свой вклад.

Ниже мы сосредоточимся на первых четырёх, поскольку они непосредственно связаны с динамикой камеры.

Обозначим: $\hat{\mathbf{x}}(t)$ – прогнозируемое состояние (решение модели), $\check{\mathbf{x}}(t)$ – истинное состояние, $\mathbf{P}(t) = \Phi(\mathbf{x}(t)) \in SE(3)$ – положение камеры, полученное из состояния с помощью известного отображения Φ (включающего \mathbf{T}_{cb}).

Тогда погрешность позы [8] камеры:

$$\delta \mathbf{P}(t) = \check{\mathbf{P}}(t) \ominus \hat{\mathbf{P}}(t), \quad (1)$$

где \ominus — операция «разности» в SE(3) (например, через логарифмическое отображение в se(3)). Линеаризуем отображение Φ в окрестности $\hat{\mathbf{x}}(t)$:

$$\delta \mathbf{P}(t) \approx \mathbf{J}_{\Phi}(t) \cdot \delta \mathbf{x}(t), \quad (2)$$

$$\text{где } \mathbf{J}_{\Phi}(t) = \left. \frac{\partial \Phi}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}(t)} \in \mathbb{R}^{6 \times n} \quad (3)$$

– якобиан преобразования состояния в позу камеры. Он имеет 6 параметров: 3 трансляции и 3 угла.

Погрешность состояния

Погрешность состояния обозначим как $\delta \mathbf{x}(t)$. Эта погрешность есть разница между фактическим и предсказанным состоянием:

$$\delta \mathbf{x}(t) = \check{\mathbf{x}}(t) - \hat{\mathbf{x}}(t). \quad (4)$$

Тогда динамика этой погрешности описывается вариационным уравнением [9], [10], (похожим на уравнение состояния в уравнении Калмана [11, 12]):

$$\delta \dot{\mathbf{x}}(t) = \mathbf{A}(t) \delta \mathbf{x}(t) + \mathbf{B}(t) \delta \mathbf{u}(t) + \varepsilon(t), \quad (5)$$

где:

$$\mathbf{A}(t) = \left. \frac{\partial f}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}(t), \mathbf{u}(t)} \text{ – матрица Якоби [13] динамики,}$$

$$\mathbf{B}(t) = \left. \frac{\partial f}{\partial \mathbf{u}} \right|_{\hat{\mathbf{x}}(t), \mathbf{u}(t)} \text{ – влияние ошибки управления,}$$

$\varepsilon(t)$ – ошибка модели (разница между истинной и модельной динамикой).

Решение этого линейного неоднородного уравнения [14]:

$$\delta \mathbf{x}(t_0 + \Delta t) = \Phi(t_0 + \Delta t, t_0) \delta \mathbf{x}_0 + \int_{t_0}^{t_0 + \Delta t} \Phi(t_0 + \Delta t, \tau) [\mathbf{B}(\tau) \delta \mathbf{u}(\tau) + \varepsilon(\tau)] d\tau, \quad (6)$$

где $\Phi(t, t_0)$ — фундаментальная матрица решений (матрица перехода) системы $\delta \dot{\mathbf{x}}(t) = \mathbf{A}(t) \delta \mathbf{x}(t)$.

Погрешность проецирования точки на изображение

Пусть $\mathbf{X}(t) \in \mathbb{R}^3$ — трёхмерная точка сцены в глобальной системе координат. Её проекция \mathbf{p} [6] на изображение:

$$\mathbf{p} = \pi(\mathbf{P}^{-1} \tilde{\mathbf{X}}) \quad (7)$$

где $\tilde{\mathbf{X}} = [\mathbf{X}^T, \mathbf{1}]^T$, $\pi(\cdot)$ - функция проецирования камеры (включает внутренние параметры). Погрешность проекции:

$$\delta \mathbf{p} \approx \mathbf{J}_{\pi}(t) \cdot \delta \mathbf{P},$$

где $\mathbf{J}_{\pi} = \frac{\partial \pi}{\partial \mathbf{P}}$ — якобиан проекции по положению камеры (размер 2×6).

Подставляя $\delta \mathbf{P} \approx \mathbf{J}_{\Phi}(t) \cdot \delta \mathbf{x}$, получаем:

$$\delta \mathbf{p} \approx \mathbf{J}_{\pi}(t) \cdot \mathbf{J}_{\Phi}(t) \cdot \delta \mathbf{x}(t_0 + \Delta t). \quad (8)$$

Или так:

$$\delta \mathbf{p} \approx \mathbf{J}_{\text{tot}}(t) \cdot \delta \mathbf{x}(t_0 + \Delta t) \quad (9)$$

где:

$$\mathbf{J}_{\text{tot}}(t) = \mathbf{J}_{\pi}(t) \cdot \mathbf{J}_{\Phi}(t). \quad (10)$$

Итоговая погрешность проекции зависит от:

- чувствительности проекции к положению камеры \mathbf{J}_{π} ,
- чувствительность положения к состоянию ЛА \mathbf{J}_{Φ} ,
- накопленной ошибки состояния $\delta \mathbf{x}(t_0 + \Delta t)$, которая, в свою очередь, зависит от начальной ошибки, ошибки модели, управления и ошибки численного интегрирования.

Оценка величины погрешности

Если предположить, что все источники ошибок ограничены:

$$\|\delta \mathbf{x}_0\| \leq \sigma_0, \quad \|\delta \mathbf{u}(t)\| \leq \sigma_u, \quad \|\delta \varepsilon(\tau)\| \leq \sigma_\varepsilon, \quad \|\delta \mathbf{T}_{cb}\| \leq \sigma_c, \quad (11)$$

и если известна оценка нормы фундаментальной матрицы:

$$\|\Phi(t, t_0)\| \leq k(t - t_0) \quad (12)$$

Худший случай, (например, экспоненциальный рост при неустойчивой динамике), то можно получить верхнюю границу:

$$\|\delta \mathbf{p}\| \leq \|\mathbf{J}_{\text{tot}}\| \cdot \left[k(\Delta t)\sigma_0 + \int_0^{\Delta t} k(\Delta t - \tau)(\|\mathbf{B}\|\sigma_u + \sigma_\varepsilon)d\tau \right] + \|\mathbf{J}_{\pi}\|\sigma_c \quad (13)$$

Эта формула показывает, что погрешность прогноза кадра растёт со временем прогноза Δt , усиливается при неустойчивой динамике (большой k), чувствительна к ошибкам в начальных условиях и управлении, зависит от геометрии сцены (через \mathbf{J}_{π} : точки, расположенные ближе к краю кадра или на большом расстоянии, дают большую погрешность).

Рассмотрим конкретный пример для упрощённой модели и численную оценку для типичных параметров перемещаемой платформы. Возьмем 3D-модель, такую, что текущее значение углов Эйлера всегда известно. Для этого примера [15-17] примем следующие допущения: состояние платформы описывается так положение центра масс: $\mathbf{r} = [x, y, z]^T \in \mathbb{R}^3$, скорость равна $\mathbf{v} = [\dot{x}, \dot{y}, \dot{z}]^T \in \mathbb{R}^3$, углы Эйлера (тангаж θ , рыскание ψ , крен ϕ): $\boldsymbol{\eta} = [\phi, \theta, \psi]^T \in \mathbb{R}^3$ и угловая скорость: $\boldsymbol{\omega} = [p, q, r]^T \in \mathbb{R}^3$.

Таким образом, движение описывается вектором: $\mathbf{x} = [\mathbf{r}^T, \mathbf{v}^T, \boldsymbol{\eta}^T, \boldsymbol{\omega}^T]^T \in \mathbb{R}^{12}$.

Углы Эйлера считаются точно известными в каждый момент времени (например, благодаря высокоточному датчику). То есть погрешность в $\boldsymbol{\eta}$ отсутствует: $\delta \boldsymbol{\eta}(t) = \mathbf{0}$

Камера жёстко закреплена в точке \mathbf{r}_c^b в системе координат корпуса, без поворота относительно корпуса (для простоты):

$$\mathbf{T}_{cb} = \begin{bmatrix} \mathbf{J}_3 & \mathbf{r}_c^b \\ 0 & 1 \end{bmatrix}$$

Динамическая модель здесь – упрощённая кинематическая модель движения центра масс и вращения:

$\dot{\mathbf{r}} = \mathbf{v}$, $\dot{\mathbf{v}} = \mathbf{a}$ – известное (или прогнозируемое) ускорение,

$\dot{\boldsymbol{\eta}} = \mathbf{W}(\boldsymbol{\eta})\boldsymbol{\omega}$, где \mathbf{W} – матрица связи угловой скорости и производных углов Эйлера,

$\dot{\boldsymbol{\omega}} = \boldsymbol{\alpha}$ — прогнозируемое угловое ускорение.

Но поскольку $\boldsymbol{\eta}(t)$ известно точно, мы можем не моделировать его динамику, а использовать как заданную функцию времени.

Погрешность возникает только из-за неточного прогноза $\mathbf{r}(t)$ и $\mathbf{v}(t)$, то есть из-за ошибок в ускорении $\mathbf{a}(t)$.

Оценим погрешность проецирования точки сцены $\mathbf{X}(t) \in \mathbb{R}^3$ на изображение в момент $t = t_0 + \Delta t$, обусловленную ошибкой прогноза положения камеры $\delta \mathbf{r}_c(t)$.

1) Позиция для съёмки

Положение камеры в глобальной системе:

$$\mathbf{r}_c(t) = \mathbf{r}(t) + \mathbf{R}(\boldsymbol{\eta}(t))\mathbf{r}_c^b \quad \text{где } \mathbf{R}(\boldsymbol{\eta}) \in SO(3) \quad (14)$$

– матрица поворота из локальной системы координат в глобальную (строится на основе углов Эйлера).

Поскольку $\boldsymbol{\eta}(t)$ известно точно, ошибка в $\mathbf{r}_c(t)$ обусловлена только ошибкой в $\mathbf{r}(t)$:

$$\delta \mathbf{r}_c(t) = \delta \mathbf{r}(t). \quad (15)$$

Ориентация камеры $\mathbf{R}_c = \mathbf{R}(\boldsymbol{\eta}(t))$ – без ошибок.

2) Прогноз местоположения и его погрешность

Пусть ускорение моделируется как $\hat{\mathbf{a}}(t)$, но истинное ускорение – $\ddot{\mathbf{a}}(t) = \hat{\mathbf{a}}(t) + \boldsymbol{\varepsilon}_a(t)$, где $\|\boldsymbol{\varepsilon}_a(t)\| \leq \sigma_a$.

Тогда ошибка в положении за Δt :

$$\delta \mathbf{r}(t_0 + \Delta t) = \int_{t_0}^{t_0 + \Delta t} \int_{t_0}^{\tau} \boldsymbol{\varepsilon}_a(t) ds d\tau \quad (16)$$

Если $\boldsymbol{\varepsilon}_a(t)$ постоянно (это в худшем случае), то:

$$\|\delta \mathbf{r}\| \leq \frac{1}{2} \sigma_a (\Delta t)^2. \quad (17)$$

Если есть ошибка и в начальной скорости $\delta \mathbf{v}_0$, то:

$$\delta \mathbf{r} = \delta \mathbf{v}_0 \Delta t + \frac{1}{2} \sigma_a (\Delta t)^2 \quad (18)$$

3) Проекция на изображение

Пусть внутренние параметры камеры заданы матрицей $\mathbf{K} \in \mathbb{R}^{3 \times 3}$. Тогда проекция точки \mathbf{X} в пиксельные координаты:

$$\tilde{\mathbf{p}} = \mathbf{K} \mathbf{R}_c^T (\mathbf{X} - \mathbf{r}_c) \quad (19)$$

где

$$\mathbf{p} = \begin{bmatrix} \tilde{p}_x \\ \tilde{p}_z \\ \tilde{p}_y \\ \tilde{p}_z \end{bmatrix} \quad (20)$$

Линеаризуем по \mathbf{r}_c . Обозначим координаты точки в системе камеры:

$$\mathbf{X}_c = \mathbf{R}_c^T (\mathbf{X} - \mathbf{r}_c) = [X_c, Y_c, Z_c]^T \quad (21)$$

f_x, f_y — фокусные расстояния (из \mathbf{K}).

Тогда якобиан проекции при смещении камеры:

$$\frac{\partial \mathbf{p}}{\partial \mathbf{r}_c} = -\frac{1}{Z_c} \begin{bmatrix} f_x & 0 & -f_x \frac{X_c}{Z_c} \\ 0 & f_y & -f_y \frac{Y_c}{Z_c} \end{bmatrix} \mathbf{R}_c^T \quad (22)$$

Обозначим эту матрицу как $\mathbf{J}_p \in \mathbb{R}^{2 \times 3}$.

Тогда погрешность проекции будет равна:

$$\delta \mathbf{p} = \mathbf{J}_p \delta \mathbf{r}_c \quad (23)$$

Численный пример.

Пусть: $\Delta t = 0.5$ с, $\sigma_a = 0.2$ м/с² (типичная ошибка в моделировании ускорения БПЛА)
 $\delta \mathbf{v}_0 = 0$ (идеальная начальная скорость), $\mathbf{r}_c^b = [0.2, 0, 0.1]^T$ м (камера смещена вперёд и вверх),
 Углы Эйлера (тангаж $\theta = 10^\circ$, рыскание $\psi = 45^\circ$, крен $\phi = 0^\circ$): вычисляем \mathbf{R}_c
 Точка съёмки: $\mathbf{X} = [100, 50, 0]^T$ м (на земле, в 100 м от старта). Камера [18], [19]:
 $f_x, f_y = 800$ пикселей, $Z_c = 95$ м (примерно).

Тогда, подставляя значения в вышеприведенные формулы, получаем:

$$\begin{aligned} \|\delta \mathbf{r}\| &\leq \frac{1}{2} \cdot 0.2 \cdot (0.5)^2 = 0.025 \text{ м} = 2,5 \text{ см}, \\ \|\mathbf{J}_p\| &\approx f_x \frac{1}{Z_c} \approx \frac{800}{95} \approx 8.4 \text{ пикс/м}, \\ \|\delta \mathbf{p}\| &\leq 8.4 \cdot 0.025 \approx 0.21 \text{ пикселя}. \end{aligned}$$

Вывод: при хорошей модели ускорения и точных углах Эйлера ошибка прогноза кадра через 0,5 с составляет менее 0,3 пикселя, что приемлемо для многих задач (стабилизации, визуального сопровождения, оптической навигации).

Но если: $\Delta t = 2$ с, $\|\delta \mathbf{r}\| \leq 0.4$ м, $Z_c = 20$ м (низкий полёт), то $f/Z_c = 40$ пиксель/м

Тогда $\|\delta \mathbf{p}\| \leq 40 \cdot 0.4 \approx 16$ пикселей, а это уже существенная ошибка.

Заключение

С качественной точки зрения погрешность прогнозирования кадра – это композиция ошибок:

1. Динамическая ошибка – накопление неточностей в модели движения.
2. Кинематическая трансформация – преобразование ошибки состояния в ошибку положения камеры.
3. Проекционная чувствительность – как эта ошибка в положении влияет на расположение пикселей.

Для минимизации погрешности важно:

- использовать точные динамические модели (возможно, с учётом аэродинамики и возмущений),
- точно оценивать начальное состояние (например, с помощью расширенного фильтра Калмана [20], [21]),
- учитывать неопределённость управления,
- калибровка внешних параметров камеры с высокой точностью.

Погрешность проекции растёт как:

$$\|\delta \mathbf{p}\| \propto \frac{1}{Z_c} \cdot \left(\|\delta \mathbf{v}_0\| \Delta t + \frac{1}{2} \sigma_a (\Delta t)^2 \right),$$

где первый множитель – это ошибка масштаба проекции, а второй – ошибка позиции.

Можно сделать такие выводы для практики:

- существует критическая высота полёта (чем ниже, тем больше погрешность в пикселях),
- погрешность увеличивается в квадрате со временем прогноза
- погрешность линейно зависит от ошибки ускорения.

При точно известных углах Эйлера основной вклад в погрешность прогноза кадра вносит неточность прогноза траектории центра масс, и, особенно, неточность ускорения. В типичных условиях [22] (умеренное время прогноза, средняя высота) погрешность может составлять менее 1 пикселя, но быстро растёт при увеличении Δt или снижении высоты.

Список литературы

1. Vukoti' V., Pinteа S-L., Raymond C., Gravier G., Jan van Gemert. One-Step Time-Dependent Future Video Frame Prediction with a Convolutional Encoder-Decoder Neural Network. [inria.hal.science, hal-01467064/document](https://hal.science/hal-01467064/document)
2. Mahjourian R., Wicke M., Angelova A. Geometry-Based Next Frame Prediction from Monocular Video. arXiv:1609.06377v2 Jun 2017
3. Ландау Л. Д., Лифшиц Е. М. Теоретическая физика, Учеб. Пособие В 10 томах, Т.1 Механика, Гл. ред. Физ.-мат. Лит. 1988.-216 с. ISBN 5-02-013650-9 (т.1)
4. Голдстейн Г. Классическая механика Наука, Физ.-мат. лит. 1975
5. M. Li, A.I. Mourikis: "Online Temporal Calibration for Camera-IMU Systems: Theory and Algorithms," International Journal of Robotics Research, 33(7), pp. 947-964, June 2014.
6. A.I. Mourikis, N. Trawny, S.I. Roumeliotis, A. Johnson, A. Ansar, L. Matthies: "Vision-Aided Inertial Navigation for Spacecraft Entry, Descent, and Landing," IEEE Transactions on Robotics, 25(2), pp. 264-280, April 2009. 2009 IEEE Transactions on Robotics Best Paper Award (King-Sun Fu Memorial Award)
7. G. Huang, A.I. Mourikis, S.I. Roumeliotis: "On the Consistency of Multi-robot Cooperative Localization," Proceedings of the Robotics: Science and Systems Conference, Seattle, WA, June 28-July 1, 2009.
8. M. Li, A.I. Mourikis, "Optimization-based Estimator Design for Vision-Aided Inertial Navigation," Proceedings of the Robotics: Science and Systems Conference, Sydney, Australia, July 2012, pp. 241-248.
9. Эльсгольц Л. Э. Дифференциальные уравнения и вариационное исчисление, М., Наука, 1969
10. Гельфанд И. М. Фомин С. В. Вариационное исчисление. Наука, Физ.-мат. Лит. 1961
11. Зейферт Г., Трефалль В., Вариационное исчисление в целом. Ижевск, Издательский дом «Удмуртский университет» 2000, 160 стр. ISBN 5-7029-0336-6
12. Julier, S.J.; Uhlmann, J.K. (2004). "Unscented filtering and nonlinear estimation" *Proceedings of the IEEE*. **92** (3): 401–422. doi:10.1109/jproc.2003.823141
13. Фихтенгольц Г. М., Курс дифференциального и интегрального исчисления, т 1-3, М, Наука, 1969
14. Понтрягин Л. С. Обыкновенные дифференциальные уравнения, М, Наука, 1970
15. Зуев, В. М. Отбор значимых данных при разработке робототехнических устройств [Текст] / В. М. Зуев, С. А. Рудской // Искусственный интеллект: теоретические аспекты и практическое применение : материалы Донецкого Международного круглого стола. Донецк : ФГБНУ «ИПИИ», 2025. 296 с. С. 90-93.
16. Зуев, В. М. Оценка собственного местоположения аппарата на основе анализа видеоизображения / В. М. Зуев, С. Б. Иванова // Проблемы искусственного интеллекта. 2024. № 2 (33). С. 21–28. ISSN 2413-7383. DOI 10.24412/2413-7383-2024-2-21-28.
17. Зуев, В. М. Улучшение визуального обнаружения объектов / Искусственный интеллект: теоретические аспекты, практическое применение : материалы Донецкого международного научного круглого стола. Донецк : ФГБНУ «ИПИИ», 2024. – 238 с. – С. 67–72.
18. Фролов, В. В. Анализ характеристик видеокамер оптического диапазона для работы в составе интеллектуальных систем обнаружения и оценки объектов беспилотными авиационными системами // Искусственный интеллект: теоретические аспекты, практическое применение : материалы Донецкого международного научного круглого стола. Донецк : ФГБНУ «ИПИИ», 2024. 328 с. С. 304–310.
19. Фролов, В. В. Методы синхронного сложения видеосигналов // Искусственный интеллект: теоретические аспекты, практическое применение : материалы Донецкого международного научного круглого стола. Донецк : ФГБНУ «ИПИИ», 2023. 252 с. С. 217–220.
20. Kalman R.E. New Results in Linear Filtering and Prediction Theory. Transactions of the ASME, Journal of Basic Engineering. March 1961, p.95-108.
21. Grewal M. S. Andrews A. Kalman filtering: Theory and practice using MATLAB. 2001. John Wiley & Sons. ISBN 0-471-39254-5
22. ОДМ 218.9.017–2019 Методические рекомендации по производству аэрофототопографических работ с использованием беспилотных летательных аппаратов при изысканиях в целях строительства и реконструкции автомобильных дорог (РОСАВТОДОР) Москва 2021 – 55с.

References

1. Vukoti'V., Pinteá S-L., Raymond C., Gravier G., Jan van Gemert. One-Step Time-Dependent Future Video Frame Prediction with a Convolutional Encoder-Decoder Neural Network. [inria.hal.science, hal-01467064/document](https://inria.hal.science/hal-01467064/document)
2. Mahjourian R., Wicke M., Angelova A. Geometry-Based Next Frame Prediction from Monocular Video. arXiv:1609.06377v2 Jun 2017
3. Landau L D Lifshits E M Theoretical Physics, Textbook. Manual In 10 volumes, Vol.1 Mechanics, Head. ed. Phys.-math. Lit. 1988.-216 p. ISBN 5-02-013650-9 (Vol.1)
4. Goldstein G. Classical Mechanics, Addison-Wesley, 2002, 638 p, ISBN 978-0-201-65702-9
5. M. Li, A.I. Mourikis: "Online Temporal Calibration for Camera-IMU Systems: Theory and Algorithms," International Journal of Robotics Research, 33(7), pp. 947 964, June 2014.
6. A.I. Mourikis, N. Trawny, S.I. Roumeliotis, A. Johnson, A. Ansar, L. Matthies: "Vision-Aided Inertial Navigation for Spacecraft Entry, Descent, and Landing," IEEE Transactions on Robotics, 25(2), pp. 264-280, April 2009. 2009 IEEE Transactions on Robotics Best Paper Award (King-Sun Fu Memorial Award)
7. G.Huang, A.I. Mourikis, S.I. Roumeliotis: "On the Consistency of Multi-robot Co operative Localization," Proceedings of the Robotics: Science and Systems Conference, Seattle, WA, June 28-July 1, 2009.
8. M.Li, A.I. Mourikis, "Optimization-based Estimator Design for Vision-Aided Inertial Navigation," Proceedings of the Robotics: Science and Systems Conference, Sydney, Australia, July 2012, pp. 241-248.
9. L. E. Elsgolts. Differential Equations and Calculus of Variations, Moscow, Nauka, 1969
10. Gelfand I. M. Fomin S. V. Variational Calculus, Moscow, Nauka, 1961
11. Seifert G., Trefall V., Calculus of Variations in General. Izhevsk, Udmurt University Publishing House 2000, 160 pages. ISBN 5-7029-0336-6
12. Julier, S.J.; Uhlmann, J.K. (2004). "[Unscented filtering and nonlinear estimation](https://doi.org/10.1109/jproc.2003.823141)" *Proceedings of the IEEE*. **92** (3): 401–422. doi:10.1109/jproc.2003.823141
13. Fikhtengolts, G. M., Course of Differential and Integral Calculus, vols. 1-3, Moscow, Nauka, 1969
14. Pontryagin L. S. Ordinary Differential Equations, Moscow, Nauka, 1970
15. Zuev, V. M. Selection of Significant Data in the Development of Robotic Devices [Text] / V. M. Zuev, S. A. Rudskoy // Artificial Intelligence: Theoretical Aspects and Practical Application : Materials of the Donetsk International Round Table. – Donetsk : FGBNU "IPII", 2025. – 296 p. – Pp. 90-93.
16. Zuev, V. M. Estimation of the Device's Own Location Based on Video Image Analysis [Text] / V. M. Zuev, S. B. Ivanova // Problems of Artificial Intelligence. – 2024. – No. 2 (33). – Pp. 21–28. – ISSN 2413-7383. – DOI 10.24412/2413-7383-2024-2-21-28.
17. Zuev, V. M. Improving Visual Object Detection [Text] / V. M. Zuev // Artificial Intelligence: Theoretical Aspects and Practical Applications : Proceedings of the Donetsk International Scientific Round Table. – Donetsk: IPII, 2024. – 238 p. – Pp. 67–72.
18. Frolov, V. V. Analysis of the Characteristics of Optical Range Video Cameras for Use in Intelligent Systems for Detecting and Evaluating Objects by Unmanned Aerial Systems [Text] / V. V. Frolov // Artificial Intelligence: Theoretical Aspects and Practical Applications : Proceedings of the Donetsk International Scientific Round Table. – Donetsk : IPII, 2024. – 328 p. – Pp. 304–310.
19. Frolov, V. V. Methods of synchronous addition of video signals [Text] / V. V. Frolov // Artificial Intelligence: Theoretical Aspects, Practical Application : Materials of the Donetsk International Scientific Round Table. – Donetsk : FGBNU "IPII", 2023. – 252 p. – Pp. 217–220.
20. Kalman R.E. New Results in Linear Filtering and Prediction Theory. Transactions of the ASME, Journal of Basic Engineering. March 1961, p.95-108.
21. Grewal M. S. Andrews A. Kalman filtering: Theory and practice using MATLAB. 2001. John Wiley & Sons. ISBN 0-471-39254-5
22. ODM 218.9.017–2019 Methodological Recommendations for the Production of Aerophototopographic Works Using Unmanned Aerial Vehicles for Surveys for the Construction and Reconstruction of Highways (ROSAVTODOR) Moscow 2021 – 55p.

RESUME

V.M. Zuev

Error in predicting the position of the next film frame

Predicting the next frame of a scene is used in object detection and detection algorithms, optical odometry, optical navigation, and other applications. The effectiveness of these algorithms largely depends on the frame prediction error. Most studies address the accuracy issue by selecting the neural network architecture, which does not allow for general conclusions.

Methods of classical mechanics, projective geometry, and variational calculus were used to estimate the errors.

A formula for calculating the prediction error of the next frame of a scene was obtained. It can be used in object detection and detection algorithms. A numerical example of the calculation is provided.

From the analysis of the obtained dependencies, the following practical conclusions can be drawn: the lower the camera position above the surface, the greater the pixel error. The error increases squared with the prediction time and is linearly dependent on the acceleration error.

РЕЗЮМЕ

В.М. Зуев

Погрешность прогнозирования положения следующего кадра съёмки

Прогнозирование следующего кадра съёмки позволяет использоваться в алгоритмах обнаружения и детекции объектов, в оптической одометрии, оптической навигации и других приложениях. Эффективность этих алгоритмов в значительной мере зависит от погрешности прогнозирования кадра съёмки. Большинство работ решает проблему точности подбором архитектуры нейросети, что не позволяет сделать обобщенные выводы.

Для оценки погрешностей применены методы классической механики, проективной геометрии и вариационного исчисления.

Получена формула, позволяющая находить погрешность прогнозирования следующего кадра съёмки. Она может использоваться в алгоритмах обнаружения и детекции объектов. Приведен численный пример расчета.

Из анализа полученных зависимостей можно сделать такие выводы для практики: чем ниже позиция камеры над поверхностью, тем больше погрешность в пикселях, погрешность увеличивается в квадрате со временем прогноза и линейно зависит от ошибки ускорения.

Зуев Владимир Михайлович - Научный сотрудник, отдел интеллектуальных робототехнических систем, Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», г. Донецк, Российская Федерация.
Область научных интересов: интеллектуальные робототехнические системы.
эл. почта zvm.ipai@mail.ru, адрес: 283048, г. Донецк, ул. Артема, д. 118 б, РФ.
Телефон: +7949 4048702

Статья поступила в редакцию 12.09.2025.

Раздел 6
Section 6

Системный анализ,
управление и обработка
информации, статистика

*System analysis,
management and information
processing, statistics*

УДК 519.4

DOI 10.24412/2413-7383-2025-4-39-183-193

П.В. Васильев¹, О.В. Даринцев², Р.А. Мунасыпов²¹Института механики им. Р.Р. Мавлютова УФИЦ РАН

450054 Россия, Республика Башкортостан, г. Уфа, Проспект Октября, 71

²ФГБОУ ВО «Уфимский университет науки и технологий»

450076, Россия, Республика Башкортостан, г. Уфа, ул. Заки Валиди, 32

ПРИМЕНЕНИЕ АССОЦИАТИВНОГО РЕГУЛЯТОРА В КОНТУРЕ УПРАВЛЕНИЯ МЕХАТРОННЫМ МОДУЛЕМ*

P.V. Vasilev¹, O.V. Darintsev², R.A. Munasypov²¹Institute of Mechanics named after R.R. Mavlyutov UFIC RAS

450054 Russia, Republic of Bashkortostan, Ufa, 71 Oktyabrya Av.

²Ufa University of Science and Technology

450076, Russia, Republic of Bashkortostan, Ufa, 32 Zaki Validi str.

APPLICATION OF AN ASSOCIATIVE CONTROLLER IN THE CONTROL CIRCUIT OF A MECHATRONIC MODULE

В статье рассматривается актуальная задача повышения эффективности систем управления мехатронными модулями в контексте концепции "Индустрия 4.0". Проводится сравнительный анализ четырех различных методов применения ассоциативного регулятора на базе нейронной сети *СМАС* (*Cerebellar Model Articulation Controller*) для улучшения качества работы классического ПИД-регулятора. Рассматриваются следующие архитектуры: параллельная работа ПИД-регулятора и ассоциативного регулятора для компенсации возмущений, динамическая настройка коэффициентов ПИД-регулятора с помощью ассоциативной памяти, гибридный метод, сочетающий оба подхода, а также полная замена ПИД-регулятора ассоциативным. Анализ проводится на примере модели сервопривода в среде *Matlab/Simulink*. В работе оценивается эффективность предложенных методов по таким показателям, как быстродействие и перерегулирование, и демонстрируются преимущества гибридных интеллектуальных систем управления.

Ключевые слова: ПИД-регулятор, ассоциативная память, настройка регулятора, *СМАС*.

The article discusses the urgent task of improving the efficiency of mechatronic module control systems in the context of the "Industry 4.0" concept. A comparative analysis of four different methods of using an associative controller based on the CMAC neural network (Cerebellar Model Articulation Controller) is carried out to improve the quality of the classic PID controller. The following architectures are considered: parallel operation of a PID controller and an associative controller to compensate for disturbances, dynamic adjustment of the coefficients of the PID controller using associative memory, a hybrid method combining both approaches, as well as the complete replacement of the PID controller with an associative one. The analysis is carried out using the example of a servo drive model in the Matlab/Simulink environment. The paper evaluates the effectiveness of the proposed methods in terms of performance and overshoot, and demonstrates the advantages of hybrid intelligent control systems.

Key words: PID controller, associative memory, adjustment of the controller, CMAC.

* Работа посвящена исследованию эффективности применения ассоциативного регулятора (АР) для улучшения качества управления мехатронными модулями.

Введение

В современных реалиях искусственный интеллект (ИИ) [1-3] становится неотъемлемой частью промышленной автоматизации, позволяя создавать "умные" производства, управляемые взаимосвязанными системами и аналитикой на основе данных. Интеграция интеллектуальных технологий в системы управления технологическим оборудованием позволяет в режиме реального времени осуществлять текущий контроль и выполнять коррекцию параметров работы оборудования. Наиболее известны успешные решения задач управления сложными динамическими объектами с использованием таких методов ИИ, как нейронные сети и нечеткая логика [4-11], позволяющие получать компактные решения для систем при отсутствии точных математических моделей [12]. Несмотря на успешность таких подходов для решения задач управления, наличие существенных недостатков: сложность настройки, объемные обучающие выборки, длительность обучения и т.д. отсутствие дообучения в режиме реального времени, - потребовали провести исследования по оценке эффективности применения альтернативных методик ИИ.

Цель работы - оценить перспективы использования ассоциативной памяти (АП).

Постановка задачи. Принимая во внимание специфику работы АП, которая позволяет осуществлять доступ к данным не по адресу, а по их содержанию [13], восстанавливать информацию по ее фрагменту или зашумленной копии [14-16], необходимо на ее базе:

1. Разработать архитектуру ассоциативного регулятора (АР);
2. Провести предварительный синтез адаптивных систем управления мехатронными модулями, обладающие способностью эффективного функционирования в условиях неопределенности и запоминания новых состояний без полного переобучения всей системы.

Эффективность предлагаемого решения будет оценена по результатам сравнительного анализа основных схем применения АР в системах управления мехатронными модулями.

Методология

В данной работе для реализации АР была выбрана нейронная сеть *СМАС* (*Cerebellar Model Articulation Controller*) в качестве главного элемента АП. Выбор в пользу *СМАС* обусловлен рядом ее ключевых преимуществ по сравнению с другими нейросетевыми архитектурами и методами нечеткой логики:

– Быстрое обучение: В отличие от сетей с полносвязной структурой, требующих обновления всех весовых коэффициентов на каждой итерации, *СМАС* обладает локальной обобщающей способностью. Это означает, что при подаче входного сигнала активируется и корректируется лишь небольшая группа нейронов, что значительно ускоряет процесс обучения. Данное свойство *СМАС* создает явные предпосылки для ее применения в системах управления, функционирующих в режиме реального времени.

– Простота вычислений: Алгоритм работы *СМАС* основан на простых операциях чтения из таблицы и суммирования, что снижает вычислительную нагрузку по сравнению со сложными вычислениями функций активации или принадлежности в многослойных перцептронах или нечетких системах.

– Хорошая способность к обобщению: Структура *СМАС* позволяет ей формировать схожие выходные сигналы для близких по значению входных векторов. Это достигается за счет того, что близкие входные сигналы активируют пересекающиеся наборы ячеек памяти, обеспечивая плавное изменение управляющего воздействия.

– Меньшая потребность в априорных знаниях: По сравнению с системами на нечеткой логике, *СМАС* не требует формирования сложной базы правил "если-то", что упрощает ее первоначальную настройку и адаптацию.

Для сравнительного анализа были выбраны четыре возможных архитектуры системы управления [12], которые различаются способом взаимодействия АР и ПИД-регулятора [17-21].

В работах [17], [19-22] описывается работа АР, *СМАС* преобразует многомерный входной сигнал, это могут быть ошибка $e(t)$ и ее производная $\Delta e(t)$, в набор адресов в нескольких таблицах весов с помощью специального механизма отображения (меток). Выходным сигналом регулятора является сумма весов, хранящихся по этим адресам, что позволяет локально настраивать реакцию системы для разных режимов работы, например, для переходного и стационарного состояний. Принцип работы и структура *СМАС* представлены в соответствии с рисунком 1.

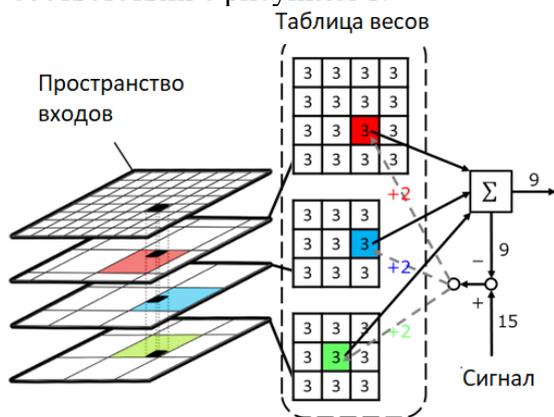


Рисунок 1 – Принцип работы и структура *СМАС*

В первом варианте архитектуры в соответствии с рисунком 2, ПИД-регулятор и АР работают параллельно. ПИД-регулятор обеспечивает базовое управление, а АР выполняет роль компенсирующего элемента. На вход АР подаются данные о состоянии системы, а ее выходной сигнал суммируется с выходным сигналом ПИД-регулятора. Основная задача АР в этой схеме — компенсировать нелинейности объекта управления и внешние возмущения, которые не могут быть эффективно обработаны линейным ПИД-регулятором.

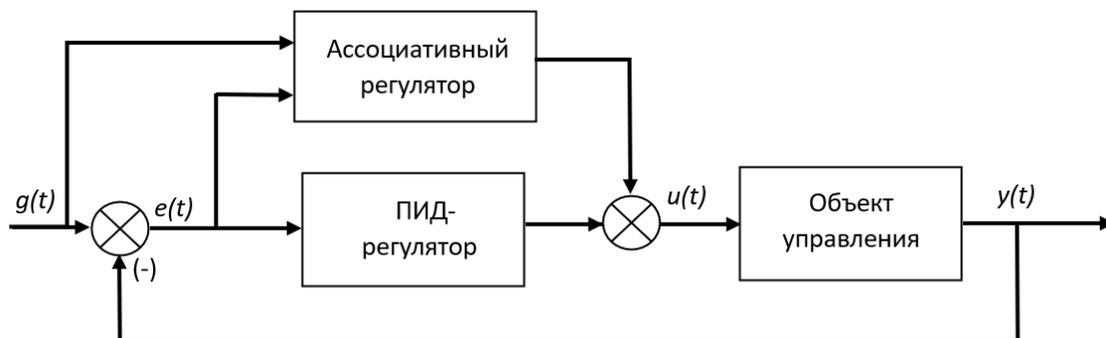


Рисунок 2 – Первый вариант архитектуры

Во втором варианте архитектуры АР в соответствии с рисунком 3, применяется в качестве "настройщика" параметров ПИД-регулятора. В отличие от систем с фиксированными коэффициентами, здесь АР динамически изменяет коэффициенты

ПИД-регулятора в зависимости от текущего режима работы системы. На вход сети подаются параметры, характеризующие динамику процесса, а на выходе нейросетью формируются значения коэффициентов K_p , K_i и K_d , для обучения которой использовался предварительно рассчитанный датасет на полном диапазоне входных сигналов.

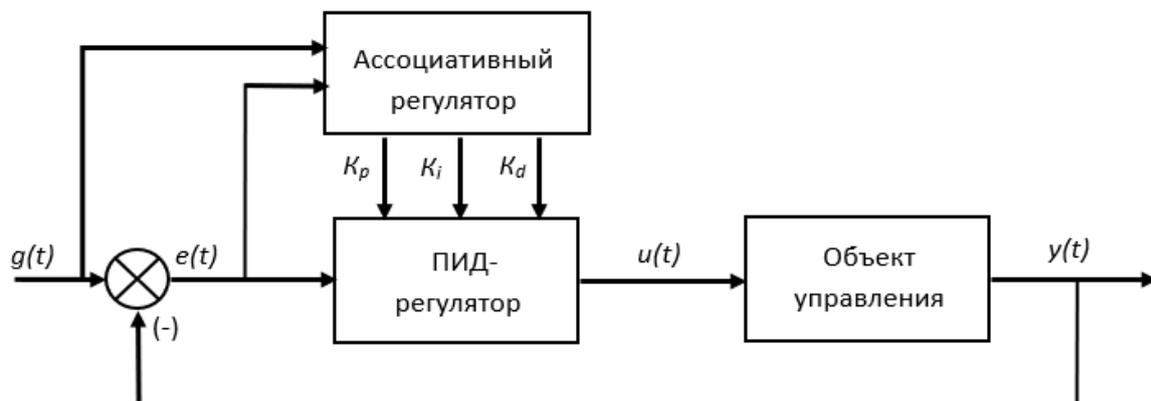


Рисунок 3 – Второй вариант архитектуры

Третий вариант архитектуры — это комбинация двух предыдущих подходов в соответствии с рисунком 4. В данном варианте системы одна сеть АР используется для динамической настройки коэффициентов ПИД-регулятора, в то же время второй АР работает параллельно основному контуру управления и корректирует его итоговый выходной сигнал.

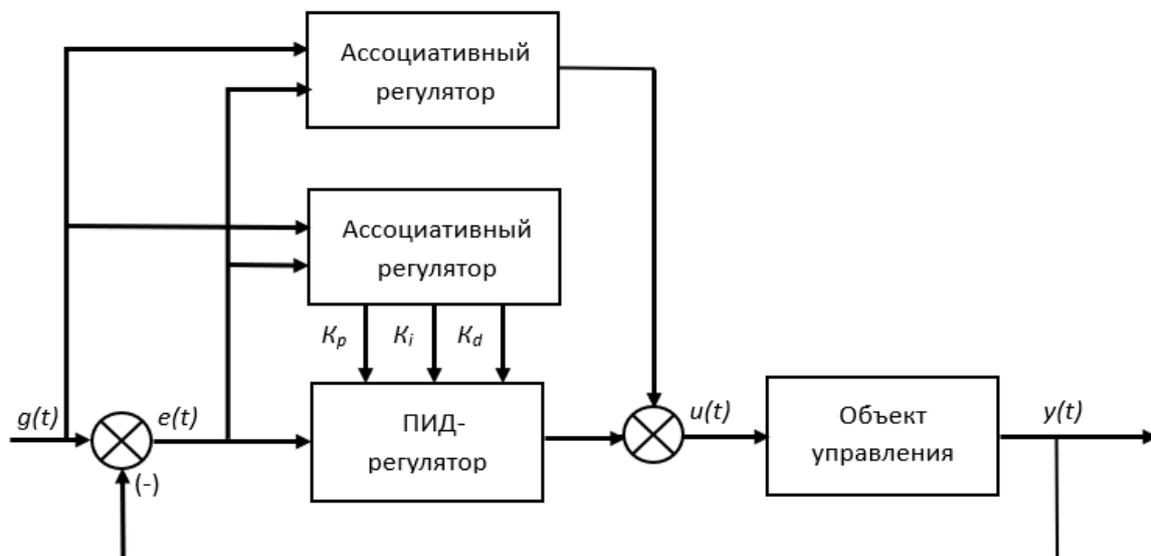


Рисунок 4 – Третий вариант архитектуры

В четвертом варианте архитектуры АР полностью заменяет собой классический ПИД-регулятор в соответствии с рисунком 5. Входными сигналами для АР служат ошибка регулирования $e(t)$, ее интеграл $\int e(t)dt$ и производная $de(t)/dt$, что функционально соответствует трем составляющим ПИД-закона. Обучение сети направлено на то, чтобы она научилась формировать управляющий сигнал, минимизирующий ошибку. Преимущество такого подхода заключается в способности АР реализовывать аппроксимированную сложную нелинейную зависимость между входами и выходом, что теоретически позволяет добиться более высокого качества управления для нелинейных объектов по сравнению с классическим линейным ПИД-регулятором.



Рисунок 5 – Четвертый вариант архитектуры

Результаты

Перед тем как проводить эксперименты на реальном объекте проведем численное моделирование. Для примера моделирования возьмем модель сервопривода, описываемого колебательным звеном:

$$W(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2},$$

где ω_n - собственная частота равная 7,5 рад/с, ξ - коэффициент демпфирования равный 0,5. Моделирование проводится с использованием ПО *Matlab/Simulink*.

Результат моделирования системы с ПИД-регулятором в соответствии с рисунком 6 и коэффициентами: $K_p=3,4$, $K_i=13,2$, $K_d=0,22$. Как видно из графика ПИД-регулятор не может полностью компенсировать перерегулирование возникающие при управлении сервоприводом, что неблагоприятно сказывается на работе системы, в частности, наблюдаются снижение точности и повторяемости позиционирования, повышение механического износа и т.д.

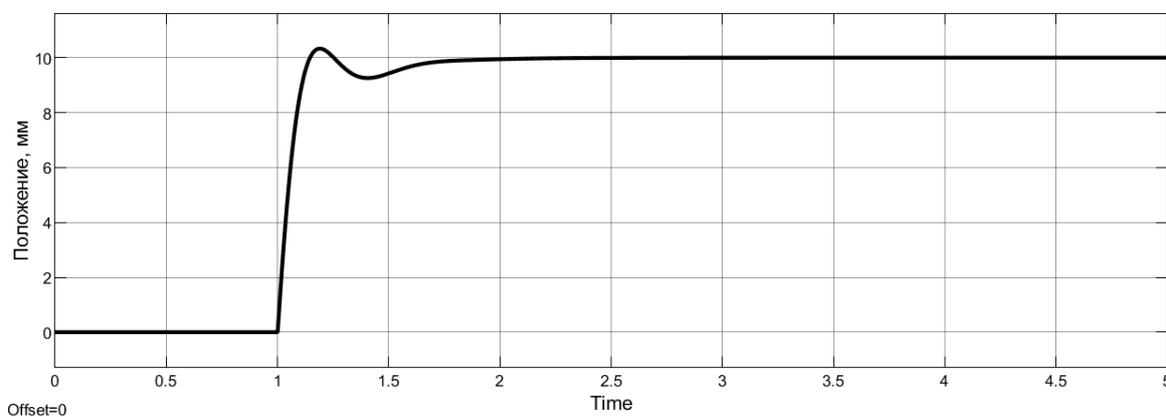


Рисунок 6 – Результат моделирования работы ПИД-регулятора

Далее согласно предложенным методам в систему управления был добавлен АР. Для работы АР было проведено его обучение в соответствии с рисунком 7, в ходе которого выполнены настройки диапазонов работы, настройка параметров квантования сигналов и коэффициентов регулятора для получения более качественных результатов.

В результате обучения наименьшую среднеквадратичную ошибку показала гибридная система управления. После обучения графики переходных процессов для разных вариантов систем управления представлены в соответствии с рисунком 8 и в таблице 1.

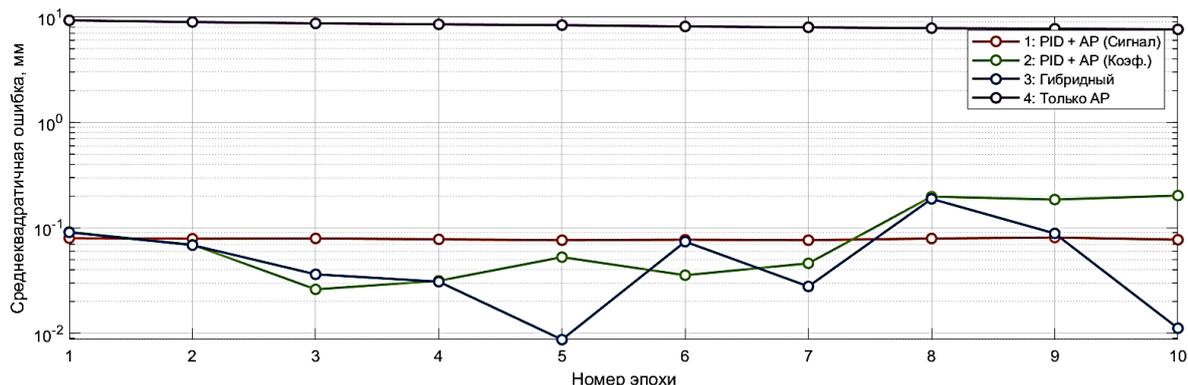


Рисунок 7 – Процесс обучения АП

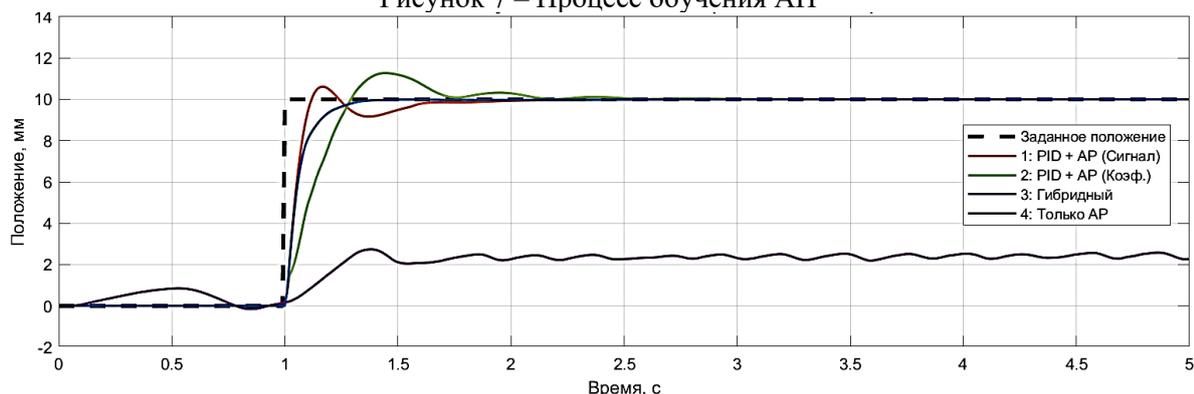


Рисунок 8 – Графики переходных процессов на последней эпохе

В таблице 1 приведены параметры для оценки качества переходного процесса, обучения AP, и оценки вычислительной мощности. Средняя квадратичная ошибка показывает, насколько качественно AP справляется с работой. Среднее время генерации команды управления показывает типичную вычислительную нагрузку при реализации алгоритма управления на протяжении всей симуляции. Максимальное время генерации команды управления показывает самое большое время, которое потребовалось на генерацию команды управления в один период управления. В ходе проведенного моделирования было принято, что длительность одного периода составляет 10 000 мкс. Параметр максимального времени генерации команды управления важен поскольку мехатронные системы могут применяться в системах реального времени, что, согласно теореме Котельникова, требует, чтобы AP работал с частотой как минимум в два раза выше собственной частоты системы.

Таблица 1 – Результаты моделирования (последняя эпоха)

№	Параметр	ПИД-регулятор	ПИД+AP Вариант 1	ПИД+AP Вариант 2	Гибридный Вариант 3	AP Вариант 4
1.	Перерегулирование, %	3,30	6,10	12,70	0	6,80
2.	Время регулирования, с	0,51	0,50	0,64	0,21	0,77
3.	Средняя квадратичная ошибка	-	0,7721	0,20251	0,01117	7,63059
4.	Среднее время генерации команды управления, мкс	-	2,57	5,93	7,46	2,01
5.	Максимальное время генерации команды управления, мкс	-	42,2	18,00	40,40	5,00

Как видно по результатам моделирования лучший результат переходного процесса показал гибридный метод, но при этом у него самые большие значения в других показателях - среднее время генерации команды управления, максимальное время генерации команды управления. Самый худший результат по данному критерию имеет система управления, где применяется только АР, что было ожидаемо, так как было реализовано недостаточное количество эпох обучения, а следовательно, недостаточно корректно была выполнена точная настройка параметров квантования.

Выводы

В ходе обработки результатов математического моделирования новых архитектур систем управления мехатронным модулем был выполнен сравнительный анализ четырех способов интеграции ассоциативного регулятора на базе нейронной сети *СМАС*. В качестве объекта управления рассматривался сервопривод, для которого получены результаты применения классического ПИД-регулятора, показавшие ограниченные возможности уменьшения колебаний в системе.

Полученные результаты показали, что наилучшее качество управления по интегральному критерию — максимальное быстродействие при отсутствии перерегулирования — достигается при использовании гибридной архитектуры, сочетающей динамическую настройку коэффициентов ПИД-регулятора и одновременную коррекцию его выходного сигнала с помощью ассоциативной памяти. Архитектура, в которой была сделана попытка заменить ПИД-регулятор на ассоциативный регулятор, продемонстрировала наименее эффективный результат, что позволило сократить количество синтезируемых вариантов при проведении натурных экспериментов.

В ходе экспериментов также была подтверждена перспективность применения интеллектуальных подходов для синтеза контуров адаптации и повышения качества систем управления. Однако, перед практической реализацией исследованных архитектур и выбором направлений дальнейшего развития необходимо провести не один цикл исследований. В первую очередь, требуется разработать четкие методические указания по настройке параметров квантования входных и выходных сигналов *СМАС*, так как этот аспект оказывает существенное влияние на точность и стабильность регулирования. Не менее актуальным направлением является исследование методов интеграции априорных знаний об объекте управления в обучаемую модель для ускорения процесса обучения и повышения его эффективности.

Список литературы

1. Денике Л.В. Зарубежный опыт цифровой трансформации предприятий за счет применения искусственного интеллекта // Прикладные экономические исследования. 2025. № 2. С. 249–255. doi: 10.47576/2949-1908.2025.2.2.028.
2. Айдосова А.А., Еремекбаева Д.Д. Роль искусственного интеллекта в цифровой модернизации экономики // Universum: экономика и юриспруденция: электрон. научн. журн. 2025. 6(128).
3. Анцыферов С.С., Фазилова К.Н., Муравьев Д.С. Интеллектуальные системы управления качеством технологических процессов и оборудования наукоёмкого производства // Искусственный интеллект: теоретические аспекты, практическое применение: материалы Донецкого международного научного круглого стола, Донецк, 28 мая 2025 года. Донецк: Федеральное государственное бюджетное научное учреждение "Институт проблем искусственного интеллекта", 2025. С. 163–165. EDN CRMCLC.
4. Бураков М. В., Коновалов А. С. Нечеткий супервизор ПИД-регулятора. Информационно-управляющие системы, 2018, № 5, с. 13–21. doi:10.31799/1684-8853-2018-5-13-21.
5. Жиллов Р.А. Постройка ПИД-регулятора с использованием нейронных сетей // Известия Кабардино-Балкарского научного центра РАН. 2022. № 5 (109). С. 38–47. DOI: 10.35330/1991-6639-2022-5-109-38-47

6. Чибисов А.В., Попов А.Г., Мозговой Н.А., Пиджаков В.А., Пиджаков В.А. Автоматическая настройка ПИД-регулятора с применением машинного обучения для обеспечения стабильной работы автономных электронных систем при термоциклировании // Труды МФТИ. 2024. №2 (62).
7. Воевода А.А., Романников Д.О. Метод синтеза регуляторов для многоканальных систем с использованием нейронных сетей. Вычислительные технологии. 2020; № 25(3): С. 111–118.
8. S. Kumar Krishnamoorthy, N. Das, P. Gudimetla and K. Emami, "Enhanced Speed Control for BLDC Motors Using WOA-Integrated PID Controller Optimization," in IEEE Access, vol. 12, pp. 162465-162475, 2024, doi: 10.1109/ACCESS.2024.3480349.
9. R. -M. Motorga et al., "Adaptive FOPID Controller with AI Techniques for the Driving Motors Driving the Furnance Hearth," 2022 10th International Conference on Control, Mechatronics and Automation (ICCMA), Belval, Luxembourg, 2022, pp. 122-128, doi: 10.1109/ICCMA56665.2022.10011614.
10. V. Revathi, Y. P. Rangaiyah, A. Dutt, A. A. Hameed, R. V. S. S. Nagini and D. K. Yadav, "Design of Small UAV Autopilot Systems Using Adaptive Controllers, PID Controllers, and AI," 2025 International Conference on Next Generation Communication & Information Processing (INCIP), Bangalore, India, 2025, pp. 20-25, doi: 10.1109/INCIP64058.2025.11020391.
11. R. Priyadarsini, A. Nayak and A. Kumar Barisal, "Frequency regulation of multi area interconnected system by using artificial intelligence based controller," 2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur, Odisha, India, 2022, pp. 1-6, doi: 10.1109/iSSSC56467.2022.10051434.
12. Васильев В.И., Ильясов Б.Г. Интеллектуальные системы управления. Теория и практика: учебное пособие. – М.: Радиотехника, 2009. 392 с.
13. Огнев И.В., Борисов В.В., Сутула Н.А. Ассоциативные память, среды, системы. – М.: Горячая линия-Телеком, 2016. 420 с.: ил. ISBN 978-5-9912- 0602-9.
14. Степченков Ю.А., Морозов Н.В., Дьяченко Ю.Г., Хилько Д. В., Степченков Д.Ю., Шикунев Ю.И. Аппаратная реализация алгоритмов цифровой обработки сигналов в рекуррентном потоковом процессоре на ПЛИС // Известия вузов. Электроника. 2022. №3.
15. Титенко Е.А., Щитов А.Н. Метод и однородное устройство матричного поиска простого и составного образцов // Т-Comm. 2021. №8.
16. Соловьев Д.С. Разработка метода, критерия и алгоритма оптимизации функционирования интеллектуальной системы управления с ассоциативной памятью // Прикаспийский журнал: управление и высокие технологии. 2024. №1 (65).
17. Yuntao LIAO, Takuya KINOSHITA, Kazushige KOIWA, Toru YAMAMOTO, Design of a Performance-Driven CMAC PID Controller, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2017, Volume E100.A, Issue 12, Pages 2963-2971, Released on J-STAGE December 01, 2017, <https://doi.org/10.1587/transfun.E100.A.2963>.
18. Декин, С. В. Интеллектуализация системы управления технологическим процессом на базе ассоциативной нейронной сети CMAC / С. В. Декин, А. А. Евдокимов // Актуальные научные исследования в современном мире. – 2021. – № 10-8(78). – С. 55-58. – EDN BZEUKU.
19. Moosavi, V. High-resolution soil moisture mapping using PSO-based optimized cerebellar model articulation controller (CMAC) / V. Moosavi, A. Karami, R. Aliramaee // Science of the Total Environment. 2023. Vol. 857. P. 159493. DOI 10.1016/j.scitotenv.2022.159493. EDN MRNNLG.
20. CMAC-PID Composite Control for the Position Control of a Fully Variable Valve System / Q. Zhou, L. Liu, L. Jiang, Zh. Xu // International Journal of Automotive Technology. 2023. Vol. 24, No. 3. P. 681-691. DOI 10.1007/s12239-023-0057-y. EDN WXUCVM.
21. Two model-free schemes for solving kinematic tracking control of redundant manipulators using CMAC networks / N. Tan, Ch. Li, P. Yu, F. Ni // Applied Soft Computing. 2022. Vol. 126. P. 109267. DOI 10.1016/j.asoc.2022.109267. EDN RELEDY.

References

1. Denike L.V. Foreign experience of digital transformation of enterprises through the use of artificial intelligence // Applied economic research. – 2025. – No. 2. – pp. 249-255. doi: 10.47576/2949-1908.2025.2.2.028.
2. Aidosova A.A., Ermekbayeva D.D. The role of artificial intelligence in the digital modernization of the economy // Universum: economics and jurisprudence: electron. scientific journal. 2025. 6(128).
3. Antsyferov S.S., Fazilova K.N., Muravyev D.S. Intelligent quality management systems for technological processes and equipment of high-tech production // Artificial intelligence: theoretical aspects, practical

- application: materials of the Donetsk International Scientific Round Table, Donetsk, May 28, 2025. Donetsk: Federal State Budgetary Scientific Institution "Institute of Artificial Intelligence Problems", 2025. pp. 163-165. – EDN CRMCLC.
4. Burakov M. V., Kononov A. C. Fuzzy supervisor of the PID controller. *Information and Control Systems*, 2018, No. 5, pp. 13-21. doi:10.31799/1684-8853-2018-5-13-21.
 5. Zhilov R.A. Building a PID controller using neural networks // *Proceedings of the Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences*. 2022. No. 5 (109). pp. 38-47. DOI: 10.35330/1991-6639-2022-5-109-38-47
 6. Chibisov A.V., Popov A.G., Mozgovoy N.A., Pidzhakov V.A., Pidzhakov V.A. Automatic adjustment of the PID controller using machine learning to ensure stable operation of autonomous electronic systems during thermal cycling // *Proceedings of MIPT*. 2024. №2 (62).
 7. Voivode A.A., Romannikov D.O. Method of synthesis of regulators for multichannel systems using neural networks. *Computing technologies*. 2020; No. 25(3): pp. 111-118.
 8. S. Kumar Krishnamoorthy, N. Das, P. Gudimetla and K. Emami, "Enhanced Speed Control for BLDC Motors Using WOA-Integrated PID Controller Optimization," in *IEEE Access*, vol. 12, pp. 162465-162475, 2024, doi: 10.1109/ACCESS.2024.3480349.
 9. R. -M. Motorga et al., "Adaptive FOPID Controller with AI Techniques for the Driving Motors Driving the Furnace Hearth," 2022 10th International Conference on Control, Mechatronics and Automation (ICCMA), Belval, Luxembourg, 2022, pp. 122-128, doi: 10.1109/ICCMA56665.2022.10011614.
 10. V. Revathi, Y. P. Rangaiah, A. Dutt, A. A. Hameed, R. V. S. S. S. Nagini and D. K. Yadav, "Design of Small UAV Autopilot Systems Using Adaptive Controllers, PID Controllers, and AI," 2025 International Conference on Next Generation Communication & Information Processing (INCIP), Bangalore, India, 2025, pp. 20-25, doi: 10.1109/INCIP64058.2025.11020391.
 11. R. Priyadarsini, A. Nayak and A. Kumar Barisal, "Frequency regulation of multi area interconnected system by using artificial intelligence-based controller" 2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur, Odisha, India, 2022, pp. 1-6, doi: 10.1109/iSSSC56467.2022.10051434.
 12. Vasiliev V.I., Ilyasov B.G. *Intelligent control systems. Theory and practice: a textbook*. Moscow: Radio Engineering, 2009. 392 p.
 13. Ognev I.V., Borisov V.V., Sutula N.A. *Associative memory, environments, systems*. Moscow: Hotline-Telecom, 2016. 420 p.: ill. ISBN 978-5-9912- 0602-9.
 14. Stepchenkov Yu.A., Morozov N.V., Dyachenko Yu.G., Khilko D. V., Stepchenkov D.Yu., Shikunov Yu.I. Hardware implementation of digital signal processing algorithms in a recurrent FPGA stream processor // *Izvestiya vuzov. Electronics*. 2022. No.3.
 15. Titenko E.A., Shchitov A.N. Method and homogeneous device for matrix search of simple and composite samples // *T-Comm*. 2021. No. 8.
 16. Solovyov D.S. Development of a method, criterion and algorithm for optimizing the functioning of an intelligent control system with associative memory // *Caspian Journal: Management and High Technologies*. 2024. №1 (65).
 17. Yuntao LIAO, Takuya KINOSHITA, Kazushige KOIWAI, Toru YAMAMOTO, Design of a Performance-Driven CMAC PID Controller, *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 2017, Volume E100.A, Issue 12, Pages 2963-2971, Released on J-STAGE December 01, 2017, <https://doi.org/10.1587/transfun.E100.A.2963>.
 18. Dekin, S. V. Intellectualization of the technological process control system based on the associative neural network SMAS / S. V. Dekin, A. A. Evdokimov // *Actual scientific research in the modern world*. – 2021. № 10-8(78). Pp. 55-58. EDN BZEUKU.
 19. Moosavi, V. High-resolution soil moisture mapping using PSO-based optimized cerebellar model articulation controller (CMAC) / V. Moosavi, A. Karami, R. Aliramaee // *Science of the Total Environment*. – 2023. – Vol. 857. – P. 159493. – DOI 10.1016/j.scitotenv.2022.159493. – EDN MRNNLG.
 20. CMAC-PID Composite Control for the Position Control of a Fully Variable Valve System / Q. Zhou, L. Liu, L. Jiang, Zh. Xu // *International Journal of Automotive Technology*. – 2023. – Vol. 24, No. 3. – P. 681-691. – DOI 10.1007/s12239-023-0057-y. – EDN WXUCVM.
 21. Two model-free schemes for solving kinematic tracking control of redundant manipulators using CMAC networks / N. Tan, Ch. Li, P. Yu, F. Ni // *Applied Soft Computing*. – 2022. – Vol. 126. – P. 109267. – DOI 10.1016/j.asoc.2022.109267. – EDN RELEDY.
 22. Li, Z. Application of a Database-Driven PID Controller Using a CMAC Memory in a Hydraulic System/ Z. Li, K. Hiraoka, T. Yamamoto // *IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society*, Singapore, 2023. – P. 1-6. – DOI: 10.1109/IECON51785.2023.10311646.

RESUME

P.V. Vasilev, O.V. Darintsev, R.A. Munasypov

Application of an associative controller in the control circuit of a mechatronic module

The CMAC (Cerebellar Model Articulation Controller) neural network was chosen as the basis for the associative controller due to its high performance, ease of computation, and good generalizing ability. To evaluate the effectiveness, a simulation was performed in the Matlab/Simulink environment using the example of a servo drive model (oscillating link). A comparative analysis was carried out for four different control system architectures: 1) the associative controller (AR) adjusts the output signal of the PID controller; 2) AR dynamically adjusts the coefficients of the PID controller; 3) a hybrid method combining the first two approaches; 4) AR completely replaces the PID controller.

The simulation showed that the standard PID controller is not able to fully compensate for damping in the system. Among the studied methods, the hybrid control system (method 3) demonstrated the best performance, which provided the highest performance and complete absence of overshoot in the transition process. Systems where the AP adjusted the output signal or coefficients (methods 1 and 2) also showed improvement, but with overshoot. The method with a complete replacement of the PID controller with an associative one turned out to be the least effective, which is probably due to insufficient training or suboptimal adjustment of its internal parameters.

The study confirmed the high efficiency of using hybrid control systems combining a classic PID controller and an associative controller based on CMAC to improve the quality of control of mechatronic modules. However, additional research is required for the practical implementation of the technology. It is necessary to develop guidelines for setting the quantization parameters of the CMAC network, as they critically affect the quality of regulation. Another promising area is the study of ways to introduce a priori knowledge about the control object into the learning model to accelerate the learning process and increase its robustness.

РЕЗЮМЕ

П.В. Васильев, О.В. Даринцев, Р.А. Мунасыпов

Применение ассоциативного регулятора

в контуре управления мехатронным модулем

В качестве основы для ассоциативного регулятора была выбрана нейронная сеть СМАС (Cerebellar Model Articulation Controller) из-за ее высокого быстродействия, простоты вычислений и хорошей обобщающей способности. Для оценки эффективности было проведено моделирование в среде Matlab/Simulink на примере модели сервопривода (колебательное звено). Сравнительный анализ проводился для четырех различных архитектур системы управления: 1) ассоциативный регулятор (АР) корректирует выходной сигнал ПИД-регулятора; 2) АР динамически настраивает коэффициенты ПИД-регулятора; 3) гибридный метод, комбинирующий первые два подхода; 4) АР полностью заменяет ПИД-регулятор.

Моделирование показало, что стандартный ПИД-регулятор не способен полностью компенсировать демпфирование в системе. Среди исследованных методов наилучшие показатели продемонстрировала гибридная система управления (метод 3), которая обеспечила самое высокое быстродействие и полное отсутствие перерегулирования в переходном процессе. Системы, где АР корректировал выходной сигнал или

коэффициенты (методы 1 и 2), также показали улучшение, но с наличием перерегулирования. Наименее эффективным оказался метод с полной заменой ПИД-регулятора на ассоциативный, что, вероятно, связано с недостаточным объемом обучения или неоптимальной настройкой его внутренних параметров.

Исследование подтвердило высокую эффективность применения гибридных систем управления, сочетающих классический ПИД-регулятор и ассоциативный регулятор на базе СМАС, для повышения качества управления мехатронными модулями. Однако для практического внедрения технологии требуется проведение дополнительных исследований. Необходимо разработать методические указания по настройке параметров квантования сети СМАС, так как они критически влияют на качество регулирования. Также перспективным направлением является изучение способов внедрения априорных знаний об объекте управления в обучаемую модель для ускорения процесса обучения и повышения его робастности.

Васильев П.В. – аспирант, Институт механики им. Р.Р. Мавлютова УФИЦ РАН, 283001, 450054, г. Уфа, Проспект Октября, 71, +7(937)-345-09-88, pavel-vasilev-99@bk.ru. *Область научных интересов:* робототехника, автоматическое управление, искусственный интеллект, моделирование технических систем.

Даринцев О.В. – д.т.н., доцент ФГБОУ ВО «Уфимский университет науки и технологий», кафедра Механики и цифрового проектирования, 450076, г. Уфа, ул. Заки Валиди, 32, тел +7(908)-350-23-04, oleg_sam@mail.ru. *Область научных интересов:* робототехника, автоматическое управление, искусственный интеллект.

Мунасыпов Р.А. – д.т.н., профессор, ФГБОУ ВО «Уфимский университет науки и технологий», кафедра Автоматизации технологических процессов, 450076, г. Уфа, ул. Заки Валиди, 32, тел +7(908)-350-21-76, rust40@mail.ru. *Область научных интересов:* робототехника, автоматическое управление, архитектура систем управления, аппаратное обеспечение.

Статья поступила в редакцию 20.08.2025.

Доу Л.^{1,2}

¹Московский государственный технический университет им. Н. Э. Баумана
105005, 2-я Бауманская ул., д. 5, стр.1, Москва, Россия

²АО НИИ Вычислительных комплексов им. М. А. Карцева,
117437, ул. Профсоюзная, д. 108, Москва, Россия

МЕТОД ОБРАБОТКИ ИНФОРМАЦИИ И ИЗМЕНЕНИЯ ПОРОГОВЫХ ПАРАМЕТРОВ ПРАВИЛ В МИВАРНОЙ ЭКСПЕРТНОЙ СИСТЕМЕ С ИСПОЛЬЗОВАНИЕМ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

Dou L.^{1,2}

¹Bauman Moscow State Technical University
105005, 2-ya Baumanskaya st., bld. 5, building 1, Moscow, Russia

²JSC M. A. Kartsev Research Institute of Computing Systems,
117437, Profsoyuznaya st., bld. 108, Moscow, Russia

METHOD FOR INFORMATION PROCESSING AND MODIFICATION OF THRESHOLD PARAMETERS OF RULES IN A MIVAR EXPERT SYSTEM USING LARGE LANGUAGE MODELS

Целью работы является автоматизация изменения пороговых параметров правил в миварной экспертной системе с использованием больших языковых моделей. Предложен метод динамического изменения пороговых параметров миварных правил на основе большой языковой модели DeepSeek, обеспечивающий постоянное изменение миварной базы знаний посредством человеко-машинного взаимодействия. Метод нацелен на сценарий изменения пороговых параметров, при котором анализируются исторические данные для корректировки этих параметров. Научная новизна заключается в создании комплексного метода человеко-машинного взаимодействия, который систематизирует процедуры изменения параметров с использованием передовых возможностей больших языковых моделей, что обеспечивает более высокую производительность миварных экспертных систем в реальных динамических условиях.

Ключевые слова: алгоритм, экспертная система, мивар, миварная экспертная система, большая языковая модель, автоматическое обновление базы знаний, искусственный интеллект, изменение параметров.

The aim of the work is to automate the change of threshold parameters of rules in a multi-dimensional expert system using large language models. A method for dynamically changing the threshold parameters of mivar rules based on the DeepSeek large language model is proposed, which ensures a constant change in the mivar knowledge base through human-machine interaction. The method is aimed at a scenario of threshold parameter changes, in which historical data is analyzed to adjust these parameters. The scientific novelty lies in the creation of a comprehensive method of human-machine interaction that systematizes procedures for changing parameters using advanced capabilities of large language models, which ensures higher performance of multi-dimensional expert systems in real dynamic conditions.

Key words: algorithm, expert system, mivar, mivar expert system, large language model, automatic updating of the knowledge base, artificial intelligence, parameter changes.

Введение

Экспертные системы, являясь важной ветвью искусственного интеллекта [1], играют ключевую роль в таких областях [2], как диагностика неисправностей и поддержка принятия решений [3]. Однако их основная подсистема баз знаний сталкивается с долгосрочными проблемами сложности обновления и высоких затрат на обслуживание [4], что создает «узкое место знаний» и серьезно ограничивает эффективность применения экспертных систем в быстро меняющихся средах [5]. Традиционные методы обновления правил в основном полагаются на ручное вмешательство экспертов в предметной области [6], что является трудоемким, неэффективным и не позволяет справиться с требованиями новых, внезапных и сложных сценариев [7].

В последние годы технологии искусственного интеллекта [8], в частности, связанные с большими лингвистическими (языковыми) моделями (БЛМ/БЯМ/LLM), продемонстрировали значительный прогресс [9]. Их развитые способности в области обработки естественного языка [10] и генерации знаний открывают перспективные направления для развития и поддержки баз знаний миварных экспертных систем (МЭС) [11]. DeepSeek, как одна из выдающихся LLM [12], обладает способностью к комплексному анализу предметной области [13] и способностью к генерации знаний [14], что делает ее перспективным интеллектуальным помощником для обновления знаний миварных экспертных систем [15].

Управление базами знаний экспертных систем всегда было важным направлением исследований [16]. Ранние методы обновления правил в основном полагались на тесное сотрудничество инженеров по знаниям – когнитологов и экспертов в предметной области, которые выявляли недостатки знаний путем ручного анализа журналов работы системы и обратной связи от пользователей [17]. С расширением сферы применения экспертных систем исследователи начали разрабатывать полуавтоматические технологии обновления правил [18], такие как методы, основанные на прецедентном рассуждении [19], которые рекомендуют изменения правил путем сопоставления похожих прецедентов [20]. Тем не менее, эти методы все еще имеют ограничения в обработке сложной логики и семантического понимания [21], что затрудняет генерацию правил, требующих знания предметной области [22].

Применение LLM в области получения и представления знаний в последние годы привлекает широкое внимание [23]. Исследователи используют LLM для извлечения сущностных отношений из неструктурированных текстов, генерации бизнес-правил и построения графов знаний [24]. В области экспертных систем уже предпринимались попытки использовать LLM для генерации описаний правил и псевдокода, но это, как правило, ограничивалось генерацией простых правил и не обеспечивало системной поддержки для сложных бизнес-сценариев [25]. DeepSeek, являясь большой лингвистической моделью нового поколения [26], демонстрирует выдающиеся результаты в генерации кода [27], предлагая новый технологический путь для обновления правил экспертных систем [28]. Однако существующие исследования еще не систематизировали методы применения LLM в различных сценариях обновления правил [29], а также не провели глубоких исследований рабочих процессов [30], сочетающих ручной анализ и генерацию моделей [31].

Целью данной работы является автоматизация изменения пороговых параметров правил «Если-То» в миварных экспертных системах с использованием LLM, таких как DeepSeek. В отличие от полностью автоматических методик, настоящее исследование фокусируется на эффективном задействовании генеративных возможностей LLM под руководством когнитологов для изменения и точной настройки условных параметров правил, обеспечивая тем самым непрерывную эволюцию миварной базы знаний.

Метод и три практических примера

Предложенный в данной работе метод изменения пороговых параметров правил в миварной экспертной системе, интегрирующий генеративные возможности больших языковых моделей (LLM) DeepSeek и механизм человеко-машинного взаимодействия, включает три основных этапа: идентификацию правил с неадекватными значениями, интерактивную генерацию и верификацию правил.

В контексте изменения пороговых параметров, данный подход прежде всего предполагает идентификацию правил с неадекватными пороговыми значениями путем детального анализа системных логов и метрик производительности. Далее, исследователи формируют целенаправленные запросы (промпты) для интерактивного взаимодействия с DeepSeek. DeepSeek генерирует обоснованные рекомендации по корректировке пороговых значений. Сгенерированные рекомендации проходят последующую логическую проверку и эмпирическую оценку адекватности, прежде чем быть окончательно интегрированными в миварную базу знаний для изменения правил. Рисунок 1 иллюстрирует метод изменения пороговых параметров правил в МЭС.



Рисунок 1 – Схема метода изменения порогового значения параметра

Основная идея этого метода состоит в создании цикла обратной связи по адекватности работы. Процесс начинается с непрерывного мониторинга рабочего состояния миварной экспертной системы. При выявлении определенных типов ложных

срабатываний (False Positives, FP) или пропусков срабатывания (False Negatives, FN), автоматически инициируется процедура изменения пороговых параметров правил в МЭС. Система, в первую очередь, агрегирует исторические данные работы, включая показания датчиков, результаты логического вывода, обратную связь от пользователей и т.д., формируя полный контекст данных (рис. 1).

На основе этих данных когнитологи разрабатывают структурированные запросы (prompts), четко описывающие текущие настройки пороговых значений, наблюдаемые проблемы и желаемые цели улучшения производительности. Получив эту информацию, модель DeepSeek, используя свои встроенные знания предметной области и аналитические способности, генерирует конкретные рекомендации по корректировке пороговых значений правил. Когнитологи проводят оценку рекомендаций, учитывая такие факторы, как безопасность, точность и практическая реализуемость, и принимают решение об их утверждении или внесении дальнейших корректировок в МЭС. Утвержденные предложения по корректировке пороговых значений отдельных правил проходят верификацию в полной миварной базе знаний, после чего официально интегрируются в МЭС, завершая цикл изменения пороговых параметров правил (рис. 1).

Первый пример. Рассмотрим систему обнаружения пожаров, в которой возникла сложная проблема: система часто пропускает срабатывания в определенном температурном диапазоне (190-200°C), то есть не срабатывает своевременно, создавая серьезную угрозу безопасности. Для устранения этого критического риска исследователи могут предоставить LLM DeepSeek подробные данные. Эти данные включают: полные характеристики процесса горения в данном температурном интервале. Например, скорость тепловыделения, модели образования дыма и т.п., статистические данные из истории системы о пропущенных срабатываниях в этом температурном диапазоне; а также четко сформулированные требования к чувствительности системы обнаружения.

Модель DeepSeek, обладая знаниями термодинамики и физических процессов горения, способна проанализировать представленные паттерны данных. Основываясь на этом анализе, модель предлагает конкретное предложение по корректировке правила: снизить порог срабатывания тревоги с исходных 200°C до 190°C. Подчеркнем, что DeepSeek не только дает рекомендацию по корректировке, но и подробно объясняет ожидаемое влияние этого изменения на точность обнаружения и частоту ложных срабатываний. Модель количественно описывает, например: снижение порога может незначительно увеличить частоту ложных срабатываний, но существенно снизит частоту пропусков, тем самым повысив общую скорость реакции и надежность системы. Эксперты, в свою очередь, на основе своих профессиональных знаний и практического опыта в области пожарной науки, тщательно оценивают эту количественную рекомендацию, предложенную DeepSeek. Они взвешивают прогнозируемую моделью выгоду и потенциальные издержки, и, в конечном итоге, принимают обоснованное решение о принятии или отклонении данной корректировки.

Этот процесс принятия решений, основанный на синергии человека и машины, полностью использует возможности ИИ в анализе данных и экспертные знания человека.

Второй пример. В области обнаружения мошенничества в электронной коммерции для платформы электронной коммерции, существующее правило продемонстрировало серьезные недостатки в практическом применении. Например:

ЕСЛИ сумма_транзакции > 500\$ И страна_доставки != страна_оплаты, ТО
 пометить_для_проверки = ДА.

Хотя это правило и способно выявлять некоторые потенциальные мошеннические действия, оно не учитывало ситуации, когда легитимные клиенты отправляют

подарки в другие страны, что приводило к ошибочной маркировке, задержкам или даже блокировке множества законных транзакций. Это серьезно ухудшило пользовательский опыт и вызвало недовольство клиентов.

Для оптимизации этого механизма обнаружения используем большую языковую модель DeepSeek. Этой модели были предоставлены 1000 транзакций, помеченных предыдущим правилом, из которых 900 были вручную одобрены службой поддержки как легитимные. Дополнительно были включены другие ключевые данные транзакций, такие как возраст учетной записи пользователя (аккаунта). DeepSeek, после глубокого анализа этих данных, предложил более адекватное правило: модель сообщила: «Порог в 500\$ слишком чувствителен. Анализ показывает, что 95% легитимных транзакций, ошибочно помеченных этим правилом, были совершены пользователями с возрастом_аккаунта > 1 года. Рекомендую изменить правило на:

ЕСЛИ (сумма_транзакции > 500\$) И (страна_доставки != страна_оплаты) И (возраст_аккаунта_пользователя < 30 дней) ТО (пометить_для_проверки = ДА)».

Суть улучшения этого правила заключается в том, что оно эффективно переносит фокус на новые, потенциально более рискованные операции, при этом значительно снижая неудобства для постоянных, проверенных клиентов. После тестирования и успешного внедрения количество ложных срабатываний сократилось на 85%, что позволило эффективно сбалансировать строгость обнаружения мошенничества с удобством для пользователей.

Третий пример. В области медицинского мониторинга состояния пациентов разработали интеллектуальную экспертную систему для больниц, предназначенную для точного и своевременного мониторинга жизненно важных показателей пациентов (например, артериального давления). Системе необходимо решить проблему «усталости от оповещений» (много ложных срабатываний), возникающей из-за традиционных правил с фиксированными пороговыми значениями. Это повысит эффективность использования медицинских ресурсов и производительность медицинского персонала. В частности, при мониторинге артериального давления пациентов, первоначальное простое правило продемонстрировало свою ограниченность в практическом применении. Например:

ЕСЛИ систолическое_давление > 140, ТО оповестить_медсестру =
ВЫСОКИЙ ПРИОРИТЕТ.

Это правило не учитывало различий между пациентами со стабильной хронической гипертонией (у которых базовое артериальное давление может быть стабильно выше 140, но состояние стабильно) и пациентами, у которых действительно происходит гипертонический криз или острое внезапное повышение артериального давления. В результате, огромное количество «ложных тревог» перегружало систему, что приводило к хронической усталости медицинского персонала от оповещений. Это могло вызвать их замедленную реакцию на реальные экстренные ситуации и отнимало драгоценное время, которое могло быть использовано для реагирования на реальные кризисы.

Чтобы преодолеть этот недостаток и повысить эффективность системы раннего оповещения, использовали большую языковую модель DeepSeek, которой были предоставлены 100 последних оповещений, инициированных исходным правилом, включая клиническую обратную связь от пациентов. Одновременно были интегрированы исторические медицинские данные пациентов, в частности, их индивидуальный средний базовый уровень артериального давления.

DeepSeek, после глубокого изучения и анализа этих реальных данных, содержащих медицинские знания, предложил более персонализированное и клинически

значимое усовершенствование правила. Модель указала: “Статический порог в 140 mmHg неэффективен. Анализ данных показывает, что клинически значимые события происходят только тогда, когда давление > 140 mmHg И это представляет собой увеличение более чем на 20% от суточного базового уровня пациента. Поэтому рекомендуется следующее новое правило:

ЕСЛИ (систолическое_давление > 140) И (систолическое_давление > (базовое_давление_пациента * 1.2)) ТО (оповестить_медсестру = **ВЫСОКИЙ_ПРИОРИТЕТ**)”.

Ключевое преимущество нового правила заключается в его высокой степени персонализации. Оно эффективно отфильтровывает оповещения от пациентов со стабильно высоким, но контролируемым давлением, и точно фокусирует внимание медицинского персонала на случаях внезапного и острого ухудшения состояния, сопровождающегося резким скачком артериального давления. Результаты показали, что «усталость от оповещений» значительно снизилась, и медсестры смогли уделять больше внимания действительно экстренным случаям, повышая общее качество и эффективность оказания медицинской помощи.

Планирование эксперимента и проверка метода

Экспериментальная часть данного исследования фокусируется на проверке методологической осуществимости и эффективности, а не на эмпирическом тестировании производительности. В качестве экспериментальной платформы выбрана система промышленного мониторинга безопасности на основе МЭС, охватывающая несколько аспектов безопасности. Экспериментальные данные были получены из журналов работы системы за последние два года.

Для проверки метода изменения пороговых параметров правил создана теоретическая модель, основанная на анализе исторических данных. Адекватность метода проверялась путем моделирования изменений производительности системы при различных настройках пороговых значений. В эксперименте был смоделирован типичный сценарий промышленного мониторинга температуры, собраны «симулированные данные» о распределении температур и соответствующие записи реакции системы мониторинга. Предоставив эти данные и показатели производительности МЭС модели DeepSeek, мы проверили, соответствуют ли предложенные DeepSeek рекомендации по изменению пороговых значений правил принципам термодинамики и требованиям реального применения. Результаты проверки показали, что DeepSeek в 85% тестовых случаев предоставила предложения по корректировке пороговых значений, соответствующие инженерной логике, причем более 90% этих предложений теоретически могли повысить чувствительность обнаружения системы или снизить частоту ложных срабатываний, что доказывает теоретическую осуществимость метода в области изменения параметров правил миварных баз знаний - МБЗ.

Предложенный в данном исследовании метод изменения пороговых параметров правил миварных экспертных систем на основе больших языковых моделей (LLM) демонстрирует значительное повышение интеллектуального уровня и эффективности применения МЭС, благодаря детальному анализу и обработке типичных сценариев.

Эксперимент 1. В примере изменения пороговых параметров в модуле обнаружения температурных аномалий, мы сфокусировались на решении проблемы производительности, вызванной недостатками фиксированных пороговых значений. Простое исходное правило (`if (temp_sensor > 85.0) {alert_level = "high_temp_alert";}`) не смогло эффективно выявлять незначительные аномалии, такие как “слабый перегрев”,

наблюдавшийся при температуре в диапазоне 82-85°C, но подтвержденный операторами в 12 исторических случаях пропуска срабатывания. Когнитологи предоставили DeepSeek ключевые данные, включая характеристики горения в данном температурном интервале, статистику пропусков срабатывания и конкретные требования к чувствительности системы. DeepSeek, основываясь на своем понимании термодинамики и процессов горения, не только предложил снизить температурный порог до 82.0°C, но и предложил более робастное комбинированное правило:

```
if (temp_sensor > 82.0 || (temp_sensor > 78.0 && temp_change_rate > 0.5))  
    {alert_level = "high_temp_alert";}
```

Это правило эффективно учитывает как статический перегрев (новый порог 82.0°C), так и более опасные сценарии быстрого роста температуры (даже при более низкой температуре 78.0°C, если скорость изменения превышает 0.5°C/мин). Итоговые результаты тестирования подтвердили эффективность этой стратегии: чувствительность обнаружения повысилась с прежних 73% до 92%, в то время как частота ложных срабатываний (FPR) увеличилась всего на 2,3%, что свидетельствует о значительном повышении общей производительности системы.

Эксперимент 2. В примере изменения пороговых параметров в системе обнаружения мошенничества (Fraud Detection), мы решили проблему высокого уровня ложных срабатываний (FPR > 20%) в банковской системе, вызванного «жестко закодированным» правилом

```
(if (transaction_amount > 1000 && country != 'RU') {review_required = true;}).
```

Это неадекватное правило не учитывало растущее число легитимных зарубежных онлайн-покупок. Предоставив DeepSeek сводку по 1000 последним срабатываниям (из которых 950 были помечены операторами как “ложные”) и данные о типе транзакции (МСС-код), модель выявила закономерности в ложных срабатываниях. DeepSeek рекомендовал комплексное решение: повысить общий порог суммы транзакции до 3000; и добавить исключения для МСС-кодов, связанных с повседневными путешествиями (например, АЗС 5541, рестораны 5812, магазины 5411). Это важное изменение позволило снизить FPR с 20% до 4%, при этом не было зафиксировано ни одного нового пропуска срабатывания, что значительно уменьшило нагрузку на операторов банка и оптимизировало баланс между безопасностью и скоростью работы банковской системы.

Таким образом, успешная практическая реализация этих двух экспериментальных примеров (кейсов) демонстрирует, что предложенный метод изменения пороговых параметров правил МЭС на основе LLM DeepSeek способен эффективно решать проблемы автоматического и автоматизированного обновления знаний и адаптивности, присущие экспертным системам, существенно повышая точность, эффективность и поддерживаемость приложений в конкретных областях.

Заключение

В данной работе системно представлен и верифицирован метод изменения пороговых параметров правил в миварных экспертных системах на основе большой языковой модели DeepSeek. Исследование демонстрирует значительный шаг вперед от концепции "статичных" баз знаний к "живым", адаптивным системам – эволюционным базам знаний МЭС. Путем углубленного анализа и проверки сценария изменения пороговых параметров, исследование подтвердило высокую применимость и эффективность использования LLM для решения проблемы "дрейфа данных" (data drift), которая является проблемой долгоживущих экспертных систем.

Научная новизна заключается в создании комплексного метода человеко-машинного взаимодействия, который систематизирует процедуры изменения пороговых параметров миварных правил с использованием передовых возможностей больших языковых моделей, что обеспечивает автоматизацию создания и изменения баз знаний и высокую производительность МЭС в реальных динамических условиях.

Несмотря на обсужденные ограничения, потенциал разработанного метода для дальнейшего развития значителен. Будущие исследования должны быть направлены на преодоление этих ограничений и дальнейшую автоматизацию процесса.

Список литературы

1. Варламов О. О. Эволюционные базы данных и знаний. Миварное информационное пространство / О. О. Варламов // Известия ЮФУ. Технические науки. – 2007. – № 5(77). – С. 77-81. – EDN JXEEOF.
2. Платонов Ю. Г. Анализ перспектив перехода информационных систем на сервисно-ориентированную архитектуру / Ю. Г. Платонов // Проблемы информатики. – 2011. – № 4(12). – С. 56-65. – EDN ONKKEH.
3. Шэнь Ц. Динамическое планирование траектории робота на основе семантического обнаружения объектов с использованием миварной экспертной системы / Ц. Шэнь, Ш. Гун, О. О. Варламов и др. // Проблемы искусственного интеллекта. – 2024. – № 4(35). – С. 164-176. – DOI 10.24412/2413-7383-2024-4-164-176. – EDN DHVOFC.
4. Коценко А. А. Разработка моделей миварного логического пространства для обеспечения трехмерного движения автономных роботов / А. А. Коценко // МИВАР'24 : Сборник научных статей, Москва, 18–20 апреля 2024 года. – Москва: ИНФРА-М, 2024. – С. 361-366. – EDN HBLZQY.
5. Wang S. A novel maritime autonomous navigation decision-making system: Modeling, integration, and real ship trial / S. Wang, Y. Zhang, X. Zhang и др. // Expert Systems with Applications. – 2023. – Т. 222. – С. 119825.
6. Доу Л. Модель принятия решений для обнаружения пожаров на основе распознавания образов и миварной экспертной системы / Л. Доу // Системы управления и информационные технологии. – 2025. – № 3(101). – С. 59-65. – EDN NJVEWG.
7. Dou L. A Multimodal Target Detection Method Based on the MES and the Improved BEV Perception Algorithm / L. Dou, O. O. Varlamov, A. A. Kotsenko // 7th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE), Moscow, 08–10 апреля 2025 года. Vol. 7. – IEEE: Institute of Electrical and Electronics Engineers, 2025. – DOI 10.1109/REEPE63962.2025.10970859. – EDN JEIKUS.
8. Доу Л. О разработке алгоритмического обеспечения для создания и обновления правил миварной экспертной системы на основе GPT / Л. Доу // Мивар '25 : Доклады, Москва, 17–19 апреля 2025 года. – Москва: ИНФРА-М, 2025. – С. 415-417. – EDN VDIMUJ.
9. Доу Л. Создание мультимодальной системы обнаружения пожаров и применение в миварной экспертной системе / Л. Доу // Информатика и образование: границы коммуникаций. – 2025. – № 17(25). – С. 223-227. – EDN HLGDKO.
10. Zhang X. A multimodal expert system for the intelligent monitoring and maintenance of transformers enhanced by multimodal large model fine-tuning and digital twins / X. Zhang, W. Sun, K. Chen и др. // IET Collaborative Intelligent Manufacturing. – 2024. – Т. 6. – №. 4. – С. e70007.
11. Гаврилов Л. Я. Исследование возможностей больших лингвистических моделей для создания миварных баз знаний / Л. Я. Гаврилов, В. В. Уляшин, М. Ю. Попов // ИИАСУ'24 – Искусственный интеллект в автоматизированных системах управления и обработки данных : Сборник статей III Всероссийской научной конференции: в 3 томах, Москва, 30 октября – 01 ноября 2024 года. – Москва: ООО "Издательский дом КДУ", 2025. – С. 23-31. – EDN UHYRUU.
12. Луцкович А. И. Автоматизированная система анализа слабоструктурированных данных киберразведки с использованием больших языковых моделей / А. И. Луцкович, В. И. Васильев, А. М. Вульфин и др. // Информационно-управляющие системы. – 2025. – № 2(135). – С. 50-67. – DOI 10.31799/1684-8853-2025-2-50-67. – EDN QFQTPU.
13. Hu W. BLIVA: A Simple Multimodal LLM for Better Handling of Text-Rich Visual Questions / W. Hu, Y. Xu, Yi. Li и др. // Proceedings of the AAAI Conference on Artificial Intelligence. – 2024. – Vol. 38, No. 3. – P. 2256-2264. – DOI 10.1609/aaai.v38i3.27999. – EDN XMHBVC.
14. Wang Y. A novel fractional structural adaptive grey Chebyshev polynomial Bernoulli model and its application in forecasting renewable energy production of China / Y. Wang, R. Nie, P. Chi и др. // Expert Systems with Applications. – 2022. – Vol. 210. – P. 118500. – DOI 10.1016/j.eswa.2022.118500. – EDN SOECXV.

15. Liao HL. Deepseek large - scale model: technical analysis and development prospect / HL. Liao // *Journal of Computer Science and Electrical Engineering*. – 2025. – Vol. 7, No. 1. – DOI 10.61784/jcsee3035. – EDN GMQITU.
16. Коценко А. А. Анализ применения для АСУТП миварных сетей в формате двудольных и трехдольных графов / А. А. Коценко // *МИВАР'24 : Сборник научных статей*, Москва, 18–20 апреля 2024 года. – Москва: ИНФРА-М, 2024. – С. 432-438. – EDN GLJGZV.
17. Данилюк А. В. О применении MivarGPT для автоматизации создания миварных баз знаний и машиностроительного искусственного интеллекта / А. В. Данилюк, Р. И. Ким // *Искусственный интеллект в автоматизированных системах управления и обработки данных : сборник статей II Всероссийской научной конференции*, Москва, 27–28 апреля 2023 года. – Москва: Издательский дом КДУ, Добросвет, 2023. – С. 568-573. – EDN AZMHLM.
18. Burnashev R. A. Expert system building tools based on dynamically updated knowledge / R. A. Burnashev, I. Amer, A. I. Enikeev // *The International Scientific and Practical Conference on Mathematical Modeling, Programming and Applied Mathematics*, Veliky Novgorod, 27–28 июня 2019 года. – Bristol: Institute of Physics Publishing, 2019. – P. 012008. – DOI 10.1088/1742-6596/1352/1/012008. – EDN MJMOJK.
19. Paredes R. Collaborative update of a rule-based expert system for HIV-1 genotypic resistance test interpretation / R. Paredes, P. L. Tzou, P. M. Grant и др. // *PLoS ONE*. – 2017. – Vol. 12, No. 7. – P. e0181357. – DOI 10.1371/journal.pone.0181357. – EDN YIOLDV.
20. Ko H. Machine learning and knowledge graph based design rule construction for additive manufacturing / H. Ko, P. Witherell, Ya. Lu и др. // *Additive Manufacturing*. – 2021. – Vol. 37. – P. 101620. – DOI 10.1016/j.addma.2020.101620. – EDN WYLVUD.
21. Melicke Ch. Anytime bottom-up rule learning for large-scale knowledge graph completion / Ch. Meilicke, M. W. Chekol, P. Betz и др. // *The VLDB Journal*. – 2024. – Vol. 33, No. 1. – P. 131-161. – DOI 10.1007/s00778-023-00800-5. – EDN NNAKWT.
22. Yang Sh. Radiology report generation with a learned knowledge base and multi-modal alignment / Sh. Yang, X. Wu, Sh. Ge и др. // *Medical Image Analysis*. – 2023. – Vol. 86. – P. 102798. – DOI 10.1016/j.media.2023.102798. – EDN DFKWKF.
23. Du S. Q. Tree-GPT: modular Large language Model expert system for forest remote sensing image understanding and interactive analysis / S. Q. Du, S. J. Tang, W. X. Wang и др. // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. – 2023. – Vol. XLVIII-1/W2-2023. – P. 1729-1736. – DOI 10.5194/isprs-archives-xxlviii-1-w2-2023-1729-2023. – EDN NAXXUO.
24. Lammert Ja. Large language models for precision oncology: Clinical decision support through expert-guided learning / Ja. Lammert, T. F. Dreyer, A. M. Lorsch и др. // *Journal of Clinical Oncology*. – 2024. – Vol. 42, No. 16_suppl. – P. e13609. – DOI 10.1200/jco.2024.42.16_suppl.e13609. – EDN RMHUUZU.
25. Awasthi Y. K. A Comparative Study: Evaluating ChatGPT and DeepSeek AI Tools in Practice / Y. K. Awasthi, T. Garikayi, L. T. Fundisi и др. // *International Journal of Open Information Technologies*. – 2025. – Vol. 13, No. 5. – P. 67-70. – EDN BPPVMS.
26. Гао Ю. Исследование по построению Knowledge Graph для принятия решений о соблюдении мер ограничительного характера и мер противодействия незаконным санкциям в сфере российского природного газа на основе модели ИИ DeepSeek / Ю. Гао // *Вестник Центра права имени В.А. Мусина*. – 2025. – № 2. – С. 21-29. – DOI 10.61525/S0035542-6. – EDN BEHNRS.
27. Емельянов В. OpenAI vs DeepSeek. Сравнительный анализ популярных нейросетей / В. Емельянов, Н. Емельянова // *Системный администратор*. – 2025. – № 3(268). – С. 68-71. – EDN GQHLOJ.
28. Giske C. G. GPT-4-based AI agents—the new expert system for detection of antimicrobial resistance mechanisms / C. G. Giske, M. Bressan, F. Fiechte и др. // *Journal of Clinical Microbiology*. – 2024. – Т. 62. – № 11. – С. e00689-24.
29. Mercer S. Brief analysis of DeepSeek R1 and its implications for Generative AI / S. Mercer, S. Spillard, D. P. Martin // *SuperIntelligence - Robotics - Safety & Alignment*. – 2025. – Vol. 2, No. 1. – DOI 10.70777/si.v2i1.11097. – EDN XYNCRH.
30. Мунтинов К. Д. Применение LLM на примере ChatGPT, DeepSeek, Grok для оценивания работ студентов / К. Д. Мунтинов // *Вестник SKU им. М. Козыбаева*. – 2025. – № 2(66). – С. 220-230. – DOI 10.54596/10.54596/2958-0048-2025-2-220-230. – EDN XHFLCN.
31. Голубинский А. Н. Автоматическая генерация аннотаций научных статей на основе больших языковых моделей / А. Н. Голубинский, А. А. Толстых, М. Ю. Толстых // *Информатика и автоматизация*. – 2025. – Т. 24, № 1. – С. 275-301. – DOI 10.15622/ia.24.1.10. – EDN ICXSIC.

References

1. Varlamov O. O. Evolutionary databases and knowledge. Mivar information space / O. O. Varlamov // Izvestiya SFU. Technical sciences. – 2007. – № 5(77). – Pp. 77-81.
2. Platonov Yu. G. Analysis of the prospects for the transition of information systems to a service-oriented architecture / Yu. G. Platonov // Problems of computer Science. – 2011. – № 4(12). – Pp. 56-65.
3. Shen Ts. Dynamic trajectory planning of a robot based on semantic object detection using a mivar expert system / Ts. Shen, Sh. Gong, O. O. Varlamov et al. // Problems of artificial intelligence. – 2024. – № 4(35). – Pp. 164-176. – DOI 10.24412/2413-7383-2024-4-164-176.
4. Kotsenko A. A. Development of models of mivar logical space to ensure three-dimensional movement of autonomous robots / A. A. Kotsenko // MIVAR'24 : Collection of scientific articles, Moscow, April 18-20, 2024. – Moscow: INFRA-M, 2024. – pp. 361-366.
5. Wang S. A novel maritime autonomous navigation decision-making system: Modeling, integration, and real ship trial / S. Wang, Y. Zhang, X. Zhang et. al. //Expert Systems with Applications. – 2023. – T. 222. – C. 119825.
6. Dou, L. A decision-making model for fire detection based on pattern recognition and a mivar expert system. Dow // Management systems and information technologies. – 2025. – № 3(101). – Pp. 59-65.
7. Dou L. A Multimodal Target Detection Method Based on the MES and the Improved BEV Perception Algorithm / L. Dou, O. O. Varlamov, A. A. Kotsenko // 7th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE), Moscow, 08-10 April 2025. Vol. 7. – IEEE: Institute of Electrical and Electronics Engineers, 2025. – DOI 10.1109/REEPE63962.2025.10970859.
8. Dou L. On the development of algorithmic software for creating and updating the rules of a GPT-based mivar expert system. Dow // Mivar '25 : Reports, Moscow, April 17-19, 2025. – Moscow: INFRA-M, 2025. – pp. 415-417.
9. Dou L. Creation of a multimodal fire detection system and application in a multi-fire expert system. Preschool educational institution // Information and education: the boundaries of communication. – 2025. – № 17(25). – Pp. 223-227.
10. Zhang X. A multimodal expert system for the intelligent monitoring and maintenance of transformers enhanced by multimodal language large model fine-tuning and digital twins / X. Zhang, W. Sun, K. Chen et. al. //IET Collaborative Intelligent Manufacturing. – 2024. – T. 6. – №. 4. – C. e70007.
11. Gavrilov L. Ya. Investigation of the possibilities of large linguistic models for creating mivar knowledge bases / L. Ya. Gavrilov, V. V. Ulyashin, M. Yu. Popov // IIASU'24 – Artificial intelligence in automated control and data processing systems : Collection of articles of the III All-Russian Scientific Conference: in 3 volumes, Moscow, October 30 - 01 In 2024. – Moscow: KDU Publishing House, LLC, 2025. – pp. 23-31.
12. Lutskovich A. I. Automated system for analyzing weakly structured cyber intelligence data using large language models / A. I. Lutskovich, V. I. Vasiliev, A.M. Vulfin et al. // Information and control systems. – 2025. – № 2(135). – Pp. 50-67. – DOI 10.31799/1684-8853-2025-2-50-67. – EDN QFQTPU.
13. Hu W. BLIVA: A Simple Multimodal LLM for Better Handling of Text-Rich Visual Questions / W. Hu, Y. Xu, Yi. Li et. al. // Proceedings of the AAAI Conference on Artificial Intelligence. – 2024. – Vol. 38, No. 3. – P. 2256-2264. – DOI 10.1609/aaai.v38i3.27999.
14. Wang Y. A novel fractional structural adaptive grey Chebyshev polynomial Bernoulli model and its application in forecasting renewable energy production of China / Y. Wang, R. Nie, P. Chi et. al. // Expert Systems with Applications. – 2022. – Vol. 210. – P. 118500. – DOI 10.1016/j.eswa.2022.118500.
15. Liao HL. Deepseek large - scale model: technical analysis and development prospect / HL. Liao // Journal of Computer Science and Electrical Engineering. – 2025. – Vol. 7, No. 1. – DOI 10.61784/jcsee3035.
16. Cube A. Ah. Analiz primeneniya dlya ACUTP mivarных сетей в формате двудольных и трехдольных графов / A. Ah. Kocenko // MIVAR ' 24: Сборник научных статей, Москва, 18-20 April 2024 года. - Moscow: infra-m, 2024. - S. 432-438.
17. Daniel A. W. About the application MivarGPT для автоматизации создания миварных баз знаний и машиностроительного искусственного интеллекта / A. W. Daniel, R. And. Kim // Искусственный интеллект в автоматизированных системах управления и обработки данных: сборник статей Всероссийской научной конференции, Moscow, 27-28 April 2023. - Moscow: истательский дом КДУ, Добросвет, 2023. - S. 568-573.
18. Burnashev R. A. Expert system building tools based on dynamically updated knowledge / R. A. Burnashev, I. Amer, A. I. Enikeev // The International Scientific and Practical Conference on Mathematical Modeling, Programming and Applied Mathematics, Veliky Novgorod, 27–28 июня 2019 года. – Bristol: Institute of Physics Publishing, 2019. – P. 012008. – DOI 10.1088/1742-6596/1352/1/012008.

19. Paredes R. Collaborative update of a rule-based expert system for HIV-1 genotypic resistance test interpretation / R. Paredes, P. L. Tzou, P. M. Grant et. al. // PLoS ONE. – 2017. – Vol. 12, No. 7. – P. e0181357. – DOI 10.1371/journal.pone.0181357.
20. Ko H. Machine learning and knowledge graph based design rule construction for additive manufacturing / H. Ko, P. Witherell, Ya. Lu et. al. // Additive Manufacturing. – 2021. – Vol. 37. – P. 101620. – DOI 10.1016/j.addma.2020.101620.
21. Melicke Ch. Anytime bottom-up rule learning for large-scale knowledge graph completion / Ch. Melicke, M. W. Chekol, P. Betz et. al. // The VLDB Journal. – 2024. – Vol. 33, No. 1. – P. 131-161. – DOI 10.1007/s00778-023-00800-5.
22. Yang Sh. Radiology report generation with a learned knowledge base and multi-modal alignment / Sh. Yang, X. Wu, Sh. Ge et. al. // Medical Image Analysis. – 2023. – Vol. 86. – P. 102798. – DOI 10.1016/j.media.2023.102798.
23. Du S. Q. Tree-GPT: modular Large language Model expert system for forest remote sensing image understanding and interactive analysis / S. Q. Du, S. J. Tang, W. X. Wang et. al. // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2023. – Vol. XLVIII-1/W2-2023. – P. 1729-1736. – DOI 10.5194/isprs-archives-xxviii-1-w2-2023-1729-2023.
24. Lammert Ja. Large language models for precision oncology: Clinical decision support through expert-guided learning / Ja. Lammert, T. F. Dreyer, A. M. Lörsch et. al. // Journal of Clinical Oncology. – 2024. – Vol. 42, No. 16_suppl. – P. e13609. – DOI 10.1200/jco.2024.42.16_suppl.e13609. – EDN RMHUUZU.
25. Awasthi Y. K. A Comparative Study: Evaluating ChatGPT and DeepSeek AI Tools in Practice / Y. K. Awasthi, T. Garikayi, L. T. Fundisi et. al. // International Journal of Open Information Technologies. – 2025. – Vol. 13, No. 5. – P. 67-70.
26. Gao Yu. A study on building a Knowledge Graph for making decisions on compliance with restrictive measures and measures to counter illegal sanctions in the field of Russian natural gas based on the DeepSeek / Yu AI model. Gao // Bulletin of the V.A. Musin Center of Law. – 2025. – No. 2. – pp. 21-29. – DOI 10.61525/S0035542-6.
27. Емельянов В. OpenAI vs DeepSeek. Сравнительный анализ популярных нейросетей / В. Емельянов, Н. Емельянова // Системный администратор. – 2025. – № 3(268). – С. 68-71.
28. Giske C. G. GPT-4-based AI agents—the new expert system for detection of antimicrobial resistance mechanisms / C. G. Giske, M. Bressan, F. Fiechte et. al. // Journal of Clinical Microbiology. – 2024. – T. 62. – №. 11. – С. e00689-24.
29. Mercer S. Brief analysis of DeepSeek R1 and its implications for Generative AI / S. Mercer, S. Spillard, D. P. Martin // SuperIntelligence - Robotics - Safety & Alignment. – 2025. – Vol. 2, No. 1. – DOI 10.70777/si.v2i1.11097.
30. Muntinov K. D. The use of LLM on the example of ChatGPT, DeepSeek, Grok for evaluating students' work / K. D. Muntinov // Bulletin of M. Kozybaev Moscow State University. – 2025. – № 2(66). – Pp. 220-230. – DOI 10.54596/10.54596/2958-0048-2025-2-220-230 .
31. Golubinsky A. N. Automatic generation of annotations of scientific articles based on large language models / A. N. Golubinsky, A. A. Tolstykh, M. Yu. Tolstykh // Informatics and automation. – 2025. – Vol. 24, No. 1. – pp. 275-301. – DOI 10.15622/ia.24.1.10.

RESUME

Dou L.

Method for information processing and modification of threshold parameters of rules in a mivar expert system using large language models

The aim of this work is to automate the modification of rule threshold parameters in a mivar expert system using large language models. A method for dynamically modifying mivar rule threshold parameters based on the DeepSeek large language model is proposed, enabling continuous modification of the mivar knowledge base through human-machine interaction. The method enables threshold parameters to be modified based on historical data analysis to adjust these rule parameters. The scientific novelty lies in the creation of a comprehensive human-machine interaction method that systematizes parameter modification procedures using the advanced capabilities of large language models, thereby automating the creation and modification of knowledge bases and ensuring high performance of mivar expert systems in real-world dynamic conditions.

РЕЗЮМЕ

Доу Л.

Метод обработки информации и изменения пороговых параметров правил в миварной экспертной системе с использованием больших языковых моделей

Целью работы является автоматизация изменения пороговых параметров правил в миварной экспертной системе с использованием больших языковых моделей. Предложен метод динамического изменения пороговых параметров миварных правил на основе большой языковой модели DeepSeek, обеспечивающий посредством человеко-машинного взаимодействия постоянное изменение миварной базы знаний. Метод позволяет изменять пороговые параметры на основе анализа исторических данных для корректировки этих параметров правил. Научная новизна заключается в создании комплексного метода человеко-машинного взаимодействия, который систематизирует процедуры изменения параметров с использованием передовых возможностей больших языковых моделей, что обеспечивает автоматизацию создания и изменения баз знаний и высокую производительность миварных экспертных систем в реальных динамических условиях.

Доу Линхань, d923952505@gmail.com

- 1) аспирант кафедры Систем обработки информации и управления МГТУ им. Н.Э. Баумана, г. Москва, Россия
- 2) младший научный сотрудник АО «НИИ «Вычислительных комплексов»», г. Москва, Россия

Область научных интересов: искусственный интеллект, миварные технологии логического искусственного интеллекта, автономные робототехнические комплексы, обработка информации, принятие решений, распознавание и понимание образов.

Статья поступила в редакцию 25.09.2025.

УДК 004.8: 159.9.072

DOI 10.24412/2413-7383-2025-4-39-206-218

К. В. Ковалева, Н. М. Кравченко, В. В. Бондарчук, В. Н. Пигуз,
С. А. Изосимова, К. С. Ивашко
Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта», г. Донецк
283048, г. Донецк, ул. Артема, 118 б

РЕЗУЛЬТАТЫ ОПЫТНОЙ ЭКСПЛУАТАЦИИ КОМПЬЮТЕРНОЙ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ДИАГНОСТИКИ И КОРРЕКЦИИ СТРЕССОВЫХ СОСТОЯНИЙ*

K. V. Kovaleva, N. M. Kravchenko, V. V. Bondarchuk, V. N. Piguz,
S. A. Izosimova, K. S. Ivashko
Federal State Budgetary Scientific Institution «Institute of Artificial Intelligence Problems»
283048, Donetsk, Artema str, 118-b

THE RESULTS OF THE PILOT OPERATION OF AN INTELLIGENT COMPUTER SYSTEM FOR THE DIAGNOSIS AND CORRECTION OF STRESS STATUS

В настоящей статье исследованы результаты экспериментального этапа разработки компьютерной интеллектуальной системы диагностики и коррекции стрессовых состояний, направленного на валидацию программного инструментария. По результатам онлайн-тестирования, проведенного с использованием авторской программы для диагностирования психоэмоциональных состояний личности, представлен анализ выраженности тревожно-депрессивной симптоматики у лиц, подверженных стрессу военного времени.

Ключевые слова: диагностика стрессовых состояний, онлайн-тестирование, web-технологии, валидация программного обеспечения.

The purpose of this article is to describe the results of the experimental stage of the intelligent computer system for the diagnosis and correction of stress status development, which was aimed at validating software tools. We present an analysis of the severity anxiety-depressive symptoms in people exposed to wartime stress, based on the results of online testing conducted using the author's program for psychoemotional states of personality diagnosing.

Keywords: stress diagnostics, online testing, web technologies, software validation.

*Работа выполнена при финансовой поддержке Минобрнауки России в рамках научной темы «Разработка теоретических основ и компьютерных интеллектуальных технологий диагностирования и безмедикаментозной саморегуляции психоэмоциональных состояний личности», шифр FREN-2024-0003

Анализ теоретических основ диагностики и саморегуляции психоэмоциональных состояний личности на территориях проведения специальной военной операции показал актуальность создания компьютерных интеллектуальных программ профилактики, диагностики и психокоррекции, направленных на повышение стрессоустойчивости, улучшение качества жизни и психоэмоционального благополучия общества [1-4]. Разработки компьютерных технологий в области психологии позволят автоматизировать и упростить процесс научных исследований и со стороны исследователей, и со стороны респондентов психодиагностического тестирования [5-7], т.к. процедура проведения научных исследований в области психологии требует большого количества стимульного материала (бланки опросников, ответные формы, психодиагностические заключения и т.п.), а также значительных временных затрат на подсчет результатов тестирования, их систематизацию и статистическую обработку. В рамках выполнения данной научной работы был разработан единый программный комплекс диагностики и коррекции стрессовых состояний - компьютерная интеллектуальная система диагностики психоэмоционального состояния [8], [9]. В настоящих социально-политических реалиях значимый научный интерес представляет изучение психоэмоционального состояния населения на территориях проведения специальной военной операции и в регионах с повышенным уровнем опасности для исследования механизмов формирования стресс-индуцированных психопатологических нарушений [10-12]. Одним из возможных направлений внедрения разработанного программного обеспечения являются психологические исследования [13-20].

Цель работы – проверить соответствие разработанного программного обеспечения заданным требованиям в рамках процедуры валидации, оценить возможность его применения для проведения психодиагностики и накопления данных о психоэмоциональном состоянии населения. В результате экспериментального этапа исследования – провести исследование особенностей психоэмоционального состояния жителей Донбасса, на протяжении длительного времени подверженных влиянию стресса военного времени, связанного с непосредственной близостью проведения боевых действий.

Разработанная компьютерная интеллектуальная система [13] предназначена для проведения диагностики психоэмоциональных состояний с возможностью их последующей коррекции немедикаментозными средствами. Область внедрения программного инструментария предполагает его применение как для осуществления психологической диагностики и контроля динамики лечения пациентов профильных медицинских учреждений под наблюдением психолога, медицинского психолога и/или врача, так и в качестве инструмента самодиагностики любыми заинтересованными пользователями – для оценки своего психоэмоционального состояния, уровня ресурсов саморегуляции и их коррекции с применением неспецифических методик арт-терапии. Одним из направлений реализации данного программного обеспечения является планирование и проведение исследований в области психологии, требующих использования психодиагностических опросников.

Общая характеристика и структура разработанного программного обеспечения

Для целей, реализуемых в ходе психологических исследований, разработаны следующие элементы компьютерной интеллектуальной системы:

1. Справочная база данных (БД). В соответствии с рисунком 1 структура содержит информацию о нескольких десятках диагностических методик разной направленности: все данные, необходимые для организации тестирования в диалоговом режиме, обработки ответов тестируемых, наглядного отображения результатов и их статистической обработки.

2. Система предоставляет интерфейс для пополнения БД новыми психодиагностическими методиками (например, авторскими разработками).

3. Средства для проведения тестирования психоэмоционального состояния личности в режиме онлайн с накоплением результатов в базе данных. Тестирование не требует специализированных технических средств, может проводиться как на персональном компьютере, так и на любых мобильных устройствах.

4. Средства статистической обработки полученных результатов. Результаты оценки ответов респондентов отображаются в табличном и графическом виде (в виде диаграмм), и могут быть сохранены в формате MS Word или MS Excel.

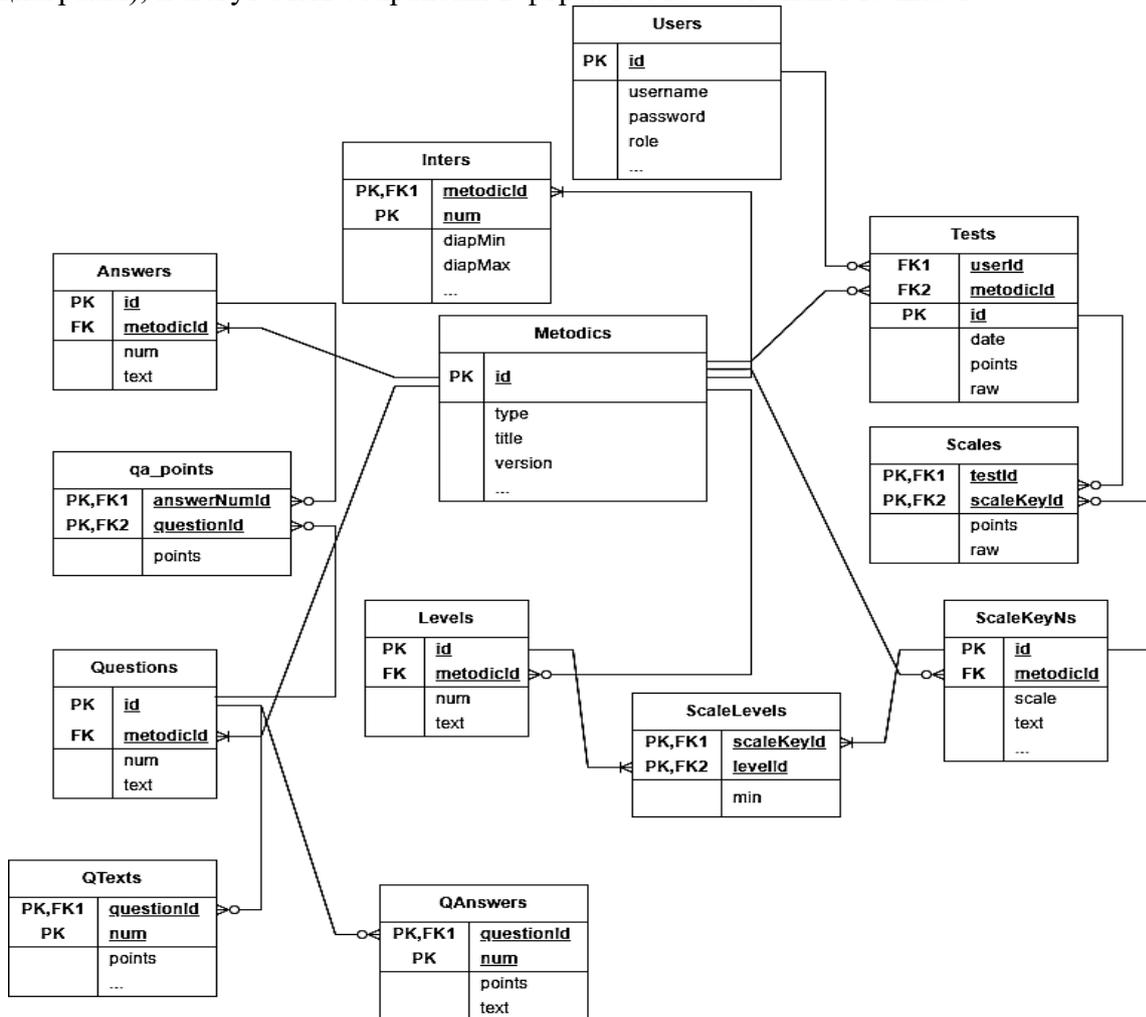


Рисунок 1 - Фрагмент схемы базы данных. Справочная информация о диагностических методиках и результаты прохождения тестов пользователями

Справочная база данных

В справочной базе данных, в соответствии с рисунком 1, хранятся следующие сведения о методике тестирования: общие сведения об авторе, версии, типе методики, категории и т.п.; список вопросов тестируемому; данные о форматировании текста вопроса при выводе; варианты ответов на вопросы; сведения о количестве сырых баллов, начисляемых за каждый вариант ответа на вопрос; информация о субшкалах (для методик, содержащих субшкалы); данные о способе преобразования сырых баллов в Т-баллы, станайны и т.п.; текстовая интерпретация полученных после преобразования результатов; данные о способе представления полученных результатов.

В разработанной компьютерной интеллектуальной системе также реализована возможность добавления в базу данных новых психодиагностических опросников как из числа уже существующих (на основании опубликованных в открытых источниках бланков, ключей и шкал для интерпретации результатов), так и авторских разработок, нуждающихся в процедуре валидации. Для этого необходимо соответствующим образом подготовить исходную информацию и заполнить предложенную форму. Пример подготовки данных и структурирования отдельных шкал опросника приведены на рисунке 2.

**МЕТОДИКА: ШКАЛА ДЛЯ КЛИНИЧЕСКОЙ ДИАГНОСТИКИ
ПТСР**

Вопрос: Случалось ли так, что Вы непроизвольно, без всякого повода, вспоминали об этих событиях? Было ли это только во сне? (Исключить, если воспоминания появлялись только во сне). Как часто это

Текст перед вопросом:

+ Добавить строку

Текст	Действия
В. ПОСТОЯННО ПОВТОРЯЮЩЕЕСЯ ПЕРЕЖИВАНИЕ ТРАВМАТИЧЕСКОГО СОБЫТИЯ	
Повторяющиеся навязчивые воспоминания о событии, вызывающие дистресс:	

Варианты ответа:

+ Добавить вариант

№ п/п	Текст варианта ответа	Число баллов	Действия
1	никогда	0	
2	один или два раза	1	
3	один или два раза в неделю	2	
4	несколько раз в неделю	3	
5	ежедневно или почти каждый день	4	

Рисунок 2 - Подготовка данных о тестовой методике. Вопрос и варианты ответа

Одной из важных функций для проведения научных исследований и валидации опросников также является возможность формирования отдельных категорий респондентов, для каждой из которых отображается ограниченный список доступных тестов, для чего при вводе или последующей корректировке данных следует отметить соответствующие категории пользователей. В настоящий момент данный инструмент доступен разработчикам и администратору, для зарегистрированных пользователей можно выбрать категорию из предложенного списка (студенты, научные работники, лица с ограниченными возможностями здоровья, участники СВО и др.).

Преимуществом разработанного коллективом ФГБНУ ИПИИ единого программного комплекса диагностики и коррекции стрессовых состояний является регулярное обновление программного обеспечения и его высокопрофессиональная техническая поддержка, позволяющая добавлять новые элементы интерфейса и функции, необходимые специалистам различных сфер (в т.ч., психология, здравоохранение, педагогика) для реализации соответствующих методологических целей и задач [7], [12].

Процедура проведения тестирования

Для проведения диагностики пользователю необходимо предварительно пройти регистрацию в данной системе. Это не противоречит анонимности опроса: система запрашивает минимум данных (например, пол, возраст, семейное положение, род занятий), которые потребуются для оценки результатов теста и последующей статистической обработки. Регистрация позволяет накапливать в БД данные о результатах психодиагностического тестирования отдельных пользователей и наблюдать за их динамикой (каждому пользователю присваивается уникальный id).

Тестирование проводится в диалоговом режиме, респондент может выбрать удобный для себя вариант: вопросы и ответы на них могут быть показаны в таблице, где следует отметить соответствующие позиции, либо каждый вопрос с вариантами ответа отображается на отдельной странице с возможностью вернуться к предыдущему вопросу и изменить ответ.

Для пользователя доступен просмотр полученных результатов тестирования, а также их интерпретация с наглядным отображением в виде шкал и диаграмм в отдельных опросниках в соответствии с рисунком 3.

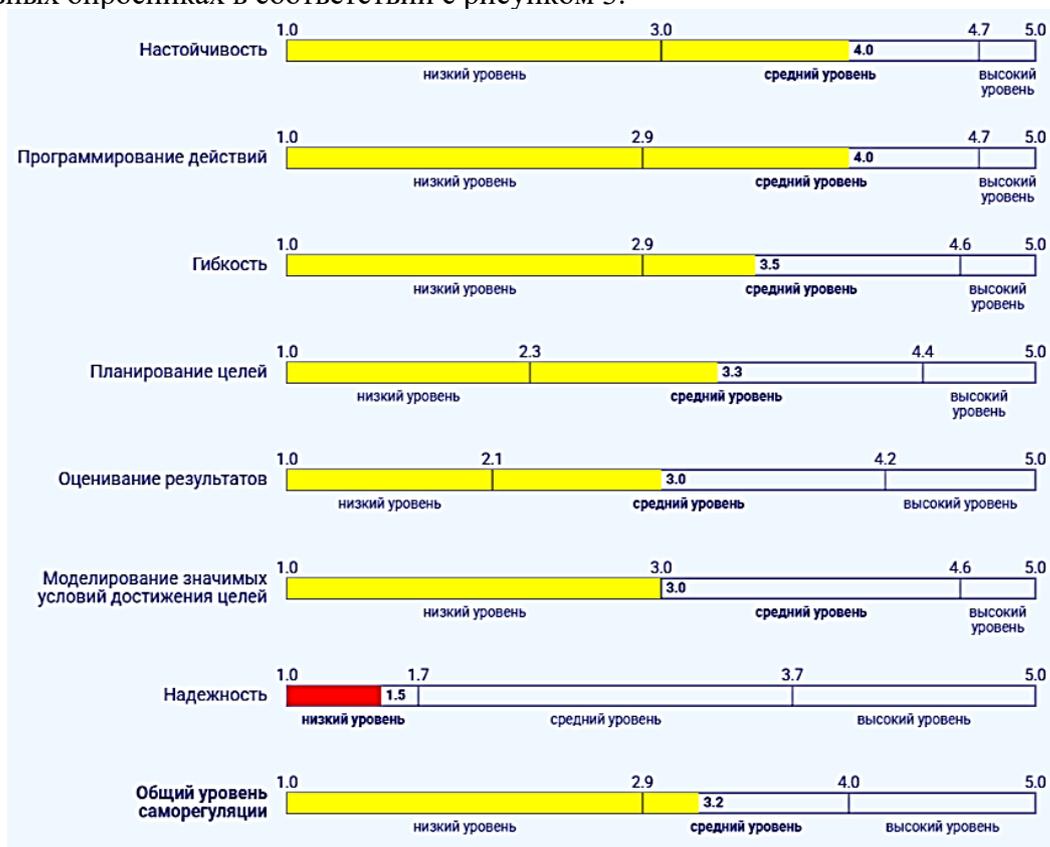


Рисунок 3 - Выводы по ответам пользователя на вопросы тестовой методики

Отображение и статистическая обработка результатов

При присвоении соответствующей роли администратором, исследователь может просмотреть результаты тестирования, выбрав интересующие его категории респондентов (например, по роду занятий, полу, возрасту, сведениям о наличии или отсутствии инвалидности) при сохранении анонимности пользователей. Информация может быть представлена в табличном виде и в виде диаграмм в соответствии с рисунком 4 и рисунком 5). При выборе определенной тестовой методики реализована

возможность распределения респондентов по произвольно выбранным возрастным группам и количественного отображения полученных для разных групп результатов. Пример отображения результатов тестирования по шкале депрессии Бека (BDI) в соответствии с рисунком 6.

№ п/п	Возраст	Пол	Семейное положение	Место жительства	Образование	Род занятий	Организация	ОВЗ или инв.	Вид	Демобилиз./участник СВО	Откуда узнали	Дата теста
41	87	жен.	замужем	Донецк	высшее	служащий	ГУ "ИПИИ"	нет		нет		23.07.2025
42	62	жен.	незамужем	Донецк	высшее	служащий		нет		нет	друзья	01.08.2025
43	30	жен.	замужем	Донецк	высшее	научный работник	TESTING	нет		нет		02.08.2025
44	30	жен.	замужем	Донецк	высшее	научный работник	TESTING	нет		нет		02.08.2025
45	30	жен.	замужем	Донецк	высшее	научный работник	TESTING	нет		нет		02.08.2025
46	22	муж.	не женат	Донецк	среднее спец.	студент	TESTING	нет		нет		15.08.2025
47	62	жен.	незамужем	Донецк	высшее	служащий		нет		нет	друзья	25.08.2025
48	26	муж.	не женат	Донецк	высшее	рабочий		нет		нет		25.08.2025
49	62	жен.	незамужем	Донецк	высшее	служащий		нет		нет	друзья	17.10.2025
50	55	жен.	замужем	Донецк	высшее	служащий	ГУ "ИПИИ"	нет		нет		05.11.2025
51	19	муж.	не женат	Донецк	среднее спец.	студент	ГУ "ИПИИ"	нет		нет		26.11.2025
52	20	жен.	незамужем	Харцызск	среднее	студент	TESTING	нет		нет		27.02.2025
53	47	жен.	незамужем	Донецк	высшее	служащий	ГУ "ИПИИ"	да		нет		13.06.2024
54	87	муж.	не женат	Донецк	высшее	служащий	ГУ "ИПИИ"	нет		нет		10.02.2025

Рисунок 4 – Отображение результатов тестирования в виде таблицы

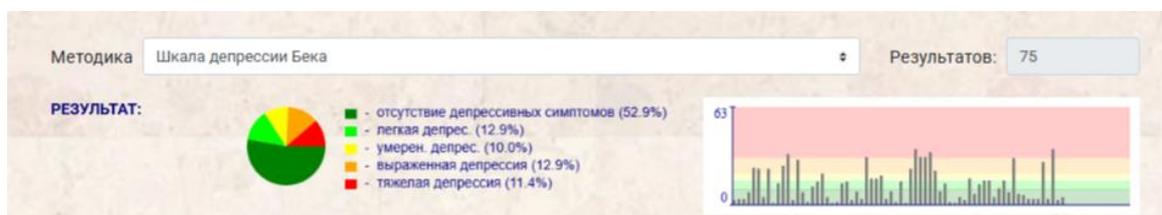


Рисунок 5 – Представление результатов в графическом виде

Границы интервалов: 20 лет, 30 лет, 50 лет

Возрастные группы: до 20 лет, 21-30 лет, 31-50 лет, 51 и старше

сокращенные названия полные названия

- ОДС - отсутствие депрессивных симптомов
- ЛД - легкая депрес.
- УД - умерен. депрес.
- ВД - выраженная депрессия
- ТД - тяжелая депрессия

Показатель	Группа																							
	до 20 лет					21-30 лет					31-50 лет					51 и старше								
	ОДС	ЛД	УД	ВД	ТД	ОДС	ЛД	УД	ВД	ТД	ОДС	ЛД	УД	ВД	ТД	ОДС	ЛД	УД	ВД	ТД				
РЕЗУЛЬТАТ:	19	3	3	2	1	28	9	3	2	2	6	22	6	2	0	2	0	10	8	1	2	3	1	15

Рисунок 6 – Результаты тестирования по сформированным возрастным группам

Для осуществления дополнительной статистической обработки в соответствии с задачами конкретного исследования в системе представлена возможность сохранения накопленных данных в форматах MS Office (Word, Excel). Экспорт результатов возможен как по отдельной методике, так и в виде сводной таблицы в соответствии с рисунком 7 по всем методикам и всем пользователям, относящимся к выбранной категории. С целью упрощения процесса интерпретации полученных данных на отдельном листе (в файле со сводной таблицей MS Excel) приведена расшифровка названий отдельных шкал и их количественная оценка по исходным материалам психодиагностического опросника в соответствии с рисунком 8.

	A	B	C	D	E	F	G	H	I	J	K
1	id	Возраст	Пол	Место жительства	Род занятий	Образование	Семейное поло	Организация	В.Бойко	PSM-25	Немчин
2	76	20 жен.	Харьцызск	студент	среднее	незамужем		4			
3	79	20 муж.	Харьцызск	студент	среднее	не женат					
4	53	42 муж.	Донецк	рабочий	высшее	не женат					
5	213	25 муж.									
6	70	20 муж.	Макеевка	студент	среднее	не женат		12			
7	98	23 муж.									
8	57	59 муж.	Донецк	служащий	высшее	женат		3			38
9	152	25 муж.	Торез	рабочий	среднее	не женат	ГУ "ИПИИ"				
10	94	21 муж.							60		42
11	227	57 жен.	Донецк	служащий	высшее	замужем	ГУ "ИПИИ"				
12	226	25 муж.	Иловайск	студент	среднее	не женат	ГУ "ИПИИ"				
13	336	56 жен.	Донецк	научный работн	высшее	замужем					
14	20	35 муж.	Горловка	служащий	среднее	не женат			58		37

Рисунок 7 – Экспорт результатов тестирования в MS Excel, данные о результатах в сводной таблице

	A	B	C	D	E	F	G	H	I
114	Тест жизнестойкости Мадди	Результат	От	До	Шкала	Описание	низкий уровень		
115		низкая		0	61	СН	принятие риска		
116		средняя		62	99	СМ	вовлеченность		
117		высокая		100	135	СН	контроль		
118									
119	Тест эмоционального интеллект	Результат	От	До	Шкала	Описание	очень низкое зн		
120		очень низкое значение		0	71	ВП	понимание своих эмоций		
121		низкое значение		72	78	ВУ	управление своими эмоци		
122		среднее значение		79	92	ВЭ	контроль экспрессии		
123		высокое значение		93	104	ВЭИ	внутриличностный ЭИ		
124		очень высокое значение		105	138	МП	понимание чужих эмоций		
125						МУ	управление чужими эмоцу		
126						МЭИ	межличностный ЭИ		
127						ПЭ	понимание эмоций		
128						УЭ	управление эмоциями		
129									
130	Стиль саморегуляции поведения	Результат	От	До	Шкала	Описание	низкий уровень		
131		низкий уровень		1.0	2.8	FL	Гибкость	1.0	
132		средний уровень		2.9	3.9	MO	Моделирование значимых	1.0	
133		высокий уровень		4.0	5.0	NA	Настойчивость	1.0	
134						OR	Оценивание результатов	1.0	
135						PL	Планирование целей	1.0	
136						PR	Программирование действ	1.0	
137						SAF	Надежность	1.0	

Рисунок 8 – Экспорт результатов тестирования в MS Excel, данные о методиках и интерпретации в сводной таблице

Средства разработки проекта

Проект реализован в клиент-серверной архитектуре (представляет собой веб-сайт), язык программирования javascript. При разработке применялись следующие технологии и программное обеспечение:

- кроссплатформенная среда выполнения Node.js – для реализации серверной части приложения [14];
- фреймворк Express.js для Node.js [15];
- система управления базами данных PostgreSQL [16];
- ORM-библиотека для работы с базами данных Sequelize [17];
- фреймворк Vue.js для создания пользовательских интерфейсов (клиентская часть приложения) [18].

Материалы и методы экспериментального этапа исследования

Основная цель экспериментального этапа исследования - валидация программного обеспечения - проверка его работоспособности и соответствия потребностям пользователя. Вначале тестирование проводилось командой разработчиков (альфа-тестирование). Затем нескольким группам внешних пользователей было предложено принять участие в валидации системы: зарегистрироваться, пройти тестирование и, по возможности, оставить свои отзывы и замечания. Тестирование осуществлялось в соответствии со следующими критериями:

- корректность работы системы на различных устройствах;
- согласованность работы различных компонентов системы (клиентского интерфейса и backend), корректность накопления информации в БД в процессе регистрации и прохождения тестов (интеграционное тестирование);
- в какой степени продукт удовлетворяет требованиям реальных пользователей (бета-тестирование);
- насколько удобен и интуитивно понятен интерфейс (usability-тестирование).

Тестирование ПО в реальных условиях подтвердило его соответствие функциональному назначению и надежность. Обратная связь от реальных пользователей позволила выявить и устранить отдельные неточности, улучшить удобство и функциональность продукта.

Контингент обследуемых лиц составили 65 респондентов в возрасте от 17 до 87 лет (средний возраст – $30,52 \pm 17,02$ лет), проживающих на территории проведения специальной военной операции (г. Донецк, Макеевка, Енакиево, Иловайск, Харцызск Донецкой Народной Республики) и пребывающих в условиях стресса военного времени, не принимавших непосредственного участия в боевых действиях и не находившихся на военной службе. Распределение респондентов по гендерному признаку было следующим: мужчины – 37 чел. (56,92%), женщины – 28 чел. (43,08%). К лицам молодого возраста были отнесены 48 обследуемых (73,85%) возрасте до 35 лет включительно (средний возраст – $21,60 \pm 3,07$ лет), из них мужчин – 30 человек (62,50%), женщин – 18 человек (37,50%). К лицам старшей возрастной группы – 17 респондентов (26,15%) в возрасте старше 35 лет (средний возраст – $55,71 \pm 14,77$ лет), из них: мужчины – 7 человек (41,18%), женщины – 10 человек (58,82%). Сравнительный анализ результатов исследования проводился по гендерному признаку, а также по возрастным группам.

Диагностика степени выраженности депрессивных симптомов осуществлялась с использованием шкалы депрессии Бека (Beck Depression Inventory) [19]. Опросник включает в себя 21 категорию симптомов и жалоб. Каждая категория состоит из 4-5 утверждений, соответствующих специфическим симптомам депрессии, которые оцениваются испытуемым от 0 (симптом отсутствует или выражен минимально) до 3 баллов (максимальная выраженность симптома). Результат интерпретировался с учетом суммарного балла по всем категориям следующим образом: 0-9 – отсутствие депрессивных симптомов; 10-15 – легкая депрессия (субдепрессия); 16-19 – умеренная депрессия; 20-29 – выраженная депрессия (средней тяжести); 30-63 – тяжелая депрессия.

Оценка уровня тревоги в качестве психопатологического феномена, существующего в настоящий момент (реактивная тревожность) и устойчивой личностной характеристики (личностная тревожность) проводилась с помощью шкалы тревоги Спилбергера (State-Trait Anxiety Inventory) в русскоязычной адаптации Ю. Л. Ханина [20]. Опросник содержит два блока утверждений по 20 вопросов, степень согласия с которыми оценивалась испытуемыми по следующей шкале: «нет, это не так»; «пожалуй, так»; «верно» и «совершенно верно». Подсчет результатов производился в соответствии с ключом и интерпретировался следующим образом: до 29 баллов – низкий, 30-45 баллов – умеренный, 46 и более – высокий уровень тревожности.

Обработка данных. Для статистической обработки данных применялся t-критерий Стьюдента для несвязанных выборок. Различия считались статистически значимыми на уровне $p \leq 0,05$. В тексте и таблицах данные представлены следующим образом: среднее арифметическое \pm стандартное отклонение (95% ДИ). Обработка данных проводилась с использованием таблиц MS Excel и лицензионных пакетов статистического анализа - «Statistica 10.0» (StatSoft) на IBM PC/AT.

Результаты и обсуждение. При анализе полученных данных среди всего контингента респондентов средний результат по шкале депрессии Бека составил $11,69 \pm 10,21$ балла (95% ДИ: 8,79-14,59) (субдепрессия/ легкая депрессия), для когнитивно-аффективной субшкалы – $7,93 \pm 7,05$ балла (95% ДИ: 5,50-10,36), для субшкалы соматизации – $4,24 \pm 4,19$ балла (95% ДИ: 2,80-5,69). По шкале тревоги Спилбергера были получены умеренные значения – $42,17 \pm 10,78$ балла (95% ДИ: 39,11-45,23) для ситуативной и $45,09 \pm 11,03$ балла (95% ДИ: 41,96-48,23) для личностной тревожности. При анализе результатов исследования было выявлено отсутствие статистически значимых различий в зависимости от гендерной принадлежности, а также между лицами молодого возраста и старшей возрастной группы (таблица 1).

Таблица 1- Результаты анализа выраженности тревожно-депрессивной симптоматики у лиц, пребывающих в условиях стресса военного времени (гендерно-возрастной аспект)

	Мужчины (n=37)	Женщины (n=28)	p	До 35 лет (n=48)	Старше 35 лет (n=17)	p
Шкала депрессии Бека	$12,97 \pm 10,86$ (8,82-17,13)	$10,00 \pm 9,21$ (5,87-14,13)	0,237	$11,58 \pm 10,59$ (8,05-15,12)	$12,00 \pm 9,33$ (6,40-17,60)	0,878
Ситуативная тревожность	$43,19 \pm 11,48$ (38,79-47,59)	$40,82 \pm 9,82$ (36,42-45,22)	0,375	$40,98 \pm 10,79$ (37,38-44,58)	$45,53 \pm 10,31$ (39,34-51,72)	0,126
Личностная тревожность	$45,03 \pm 11,59$ (40,59-49,47)	$45,18 \pm 10,44$ (40,50-49,86)	0,957	$44,13 \pm 10,98$ (40,46-47,79)	$47,82 \pm 11,01$ (41,22-54,43)	0,240

Таким образом, вне зависимости от гендерно-возрастных различий, мирные жители Донбасса обнаруживают признаки легкой депрессии, а именно, подавленное настроение, снижение уровня энергии, повышенную утомляемость, нарушения сна и аппетита, неспособность получать удовольствие привычными способами, которые в настоящий момент не оказывают значимого влияния на общий уровень жизнедеятельности и работоспособность, однако, сопровождаются субъективно тягостными внутренними ощущениями и переживаниями. Также для них характерны чувство обеспокоенности, нервозность, повышенная возбудимость и раздражительность, связанные с актуальной стрессовой ситуацией, а также наличие устойчивых личностных паттернов восприятия и интерпретации широкого спектра жизненных событий как угрожающих, несущих негативные последствия для самооценки и собственной жизни.

Заключение

В результате экспериментального этапа исследования и валидации компьютерной интеллектуальной системы диагностики психоэмоционального состояния осуществилась проверка его работоспособности и соответствия потребностям пользователя путем анонимного тестирования группы пользователей (студенты вузов и колледжей ДНР, научные сотрудники): удалось выявить и устранить ряд неточностей и недостатков ПО, улучшить пользовательский интерфейс с учетом предложений участников процесса валидации (обратной связи по работе с системой со стороны исследователя и респондентов). Данные результатов диагностики были экспортированы в MS Excel и

обработаны с помощью пакетов статистического анализа, что позволило получить фактическую информацию о психоэмоциональном состоянии лиц, находящихся под воздействием стресса военного времени, которая может быть использована профильными специалистами для разработки стратегий психокоррекционного вмешательства.

Список литературы

1. Теоретические аспекты диагностики и саморегуляции психоэмоциональных состояний личности / Ивашко К.С., Изосимова С.А., Пигуз В.Н., Ковалева К.В. // «Проблемы искусственного интеллекта». 2025. №1 (36). С. 58–67. ISSN 2413-7383. DOI 10.24412/2413-7383-58-67.
2. Ивашко, К. С. Интеллектуальные системы саморегуляции психоэмоционального состояния в условиях военных конфликтов / С. Ивашко, С. А. Изосимова, В. Н. Пигуз, К. В. Ковалева // Инновационные подходы в современной науке. сб. ст. по материалам СХСIV междунар. науч.-практ. конф. № 14 (194). М., Изд. «Интернаука», 2025. 78 с. С. 32-36. ISSN 2587-8603
3. Перспективы разработки теоретических основ и компьютерных интеллектуальных технологий диагностирования и безмедикаментозной саморегуляции психоэмоциональных состояний личности / Изосимова С.А., Пигуз В.Н., К.С. Ивашко, К.В. Ковалева. // Личность и семья в условиях новой нормальности: жизнедеятельность, совладание и психологическая помощь: материалы Международной научно-практической конференции «Человек. Семья. Культура: активность, ресурсы и помощь в условиях новой нормальности» (Москва, 1–2 октября 2025 г.) / Под науч. ред. Е.В. Куфтяк. М. : Издательство Московского института психоанализа, 2025. 312 с. С. 257-262. ISBN 978-5-904677-53-4
4. Исследование эффективного использования безмедикаментозных методов и средств интеллектуальной терапии и возможностей эмоционального искусственного интеллекта для целей саморегуляции психоэмоциональных состояний / И. С. Сальников, Р. И. Сальников, С. А. Изосимова, В. Н. Пигуз, К. С. Ивашко, И. М. Коваленко // Ежемесячный научно-практический журнал «Психическое здоровье [Mental Health]». 2024. Т. 19, № 5. С. 87–89. ISSN 2074-014X. DOI: 10.25557/2074-014X.2024.05.87-89
5. Автоматизированный анализ психологического состояния личности: подходы и перспективы / К. С. Ивашко, С. А. Изосимова, В. Н. Пигуз, К. В. Ковалева // Молодой исследователь: вызовы и перспективы : Сборник статей по материалам CDXXIII международной научно-практической конференции, Москва, 22 сентября 2025 года. Москва: Общество с ограниченной ответственностью "Интернаука", 2025. С. 15-20. EDN TNXVIC. URL: <https://www.elibrary.ru/item.asp?id=82988723> (дата обращения: 03.12.2025).
6. Применение искусственного интеллекта в диагностике и коррекции психических состояний / К. С. Ивашко, С. А. Изосимова, В. Н. Пигуз, К. В. Ковалева // Молодой исследователь: вызовы и перспективы : сб. ст. по материалам CDXXV междунар. науч.-практ. конф. № 38(425). М., Изд. «Интернаука», 2025. 107 с. С. 25-29. ISSN 2415-8771
7. Роль безмедикаментозной компьютерной интеллектуально-духовной терапии в процессе обучения и развития творческих способностей – проблемы и перспективы / И. С. Сальников, С. А. Изосимова, В. Н. Пигуз, К. С. Ивашко // «Клиническая психология в здравоохранении, образовании и социальной сфере: методы арт-терапии в практике психологической помощи» : сб. научных трудов III Всероссийской межведомственной научно-практической конференции, Нижний Новгород, 15–16 февраля 2024 г. / ННГУ им. Н.И. Лобачевского; Факультет социальных наук; Редкол.: Т.В. Маркелова, Ж.А. Шуткина, А.Ю. Акимова. Нижний Новгород: Изд-во ННГУ, 2024. Т. 2. 215 с. Текст: электронный. С. 317–328. ISBN 978-5-91326-878-5
8. Свидетельство о государственной регистрации программы для ЭВМ № 2025611192 Российская Федерация. Программа для диагностирования психоэмоциональных состояний личности : заявл. 26.12.2024 : опубл. 16.01.2025 / В. В. Бондарчук, Н. М. Кравченко, Т. Д. Ключанова, К. В. Ковалева ; заявитель Федеральное государственное бюджетное научное учреждение "Институт проблем искусственного интеллекта". EDN GZLNQIN.
9. Аналитический обзор разрабатываемой компьютерной системы диагностики психоэмоциональных состояний личности, пребывающей в условиях хронического стресса военного времени / Изосимова С.А., Сальников И.С., Пигуз В.Н., Ивашко К.С., Ковалева К.В. // Донецкие чтения – 2025: образование, наука, инновации, культура и вызовы современности: Материалы X Международной научной конференции, посвященной 60-летию создания Донецкого научного центра (Донецк, 5–7 ноября 2025 г.). Том 2: Физические, технические и компьютерные науки / Под общей редакцией проф. С.В. Беспаловой. Донецк: Изд-во ДонГУ, 2025. 470 с. С. 336-339. ISSN 2664-7370 (online); ISSN 2664-7362 (print).
10. Психоэмоциональное реагирование личности в условиях хронического стресса военного времени / Сальни-ков И.С., Изосимова С.А., Пигуз В.Н., Ивашко К.С. // Российский научный журнал «Телескоп: журнал социологических и маркетинговых исследований». 2025. № 1(17). С. 47-55. DOI: 10.24412/1994-3776-2025-1-47-55. ISSN 1994-3776.
11. Анализ и систематизация процесса социальной адаптации людей, пострадавших от хронического стресса военного времени / И. С. Сальников, С. А. Изосимова, В. Н. Пигуз, К.С. Ивашко // Российский научный журнал «Телескоп: журнал социологических и маркетинговых исследований». 2024. № 2 (14). С. 6–12. – 142 с. DOI: 10.24412/1994-3776-2024-2-6-12. ISSN 1994-3776

12. Когнитивный аспект исследований стресса военного времени / В. Н. Пигуз, К. С. Ивашко, С. А. Изосимова // Проблемы искусственного интеллекта. 2024. № 3 (34). С. 86–101. ISSN 2413-7383. DOI 10.24412/2413-7383-2024-3-88-101
13. Система психодиагностики и саморегуляции : сайт. Донецк. URL: <http://psydiag.guiaidn.ru/home>
14. Node.js — Run JavaScript Everywhere : сайт. URL: <https://nodejs.org/en> (дата обращения: 04.12.2025).
15. Express - Node.js web application framework : сайт. URL: <https://expressjs.com/> (дата обращения: 04.12.2025).
16. PostgreSQL: The world's most advanced open source database : сайт. URL: <https://www.postgresql.org/> (дата обращения: 04.12.2025).
17. Sequelize : сайт. URL: <https://sequelize.org> (дата обращения: 04.12.2025).
18. Vue.js - The Progressive JavaScript Framework | Vue.js : сайт. URL: <https://vuejs.org/> (дата обращения: 04.12.2025).
19. Батаршев, А. В. Базовые психологические свойства и самоопределение личности : практическое руководство по психологической диагностике. Санкт-Петербург : Речь, 2005. С. 44-49. Текст : непосредственный.
20. Тарабрина, Н. В. Практикум по психологии посттравматического стресса / Н. В. Тарабрина. – Санкт-Петербург : Питер, 2001. 272 с.: ил. Серия «Практикум по психологии». – Текст : непосредственный.

References

1. Ivashko K.S., Izosimova S.A., Piguz V.N., Kovaleva K.V. Theoretical aspects of diagnosis and self-regulation of psychoemotional states of personality // "Problems of artificial intelligence". 2025. No. 1 (36). pp. 58-67. ISSN 2413-7383. DOI 10.24412/2413-7383-58-67.
2. Ivashko, K. S. Intellectual systems of self-regulation of the psycho-emotional state in conditions of military conflicts / S. Ivashko, S. A. Izosimova, V. N. Piguz, K. V. Kovaleva // Innovative approaches in modern science. collection of articles based on the materials of the CXCV International Scientific and Practical Conference No. 14 (194). Moscow, Internauka Publishing House, 2025. 78 p. pp. 32-36. ISSN 2587-8603
3. Izosimova S.A., Piguz V.N., K.S. Ivashko, K.V. Kovaleva Prospects for the development of theoretical foundations and computer intelligent technologies for the diagnosis and drug-free self-regulation of psychoemotional states of personality. / Personality and family in the context of the new normality: vital activity, coping and psychological assistance: proceedings of the International Scientific and Practical Conference "Man. Family. Culture: activity, resources and assistance in the context of the new normality" (Moscow, October 1-2, 2025) / Under the scientific editorship of E.V. Kuftyak. : Publishing House of the Moscow Institute of Psychoanalysis, 2025. 312 p. pp. 257-262. ISBN 978-5-904677-53-4
4. Salnikov I. S., Salnikov R. I., Izosimova S. A., Piguz V. N., Ivashko K. S., Kovalenko I. M. Investigation of the effective use of drug-free methods and means of intellectual therapy and the possibilities of emotional artificial intelligence for the purposes of self-regulation of psychoemotional states / I. S. Salnikov, Izosimova R. I., Piguz V. N., Ivashko K. S., Kovalenko I. M. // Monthly scientific and practical journal "Mental Health". 2024. Vol. 19, No. 5. pp. 87-89. ISSN 2074-014X. DOI: 10.25557/2074-014X.2024.05.87-89
5. Automated analysis of the psychological state of a personality: approaches and prospects / K. S. Ivashko, S. A. Izosimova, V. N. Piguz, K. V. Kovaleva // Young researcher: challenges and prospects : A collection of articles based on the materials of the XXXIII International scientific and practical conference, Moscow, September 22, 2025. Moscow: Internauka Limited Liability Company, 2025. pp. 15-20. EDN TNXVIC. URL: <https://www.elibrary.ru/item.asp?id=82988723> (date of request: 03.12.2025).
6. Application of artificial intelligence in the diagnosis and correction of mental states / K. S. Ivashko, S. A. Izosimova, V. N. Piguz, K. V. Kovaleva // Young researcher: challenges and prospects : collection of articles based on the materials of the CDXXV International Scientific and practical Conference No. 38(425). Moscow, Ed. "Internauka", 2025. 107 p. pp. 25-29. ISSN 2415-8771
7. The role of drug-free computer intellectual and spiritual therapy in the process of learning and developing creative abilities – problems and prospects / I. S. Salnikov, S. A. Izosimova, V. N. Piguz, K. S. Ivashko // "Clinical psychology in healthcare, education and the social sphere: methods of art therapy in the practice of psychological assistance": Collection of scientific papers of the III All-Russian Interdepartmental Scientific and Practical Conference, Nizhny Novgorod, February 15-16, 2024 / N.I. Lobachevsky National Research University; Faculty of Social Sciences; Editors: T.V. Markelova, J.A. Shutkina, A.Y. Akimova. Nizhny Novgorod: UNN Publishing House, 2024. Vol. 2. 215 p. Text: electronic. pp. 317-328. ISBN 978-5-91326-878-5
8. Certificate of state registration of the computer program No. 2025611192 Russian Federation. A program for diagnosing psychoemotional states of personality : application. 12/26/2024 : published. 01/16/2025 / V. V. Bondarchuk, N. M. Kravchenko, T. D. Klyushanova, K. V. Kovaleva ; applicant Federal State Budgetary Scientific Institution "Institute of Artificial Intelligence Problems". – EDN GZLNQIN.
9. An analytical review of the computer system being developed for the diagnosis of psychoemotional states of a person living under conditions of chronic wartime stress / Izosimova S.A., Salnikov I.S., Piguz V.N., Ivashko K.S., Kovaleva K.V. // Donetsk Readings – 2025: education, science, innovation, culture and modern challenges: Proceedings of the X International Scientific Conference dedicated to the 60th anniversary of the establishment of the Donetsk Scientific Center (Donetsk, November 5-7, 2025) Volume

- 2: Physical, Technical and Computer Sciences / Edited by Professor S.V. Bespalova. Donetsk: DonGU Publishing House, 2025. 470 p. Pp. 336-339. ISSN 2664-7370 (online); ISSN 2664-7362 (print).
10. Salnikov I.S., Izosimova S.A., Piguz V.N., Ivashko K.S. Psychoemotional response of personality in conditions of chronic stress of wartime // Russian scientific journal "Telescope: journal of sociological and marketing research". 2025. No. 1(17). PP. 47-55. DOI: 10.24412/1994-3776-2025-1-47-55. ISSN 1994-3776.
 11. Analysis and systematization of the process of social adaptation of people affected by chronic wartime stress / I. S. Salnikov, S. A. Izosimova, V. N. Piguz, K.S. Ivashko // Russian scientific journal "Telescope: journal of sociological and marketing research". 2024. No. 2 (14). pp. 6-12. – 142 p. DOI: 10.24412/1994-3776-2024-2-6-12. ISSN 1994-3776
 12. The cognitive aspect of wartime stress research / V. N. Piguz, K. S. Ivashko, S. A. Izosimova // Problems of artificial intelligence. 2024. No. 3 (34). pp. 86-101. ISSN 2413-7383. DOI 10.24412/2413-7383-2024-3-88-101
 13. The system of psychodiagnostics and self-regulation : website. Donetsk. URL: <http://psydiag.guide.ru/home>
 14. Node.js — Run JavaScript Everywhere : website. URL: <https://nodejs.org/en> (date of access: 04.12.2025).
 15. Express - Node.js web application framework : website. URL: <https://expressjs.com/> (accessed: 12/04/2025).
 16. PostgreSQL: The world's most advanced open source database : website. URL: <https://www.postgresql.org/> (accessed: 12/04/2025).
 17. Sequelize : website. URL: <https://sequelize.org> (date of request: 04.12.2025).
 18. Vue.js - The Progressive JavaScript Framework | Vue.js : the website. URL: <https://vuejs.org/> (date of reference: 12/04/2025).
 19. Batarshchikov, A.V. Basic psychological properties and self-determination of personality : a practical guide to psychological diagnostics / A.V. Batarshchikov. – St. Petersburg : Speech, 2005. – pp. 44-49. – Text : direct.
 20. Tarabrina, N. V. Practicum on the psychology of post-traumatic stress / N. V. Tarabrina. – St. Petersburg : Peter, 2001. – 272 p. : ill. – Series "Practical psychology". – Text : direct.

RESUME

Kovaleva K.V., Kravchenko N.M., Bondarchuk V.V. Piguz V.N., Izosimova S.A., Ivashko K.S. The Results Of The Pilot Operation Of An Intelligent Computer System For The Diagnosis And Correction Of Stress Status

An analysis of the theoretical foundations of diagnosing and self-regulating psychoemotional states of personality in the territories currently affected by the special military operation has demonstrated the need to develop computer-based intelligent programs for prevention, diagnosis, and psychocorrection aimed at enhancing stress resistance and improving quality of life.

The project is built using a client-server architecture (a website) and the JavaScript programming language. The software was initially tested through alpha testing (by the development team) and then beta testing (by several groups of external users) to determine how well the product meets user needs. The third stage involved evaluating the interface's usability and intuitiveness.

The results of the experimental stage of the development of an intelligent computer system for the diagnosis and correction of stress conditions aimed at validating software tools are studied. We present an analysis of the severity anxiety-depressive symptoms in people exposed to wartime stress, based on the results of online testing conducted using the author's program for psychoemotional states of personality diagnosing.

РЕЗЮМЕ

К. В. Ковалева, Н. М. Кравченко, В. В. Бондарчук, В. Н. Пигуз, С. А. Изосимова, К. С. Ивашко Результаты опытной эксплуатации компьютерной интеллектуальной системы диагностики и коррекции стрессовых состояний

Анализ теоретических основ диагностики и саморегуляции психоэмоциональных состояний личности в настоящее время на территориях проведения специальной военной операции показал актуальность создания компьютерных интеллектуальных программ профилактики, диагностики и психокоррекции, направленных на повышение стрессоустойчивости, улучшение качества жизни и психоэмоционального благополучия общества.

Проект реализован в клиент-серверной архитектуре (представляет собой веб-сайт), язык программирования javascript. Тестирование программного обеспечения проводилось вначале методом альфа-тестирования (группой разработчиков), затем бета-тестирования (несколькими группами внешних пользователей) – в какой степени продукт удовлетворяет требованиям реальных пользователей; третий этап – насколько удобен и интуитивно понятен интерфейс (usability-тестирование).

Исследованы результаты экспериментального этапа разработки компьютерной интеллектуальной системы диагностики и коррекции стрессовых состояний, направленного на валидацию программного инструментария. По результатам онлайн-тестирования, проведенного с использованием авторской программы для диагностирования психоэмоциональных состояний личности, представлен анализ выраженности тревожно-депрессивной симптоматики у лиц, подверженных стрессу военного времени

Ковалева Ксения Витальевна, младший научный сотрудник, ФГБНУ «Институт проблем искусственного интеллекта»; Российская Федерация, г. Донецк, E-mail: kv_kovaleva@mail.ru, https://elibrary.ru/author_profile.asp?authorid=1012018 SPIN-код: 7315-3231, AuthorID: 1012018, ORCID: 0009-0002-8218-1898. Врач-психиатр ГБУ ДНР "РКПБ-ПТЦ", *Область научных интересов*: когнитивные исследования и автоматизированные системы диагностики.

Кравченко Наталья Михайловна, к.т.н., с.н.с., ФГБНУ «Институт проблем искусственного интеллекта», Российская Федерация, г. Донецк, E-mail: natali.kravchenko70@mail.ru, тел +7(949)-431-51-70, https://elibrary.ru/author_profile.asp?authorid=1183092 SPIN-код: 3950-1613, AuthorID: 1183092, ORCID: 0009-0001-1661-1872 *Область научных интересов*: когнитивные исследования, автоматизированные системы диагностики, системы поддержки принятия решений в медицине, математическое моделирование проветривания подземных сооружений

Бондарчук Виктория Валерьевна, к.т.н., с.н.с., ФГБНУ «Институт проблем искусственного интеллекта», Российская Федерация, г. Донецк ул. Артема, 118 б, телефон +7 949 315-79-53, E-mail: vv_bondar@mail.ru, https://elibrary.ru/author_profile.asp?authorid=278218, SPIN-код: 8247-9620, AuthorID: 278218, ORCID iD [0009-0007-1794-5309](https://orcid.org/0009-0007-1794-5309). *Область научных интересов*: проектирование автоматизированных систем управления, разработка алгоритмов косвенной критериальности мультимодальных процессов контроля и управления параметрическим анализом когнитивных систем.

Пигуз Валентина Николаевна, научный сотрудник, ФГБНУ «Институт проблем искусственного интеллекта», Российская Федерация, г. Донецк +7-949-405-26-76, E-mail: maxpvn77@gmail.com, https://elibrary.ru/author_profile.asp?authorid=1075640, SPIN-код: 1230-3345, AuthorID: 1075640, ORCID iD [0009-0003-3968-3144](https://orcid.org/0009-0003-3968-3144). *Область научных интересов*: когнитивные исследования и автоматизированные системы диагностики.

Изосимова Снежана Александровна, младший научный сотрудник, ФГБНУ «Институт проблем искусственного интеллекта», Российская Федерация, г. Донецк, E-mail: izosimova.snezhana@mail.ru, https://elibrary.ru/author_profile.asp?authorid=1175226 SPIN-код: 3162-6912, AuthorID: 1175226, ORCID [0000-0002-1521-5579](https://orcid.org/0000-0002-1521-5579), *Область научных интересов*: когнитивные исследования и автоматизированные системы диагностики.

Ивашко Кристина Сергеевна, младший научный сотрудник, ФГБНУ «Институт проблем искусственного интеллекта», Российская Федерация, г. Донецк, E-mail: kristi_8@mail.ru, https://elibrary.ru/author_profile.asp?id=1222071 SPIN-код: 1231-5701, AuthorID: 1222071, ORCID iD [0000-0002-0195-2279](https://orcid.org/0000-0002-0195-2279) *Область научных интересов*: когнитивные исследования и автоматизированные системы диагностики.

Статья поступила в редакцию 05.11.2025.

УДК 004.932.2

DOI 10.24412/2413-7383-2025-4-39-219-228

Е. Г. Курган, К. С. Петрова, А. С. Собчишен

Федеральное государственное бюджетное научное учреждение
«Институт научно-технической информации»

Российская Федерация, ДНР, 283001 г. Донецк, ул. Артема, 60

РАЗРАБОТКА ЭЛЕКТРОННОЙ БАЗЫ ДАННЫХ КАК ИНСТРУМЕНТ ЦИФРОВИЗАЦИИ ЦЕНТРАЛЬНОЙ НАУЧНО-ТЕХНИЧЕСКОЙ БИБЛИОТЕКИ

E. G. Kurgan, K. S. Petrova, A. S. Sobchishen

Federal state budgetary scientific institution "Institute of scientific and technical information"
Russian Federation, DPR, 283001 Donetsk, Artema St., 60

DEVELOPMENT OF AN ELECTRONIC DATABASE AS A TOOL FOR DIGITALIZATION OF THE CENTRAL SCIENTIFIC AND TECHNICAL LIBRARY

В работе представлена разработка электронной базы данных «Путеводитель по депозитарному фонду библиотеки», обладающая свойством масштабируемости и интуитивным интерфейсом, обеспечивающая удобство поиска литературных источников по тематическому принципу, централизованное хранение метаданных. Обоснована необходимость цифровизации ЦНТБ ФГБНУ «ИНТИ». На основе сравнительного анализа выбрана СУБД PostgreSQL. Приведено описание реляционной структуры из таблиц и веб-интерфейса для управления данными. Разработанная система, как инструмент цифровизации, служит основой для интеграции библиотеки в единое информационное пространство Российской Федерации.

Ключевые слова: электронная база данных, цифровизация, библиотека, доступность данных.

This article examines the digitalization of libraries. The effects of digitalization in the library sector are examined. The need for digitalization of the Central Scientific and Technical Library of the Federal State Budgetary Scientific Institution «ISTI» is substantiated. An electronic database «Guide to the Library's Depository Fund» has been developed, which has the property of scalability and an intuitive interface, providing convenient search for literary sources based on a thematic principle and centralized storage of metadata. Based on a comparative analysis, the PostgreSQL DBMS was selected. A description of the relational structure consisting of tables and a web interface for data management is provided. The developed system serves as the basis for integrating the library into the unified information space of Russia.

Keywords: electronic database, digitalization, library, data accessibility.

Введение

Цифровизация, являющаяся общемировым трендом, активно внедряется в библиотечном пространстве Российской Федерации. Перед центральной научно-технической библиотекой ФГБНУ «Институт научно-технической информации» (ЦНТБ), старейшей научно-технической библиотеки на территории воссоединенных субъектов, также стоит задача интеграции в библиотечную среду РФ, в ГСНТИ. В связи с этим актуальной является задача разработки инструментов цифровизации библиотеки.

Цель исследования – разработка электронной базы данных «Путеводитель по депозитарному фонду библиотеки».

Теоретические основы

Цифровизация Российской Федерации определена как один из ключевых элементов стратегического развития библиотечной сферы, что отражено в нормативной правовой базе [1-3]. Анализ исследований по данной тематике позволил определить, что цифровизация библиотечной сферы и, в частности, научно-технических библиотек, определяет ряд эффектов:

1. Процесс формирования цифровой экономики [4] охватывает в том числе и сферу научно-технической информации как основы достижения целей научно-технологического развития страны.
2. Цифровизация определяет механизмы модернизации библиотечной отрасли в соответствии с вызовами современности [5];
3. Цифровизация способствует формированию единого пространства научных знаний [6], информационного рынка [7], [8], электронного библиотековедения [9] и цифровых интегрированных систем [10];
4. Применение цифровизации способствует реализации клиентоориентированного подхода при оказании библиотечных услуг [11], изменение подхода к хранению и предоставлению информации пользователям, а также повышению эффективности деятельности библиотек [12-14].

В контексте интеграции в российское библиотечное пространство ЦНТБ в первую очередь решает вопрос цифровизации, что позволит модернизировать библиотеку, сделать ее фонды доступными для широкого круга пользователей, осуществить переход к модели электронной библиотеки и обеспечить возможность включения научно-технической информации, хранящейся в ЦНТБ, в библиотечные интегрированные системы. В связи с военными действиями, начавшимися в 2014 г., восстановление деятельности библиотеки и ее развитие в соответствии с российскими тенденциями ограничено рядом проблем, отраженных в [15]. На данном этапе оцифровка фондов для предоставления онлайн доступа к полнотекстовым версиям изданий невозможна по причине отсутствия технической оснащенности библиотеки. В рамках имеющихся возможностей была разработана электронная база данных, содержащая информацию о депозитарном фонде ЦНТБ.

Основой для базы данных разработки послужило тематическое структурирование депозитарного фонда ЦНТБ [15]. Цель создания – разработать централизованную и масштабируемую электронную базу данных «Путеводитель по депозитарному фонду библиотеки» для учёта, хранения и быстрого поиска метаданных ЦНТБ с интуитивным веб-интерфейсом и возможностью интеграции в внешние системы. Методологической основой реализации проекта стали исследования [16], [17]. Реализация описанных выше задач – от оцифровки фондов до создания интегрированного пространства –

требует надёжного и функционального инструментария для работы с данными. Критически важным является выбор системы управления базами данных (СУБД), который предопределяет успех всего проекта модернизации ЦНТБ. В контексте построения сложных информационных систем, способных к адаптации и семантической организации данных [18], а также с учетом функциональных возможностей современных онтологических платформ и графовых СУБД [19], такой выбор должен обеспечивать не только хранение, но и возможность последующего развития системы в сторону интеллектуальной обработки знаний и их интеграции в более широкие информационные пространства.

Анализ систем управления базами данных

Для определения наиболее подходящей системы управления базами данных был выполнен сравнительный анализ широко используемых систем, оценённых по ключевым параметрам: функциональные характеристики, возможности масштабирования, соответствие стандартам и особенности лицензирования. Результаты исследования отражены в соответствии с таблицей 1.

Таблица 1 – Анализ СУБД

Параметр сравнения	Вид СУБД				
	PostgreSQL [18]	Microsoft SQL Server [19]	MySQL [20]	Oracle Database [21]	MongoDB [22]
Лицензия	Open Source. PostgreSQL. Бесплатная.	Proprietary. Microsoft. Платная.	Open Source. GPL. Бесплатная.	Proprietary. Oracle. Платная.	Open Source. SSPL/AGPL. Бесплатная.
Поддержка JSON формата	Глубокая поддержка JSON	Ограниченная поддержка	Основная поддержка	Поддержка через плагины	Основная функция
Степень соблюдения стандартов	Высокая	Высокая	Средняя	Очень высокая	Нет (NoSQL)
Масштабируемость	Высокая	Высокая	Ограниченная (федерация)	Очень высокая	Массивная (за счёт кластера)
Расширяемость	Высокая (модули, плагины)	Средняя	Низкая	Низкая	Высокая
Сообщество	Большое сообщество	Большое сообщество	Большое сообщество	Большое сообщество	Активное сообщество
Поддержка сложных запросов	Отличная	Отличная	Ограниченная	Отличная	Ограниченная
Локальная версия	Есть (Postgres Pro)	Нет	Нет	Нет	Нет

Основываясь на проведённом анализе, выбрана СУБД PostgreSQL, так как она собрала все необходимые преимущества для проекта, а именно:

1. Лицензия: PostgreSQL – это полностью бесплатная open-source СУБД, что дает свободу использования и изменения без финансовых ограничений;
2. Поддержка JSON: глубокая поддержка JSON делает PostgreSQL идеальной для работы с гибридными данными;

3. Соответствие стандартам: высокая степень соблюдения SQL-стандартов гарантирует надежность и портативность кода;
4. Масштабируемость: PostgreSQL легко справляется с ростом данных и увеличением нагрузки;
5. Расширяемость: экосистема модулей и плагинов позволяет адаптировать СУБД под нужды проекта;
6. Сообщество: активное сообщество поддерживает разработку и решение проблем, что обеспечивает регулярные обновления;
7. Поддержка сложных запросов: отличная оптимизация сложных запросов обеспечивает эффективность работы с данными;
8. Наличие локальной версии и отечественной разработки: возможность развертывания локальной версии PostgreSQL Pro, сертифицированной в РФ и включенной в реестр отечественного ПО Минцифры, обеспечивает дополнительную гибкость и соответствует требованиям информационной безопасности и стратегии импортозамещения.

Разработка базы данных

Далее рассмотрим основные элементы и особенности структуры разработанной базы данных, схема которой изображена в соответствии с рисунком 1.

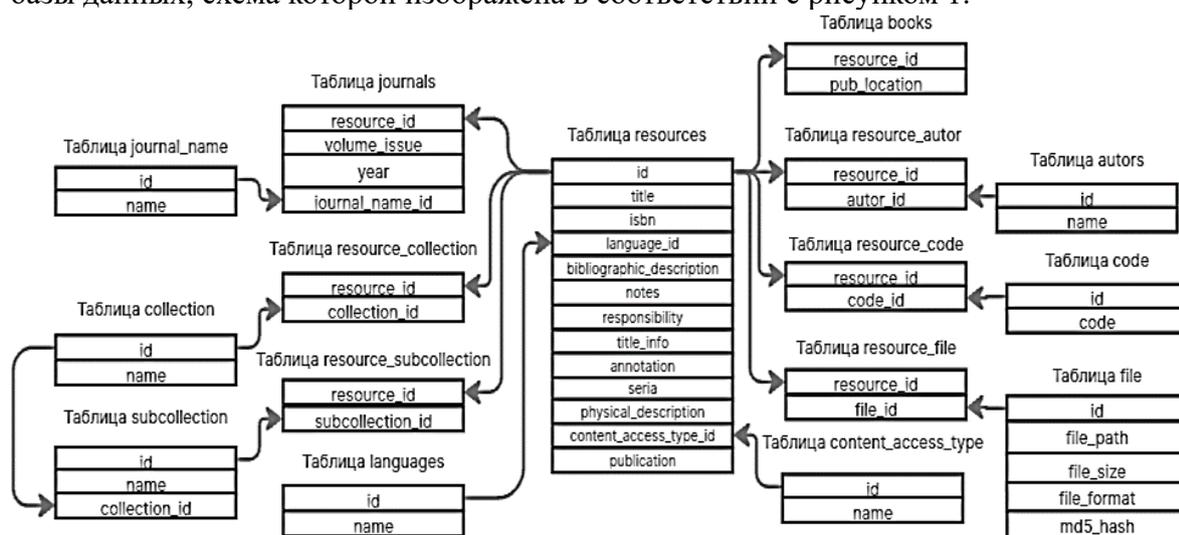


Рисунок 1 – Схема базы данных «Путеводитель по депозитарному фонду библиотеки»

Разработанная база данных состоит из связанных между собой таблиц, которые более подробно рассмотрим далее.

Таблица *resources* является центральной. Основная таблица для хранения метаданных ресурсов. Содержит следующие данные: уникальный идентификатор, основное заглавие, международный стандартный номер, полное библиографическое описание, сведения, относящиеся к заглавию, сведения о серии, аннотацию, идентификатор вида содержания и средства доступа (внешний ключ таблицы ссылается на таблицу *content_access_type*), область физической характеристики, примечания, сведения об ответственности, сведения об издании, язык.

Таблица *authors* содержит информацию об авторах, а именно уникальный идентификатор автора и ФИО автора.

Связующая таблица *resource_actors* делает возможной связь «многие ко многим», что позволяет одной книге иметь несколько авторов и одному автору быть связанным с несколькими книгами. Полная структура сущности состоит из атрибутов *resource_id*, представляющего собой идентификатор ресурса (внешний ключ таблицы ссылается на таблицу *resources.id*) и *actor_id* - идентификатор автора (внешний ключ таблицы ссылается на таблицу *actors.id*).

Связующая таблица *resource_file* делает возможной связь «один ко многим», что позволяет одной книге иметь несколько файлов. Представлена атрибутами *resource_id* и *file_id* (идентификатор коллекции (внешний ключ таблицы ссылается на таблицу *file.id*)).

Таблица *file* содержит информацию о файлах книг: уникальный идентификатор места публикации, путь к файлу, размер файла, формат файла, хеш файла для проверки целостности.

Таблица *resource_collection* связывает книги с коллекциями, что позволяет одной книге быть сразу в нескольких коллекциях и одной коллекции быть связанным с несколькими книгами. Содержит атрибуты *resource_id* и *collection_id* – идентификатор коллекции (внешний ключ таблицы ссылается на таблицу *collection.id*).

Таблица *collection* содержит информацию о коллекциях: уникальный идентификатор коллекции и название коллекции.

Таблица *languages* содержит информацию о языках, на которых написана книга/журнал: уникальный идентификатор языка и язык.

Таблица *subcollection* содержит информацию о подколлекциях: уникальный идентификатор подколлекции, название подколлекции, название коллекции, в которую входит подколлекция (внешний ключ таблицы ссылается на таблицу *collection.id*).

Связующая таблица *resource_subcollection* помогает организовать связь «многие-ко-многим», таким образом одна книга/журнал может быть в нескольких подколлекциях. Содержит атрибуты *resource_id* и *subcollection_id* – идентификатор книги/журнала (внешний ключ таблицы ссылается на таблицу *subcollection.id*)

Таблица *code* предназначена для хранения кода бронирования заказа. Содержит информацию об уникальном идентификаторе бронирования и коде бронирования.

Таблица *resource_code* содержит информацию о том, что входит в заказ пользователя. Представлена атрибутами *resource_id* и *code_id* – идентификатор кода бронирования (внешний ключ таблицы ссылается на таблицу *code.id*).

Таблица *books* содержит уникальные поля для книги: *resource_id* и *pub_location* – сведения о месте издания, издательстве и годе издания.

Таблица *content_access_type* содержит информацию об области вида содержания и средства доступа: уникальный идентификатор и название области

Таблица *journals* содержит информацию об области вида содержания и средства доступа. Содержит атрибуты *resource_id*, *volume_issue* – № тома, выпуска журнала, *year* – год выпуска, *journal_name_id* – идентификатор ресурса (внешний ключ таблицы ссылается на таблицу *journal_name.id*)

Таблица *journal_name* содержит информацию об области вида содержания и средства доступа: уникальный идентификатор и название журнала.

Также при разработке базы данных особое внимание было уделено созданию удобного пользовательского интерфейса для наполнения базы данных. Он спроектирован таким образом, чтобы пользователи могли легко добавлять, редактировать и управлять данными без необходимости обладать глубокими техническими знаниями. Интерфейс значительно упрощает взаимодействие с системой, делая её доступной и интуитивно понятной, что является ключевым элементом для успешной реализации проекта.

Форма разделена на две специализированные части (для книги и журнала/периодики) и общую секцию, применимую к обоим типам.

Часть для книги позволяет указать следующие параметры: полное библиографическое описание (будет формироваться автоматически), заголовок, основное заглавие, сведения, относящиеся к заголовку, сведения об ответственности, сведения об издании, сведения о месте издания, издательстве и годе издания, область физической характеристики, сведения о серии, международный стандартный номер (ISBN), область вида содержания и средства доступа (выбор из вариантов), язык (выбор из вариантов), примечание, аннотация. Часть для журнала/периодики позволяет указать следующие параметры: полное библиографическое описание (будет формироваться автоматически), заголовок, основное заглавие, сведения, относящиеся к заголовку, сведения об ответственности, об издании, область вида содержания и средства доступа (выбор из вариантов), название журнала, год выпуска, № тома/выпуска/журнала, о рубрике, стандартный номер, область физической характеристики, язык (выбор из вариантов), примечание, аннотация. Далее идёт общая секция как для книги, так и для журнала/периодики в ней указывается: коллекция, подколлекция.

Это значительно упрощает процесс наполнения базы данных, так как пользователи могут легко отслеживать различные версии ресурсов, и сохранять целостность данных. Благодаря такой функциональности, система обеспечивает более точную каталогизацию, облегчает поиск информации и поддерживает высокую степень интеграции данных, что особенно важно для библиотек и архивов, работающих с большими объемами материалов. Вверху страницы есть небольшое меню – две кнопки, которые обеспечивают следующий функционал: первая кнопка имеет название «Ввод данных» и обеспечивает переход на страницу ввода, и вторая – «Редактирование и просмотр всех записей» – обеспечивает переход на вторую страницу, где можно просмотреть все записи в БД. Есть также вторая страница создаваемой формы. Основным элементом этой страницы выступают структурированные информационные блоки, которые включают все сведения о записи. Любое поле в карточке можно изменить, нажав на значок карандаша. При редактировании коллекций и подколлекций открывается модальное окно для внесения изменений.

Разработанный удобный пользовательский интерфейс, включая форму для добавления и редактирования данных, значительно упрощает работу с системой, обеспечивая лёгкость управления информацией. Это делает процесс оцифровки и систематизации данных более эффективным и доступным.

Заключение

Авторами разработана база данных, обеспечивающая полноценное хранение и управление метаданными библиотечных ресурсов, таких как книги и журналы. Система позволяет структурированно учитывать авторов, библиографические описания, языковые и физические характеристики изданий, а также связанные с ними файловые материалы. Предусмотрена поддержка коллекций и подколлекций, что создаёт гибкую и удобную структуру организации фонда. Реализованные связи, включая отношения типа «многие ко многим», позволяют эффективно моделировать сложные взаимосвязи между сущностями каталога. База данных интегрируется с программным обеспечением для ввода и управления метаданными, используемым библиотекарями, и обеспечивает функционал бронирования материалов пользователями-читателями. При этом система не содержит персональных данных, что упрощает требования к безопасности. Реализация рассчитана на работу на персональном компьютере с архитектурой x64, что делает систему совместимой с современными рабочими станциями библиотек.

Список литературы

1. Правительство Российской Федерации «Стратегия развития библиотечного дела в Российской Федерации на период до 2030 года»: утверждена распоряжением Правительства Российской Федерации от 13 марта 2021 г. № 608-р. Режим доступа: <http://static.government.ru/media/files/NFWPpXpAAAEbPW60HiZiDvdZZ8AcSNuu.pdf>
2. Перечень поручений по реализации Послания Президента Федеральному Собранию (утв. Президентом Российской Федерации 30 марта 2024 г. № Пр-616). Режим доступа: <https://www.garant.ru/products/ipo/prime/doc/408704827/>
3. Каптерев, А. И. Отражение библиотечной проблематики в региональных стратегиях цифровой трансформации субъектов РФ / А. И. Каптерев, И. П. Тикунова. *Научные и технические библиотеки*. 2025. № 3. С. 161–180. <https://doi.org/10.33186/1027-3689-2025-3-161-180>
4. Миньковская, М. В. Расширение политики цифровизации в экономических процессах России / М. В. Миньковская *Проблемы искусственного интеллекта*. 2024. № 1. С. 75–93
5. Кузнецова, Т. Я. Инновационные библиотечные практики в контексте модернизации отрасли: опыт исторического анализа / Т. Я. Кузнецова. *Научные и технические библиотеки*. 2023. № 4. С. 62–82. <https://doi.org/10.33186/1027-3689-2023-4-62-82>
6. Куулар, М. Ч. Научная библиотека в едином пространстве научных знаний / М. Ч. Куулар, Н. И. Подкорытова. *Труды ГПНТБ СО РАН*. 2024. № 3. С. 20–28.
7. Шрайберг, Я. Л. Информационный рынок, образовательная и библиотечная среда в современном цифровом окружении: новые веяния и ожидаемые результаты (Ежегодный доклад Восьмого международного профессионального форума «Книга. Культура. Образование. Инновации» («Геленджик 2024»)) / Я. Л. Шрайберг *Научные и технические библиотеки*. 2024. № 9. С. 15–65.
8. Волкова, К. Ю. Книжный рынок в эпоху цифровизации: что нужно знать библиотекам. (По материалам Лондонской книжной ярмарки 2019 г.) / К. Ю. Волкова, А. И. Земсков, Я. Л. Шрайберг *Научные и технические библиотеки*. 2020. № 6. С. 49–68.
9. Шрайберг, Я. Л. Электронные библиотеки как главная структурная компонента электронного библиотековедения / Я. Л. Шрайберг *Научные и технические библиотеки*. 2023. № 12. С. 66–96.
10. Максимова, С. В. Новая стратегия развития цифровой интегрированной системы Национальной библиотеки Республики Саха (Якутия) / С. В. Максимова. *Научные и технические библиотеки*. 2024. № 3. С. 127–141.
11. Нестеренок Н. В. Клиентоориентированные библиотечно-информационные услуги: от теории к практике [Электронный ресурс] Режим доступа: <https://rlst.by/informational-resources/izdania/ib-2021/klientoorientirovannye-bibliotechno-informatsionnye-uslugi/>
12. Малиновский, М. П. Научно-техническая библиотека в едином информационном пространстве и бизнес-процессах вуза / М. П. Малиновский. *Научные и технические библиотеки*. 2023. № 4. С. 107–130.
13. Бескаравайная, Е. В. Современные методы хранения данных в библиотеках: ключевые аспекты и решения / Е. В. Бескаравайная, И. А. Митрошин. *Научные и технические библиотеки*. 2024. № 8. С. 78–97.
14. Подсекин, А. М. Повышение эффективности деятельности библиотеки на основе цифровизации / А. М. Подсекин, Т. А. Минеева. *Вестник науки*. 2023. № 12 (69). Режим доступа: <https://cyberleninka.ru/article/n/povyshenie-effektivnosti-deyatelnosti-biblioteki-na-osnove-tsifrovizatsii>
15. Курган, Е. Г. Депозитарный фонд Центральной научно-технической библиотеки Института научно-технической информации Донецкой Народной Республики [Текст] / Е. Г. Курган, В. Н. Никонова, Т. В. Игумнова. *Труды ГПНТБ СО РАН*. 2024. № 3. С. 47–55.
16. Гаркуша, Д. А. Функциональные особенности реализованных онтологических платформ / Д. А. Гаркуша *Проблемы искусственного интеллекта*. 2023. № 4 (31). – С. 4-11
17. Анцыферов, С. С. Методология развития интеллектуальных систем / С. С. Анцыферов, А. С. Сигов, К. Н. Фазилова *Проблемы искусственного интеллекта*. 2022. № 2 (25). – С. 42-47
18. PostgreSQL Global Development Group. PostgreSQL 16 Documentation [Электронный ресурс]. Режим доступа: <https://www.postgresql.org/docs/16/index.html>
19. Microsoft. Документация по SQL Server [Электронный ресурс]. Режим доступа: <https://learn.microsoft.com/ru-ru/sql/sql-server/?view=sql-server-ver17>
20. Oracle. MySQL Reference Manual [Электронный ресурс]. Режим доступа: <https://dev.mysql.com/doc/refman/8.4/en/>.
21. Oracle. Oracle Database Documentation [Электронный ресурс]. Режим доступа: <https://docs.oracle.com/en/database/oracle/oracle-database/index.html>
22. MongoDB, Inc. MongoDB Documentation [Электронный ресурс]. Режим доступа: <https://www.mongodb.com/docs/>.

References

1. Government of the Russian Federation "Strategy for the Development of Librarianship in the Russian Federation for the period up to 2030": approved by Decree of the Government of the Russian Federation dated March 13, 2021, No. 608-R.
<http://static.government.ru/media/files/NFWPpXpAAAEbPW60HiZiDvdZZ8AcSNuu.pdf> Rezhim dostupa - Zaglavie s ekrana.
2. List of instructions for the implementation of the President's Address to the Federal Assembly (approved by the President of the Russian Federation on March 30, 2024, No. Pr-616).
<https://www.garant.ru/products/ipo/prime/doc/408704827/> Rezhim dostupa - Zaglavie s ekrana.
3. Kapterev, A. I. *Reflection of library issues in regional strategies of digital transformation of subjects of the Russian Federation* / A. I. Kapterev, I. P. Tikunova // Scientific and technical libraries. 2025. No. 3. pp. 161-180.
4. Minkovskaya M. V. *Expansion of digitalization policy in the economic processes of Russia* // Problems of artificial intelligence. 2024. No. 1. pp. 75-93
5. Kuznetsova, T. Ya. *Innovative library practices in the context of modernization of the industry: the experience of historical analysis* / T. Ya. Kuznetsova // Scientific and technical libraries. 2023. No. 4. pp. 62-82.
6. Kuular, M. Ch. *Scientific library in a single space of scientific knowledge* / M. Ch. Kuular, N. I. Podkorytova // Proceedings of the State Scientific and Technical University of the Siberian Branch of the Russian Academy of Sciences. 2024. No. 3. pp. 20-28.
7. Schraiberg, Ya. L. *Information market, educational and library environment in the modern digital environment: new trends and expected results* (Annual report of the Eighth International Professional Forum "Book. Culture. Education. Innovations" (Gelendzhik 2024)) / J. L. Schraiberg // Scientific and technical libraries. 2024. No. 9. pp. 15-65.
8. Volkova, K. Y. *The book market in the era of digitalization: what libraries need to know*. (Based on the materials of the London Book Fair 2019) / K. Y. Volkova, A. I. Zemskov, Ya. L. Shraiberg // Scientific and Technical Libraries. 2020. No. 6. pp. 49-68.
9. Shraiberg, Ya. L. *Electronic libraries as the main structural component of electronic library science* / Ya. L. Shraiberg // Scientific and technical libraries. 2023. No. 12. pp. 66-96.
10. Maksimova, S. V. *A new strategy for the development of the digital integrated system of the National Library of the Republic of Sakha (Yakutia)* / S. V. Maksimova // Scientific and technical libraries. 2024. No. 3. pp. 127-141.
11. Nesterenok N. V. *Customer-oriented library and information services: from theory to practice* [Electronic resource] Access mode: <https://rlst.by/informational-resources/izdania/ib-2-2021/klientoorientirovannyye-bibliotечно-informatsionnye-uslugi/>
12. Malinovsky, M. P. *Scientific and technical library in the unified information space and business processes of the university* [Text] / M. P. Malinovsky // Scientific and technical libraries. 2023. No. 4. pp. 107-130.
13. Beskaravainaya, E. V. *Modern methods of data storage in libraries: key aspects and solutions* / E. V. Beskaravainaya, I. A. Mitroshin // Scientific and Technical libraries. 2024. No. 8. pp. 78-97.
14. Podsekin, A.M. *Improving the efficiency of library activities based on digitalization* / A.M. Podsekin, T. A. Mineeva // Bulletin of Science. 2023. № 12 (69). Access mode: <https://cyberleninka.ru/article/n/povyshenie-effektivnosti-deyatelnosti-biblioteki-na-osnove-tsifrovizatsii>
15. Kurgan, E. G. *The Depository Fund of the Central Scientific and Technical Library of the Institute of Scientific and Technical Information of the Donetsk People's Republic* / E. G. Kurgan, V. N. Nikonova, T. V. Igunnova // Proceedings of the State Scientific and Technical University of the Siberian Branch of the Russian Academy of Sciences. 2024. No. 3. pp. 47-55.
16. Garkusha, D. A. *Functional features of implemented ontological platforms* [Electronic resource] / D.A. Garkusha // Problems of artificial intelligence. 2023. No. 4 (31). pp. 4-11.
17. Antsyferov, S. S., Sigov, A. S., Fazilova, K. N. *Methodology of intellectual systems development* [Electronic resource] / S. S. Antsyferov, A. S. Sigov, K. N. Fazilova // Problems of artificial intelligence. 2022. No. 2 (25). pp. 42-47
18. PostgreSQL Global Development Group. *PostgreSQL 16 Documentation*
<https://www.postgresql.org/docs/16/index.html> Rezhim dostupa - Zaglavie s ekrana.
19. Microsoft. *SQL Server documentation* – <https://learn.microsoft.com/ru-ru/sql/sql-server/?view=sql-server-ver17> Rezhim dostupa - Zaglavie s ekrana.
20. Oracle. *MySQL Reference Manual* <https://dev.mysql.com/doc/refman/8.4/en/> / Rezhim dostupa - Zaglavie s ekrana.
21. Oracle. *Oracle Database Documentation* <https://docs.oracle.com/en/database/oracle/oracle-database/index.html> Rezhim dostupa - Zaglavie s ekrana.
22. MongoDB, Inc. *MongoDB Documentation* <https://www.mongodb.com/docs/> / Rezhim dostupa - Zaglavie s ekrana.

RESUME

E. G. Kurgan, K. S. Petrova, A. S. Sobchishen

Development of an electronic database as a tool for digitalization of the Central scientific and technical library

Reference information: the Central Scientific and Technical Library of the Federal State Budgetary Scientific Institution "INTI" in Donetsk lacks the necessary technical equipment to fully digitize its collections, necessitating the development of an interim solution – the electronic database "Guide to the Depository Collection."

Materials and Methods: data from the Central Scientific and Technical Library's depository collection, structured by topic. A comparative analysis of modern database management systems (DBMS), including PostgreSQL, Microsoft SQL Server, MySQL, Oracle Database, and MongoDB, was conducted. The comparison criteria included license type, JSON support, compliance with SQL standards, scalability, extensibility, community activity, and the ability to execute complex queries.

Results: a scalable relational database, "Depository Collection Guide," was developed, consisting of 16 tables with centralized metadata storage and flexible relationships between entities (many-to-many and one-to-many) to accurately reflect library relationships. Efficient processing of complex queries and JSON data was ensured via the PostgreSQL DBMS. An intuitive web interface was created, allowing staff without technical training to easily manage the catalog.

Conclusion: the PostgreSQL-based electronic database is a reliable tool for digitalization of the Central Scientific and Technical Library and the foundation for integrating the library into the unified information space of the Russian Federation. Its architecture and interface allow for further expansion of functionality and the transition to full access to full-text resources.

РЕЗЮМЕ

Е. Г. Курган, К. С. Петрова, А. С. Собчишен

Разработка электронной базы данных как инструмент цифровизации центральной научно-технической библиотеки

Справочная информация: центральная научно-техническая библиотека ФГБНУ «ИНТИ» в Донецке ограничена в технической оснащённости для полного оцифрования фондов, что требует разработки промежуточного решения – электронной базы данных «Путеводитель по депозитарному фонду».

Материалы и методы: данные депозитарного фонда ЦНТБ, структурированные по тематическому принципу. Сравнительный анализ современных систем управления базами данных (СУБД), включая PostgreSQL, Microsoft SQL Server, MySQL, Oracle Database и MongoDB. Критериями для сравнения послужили: тип лицензии, поддержка формата JSON, соответствие стандартам SQL, масштабируемость, расширяемость, активность сообщества и возможности выполнения сложных запросов.

Результаты: разработана масштабируемая реляционная БД «Путеводитель по депозитарному фонду» из 16 таблиц с централизованным хранением метаданных, с гибкими связи между сущностями («многие-ко-многим», «один-ко-многим») для точного отражения библиотечных отношений. Обеспечена эффективная работа со сложными запросами и JSON-данными через СУБД PostgreSQL. Создан интуитивный веб-интерфейс, позволяющий без технической подготовки легко управлять каталогом.

Вывод: электронная база данных на базе PostgreSQL является надёжным инструментом цифровизации ЦНТБ и основой для интеграции библиотеки в единое информационное пространство Российской Федерации. Ее архитектура и интерфейс позволяют дальнейшее расширение функционала и переход к полноценному доступу к полнотекстовым ресурсам.

Курган Елена Геннадьевна – к.э.н., доцент, заместитель директора по научной работе в сфере интеллектуальной собственности, Федеральное государственное бюджетное научное учреждение «Институт научно-технической информации», 283001, Донецк, ул. Артема, д. 60. Эл. почта: kurganelena@mail.ru. *Область научных интересов:* цифровизация, интеллектуальная собственность, научно-технологическое развитие, стратегическое управление, инновационная деятельность.

Петрова Ксения Сергеевна – инженер-программист, отдел «Молодежная научная лаборатория», Федеральное государственное бюджетное научное учреждение «Институт научно-технической информации», Донецк, ул. Артема, д. 60. Эл. почта: petrovakseniya2003@mail.ru. *Область научных интересов:* фронтенд-разработка, проектирование и оптимизация реляционных и нереляционных баз данных, автоматизированные системы управления.

Собчишен Андрей Сергеевич – инженер-программист 1 категории, отдел «Молодежная научная лаборатория», Федеральное государственное бюджетное научное учреждение «Институт научно-технической информации», Донецк, ул. Артема, д. 60. Эл. почта: asobchishen@list.ru. *Область научных интересов:* искусственный интеллект, обработка естественного языка, машинное обучение.

Статья поступила в редакцию 03.09.2025.

К. А. Никитенко, А. В. Звягинцева
ФГБОУ ВО «Донецкий государственный университет»
283001, г. Донецк, ул. Университетская, 24

ОЦЕНКА СЕМАНТИЧЕСКОЙ СХОЖЕСТИ ПРЕДЛОЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ ВЕРОЯТНОСТНЫХ МЕР ЭМБЕДДИНГОВ

K. A. Nikitenko, A. V. Zviagintseva
Federal State Educational Institution of Higher Education «Donetsk State University»
283001, Donetsk, University str, 24

ASSESSMENT OF SEMANTIC SIMILARITY OF SENTENCES USING PROBABILISTIC MEASURES BASED ON EMBEDDINGS

К. А. Нікітенко, Г. В. Звягінцева
ФДБОУ ВО «Донецький державний університет»
283001, м. Донецьк, вул. Університетська, 24

ОЦІНКА СЕМАНТИЧНОЇ СХОЖОСТІ РЕЧЕНЬ З ВИКОРИСТАННЯМ ІМОВІРНІСНИХ МІР ЕМБЕДДИНГІВ

В статье рассматривается задача оценки семантической схожести предложений с использованием вероятностных мер эмбедингов слов. Предложен метод, основанный на сопоставлении вероятностей независимых и зависимых событий, соответствующих словам в предложении. Проведен анализ методов построения матрицы внимания и оцифровки слов, обоснован выбор вероятностных эмбедингов как основы для количественного описания смысловых зависимостей. Предложена процедура сопоставления вероятностей, вычисленных на основе эмбедингов и весов внимания, что позволяет формализовать смысловые связи между словами и предложениями.

Ключевые слова: семантическая схожесть, вероятностная модель, матрица внимания, эмбединг, вероятностный анализ текста.

The article addresses the problem of evaluating the semantic similarity of sentences using probabilistic measures of word embeddings. An method is proposed based on comparing the probabilities of independent and dependent events corresponding to the words within a sentence. The methods for constructing attention matrices and digitizing words are analyzed, and the choice of probabilistic embeddings as the core of the semantic representation is justified. A procedure is proposed for comparing probabilities derived from embeddings and attention weights, which enables formalization of semantic relations between words and sentences.

Key words: semantic similarity, probabilistic model, attention matrix, embedding, probabilistic text analysis.

У статті розглядається задача оцінки семантичної схожості речень з використанням імовірнісних мір ембедингів слів. Запропоновано метод, який засновано на зіставленні ймовірностей незалежних і залежних подій, що відповідають словам у реченні. Проведено аналіз методів побудови матриці уваги та оцифровки слів, обґрунтовано вибір імовірнісних ембедингів як основи для кількісного опису смислових залежностей. Запропоновано процедуру зіставлення ймовірностей, які обчислено на основі ембедингів та ваг уваги, що дозволяє формалізувати смислові зв'язки між словами та реченнями.

Ключові слова: семантична схожість, імовірнісна модель, матриця уваги, ембединг, імовірнісний аналіз тексту.

Введение

Современные методы анализа естественного языка всё активнее используют вероятностные и геометрические модели представления текстовой информации. Наиболее распространенные инструменты семантического анализа – распределенные представления слов, а также механизмы внимания, лежащие в основе архитектуры трансформеров [1]. Эти подходы обеспечивают высокие результаты в задачах машинного перевода, текстовой классификации, поиска и генерации ответов, однако сохраняют ряд теоретических и интерпретационных ограничений [2].

Одной из актуальных проблем современной лингвистической информатики является **оценка семантической схожести предложений**, то есть количественное измерение степени смысловой близости между высказываниями [3], [4]. В классических моделях распределенного представления слов, таких как Word2Vec, GloVe или FastText, семантические связи выражаются через геометрическую близость векторных представлений. Однако подобная близость не всегда имеет прямое вероятностное толкование. В то же время механизм внимания в трансформерах формирует матрицу весов, нормализованных по функции *softmax*, которые могут быть интерпретированы как вероятности условной зависимости между словами в предложении. Таким образом, возникает возможность сопоставления двух вероятностных пространств: одного, заданного эмбедингами слов, и другого, определяемого распределением внимания.

В данной работе предлагается метод, основанный на сравнении **независимых и зависимых** вероятностных событий, соответствующих словам и их совместным появлениям в предложении. Если независимые события можно рассматривать как статистические соотношения слов без смысловой связи, то зависимые отражают их реальную связанность в контексте. Такое сопоставление позволяет количественно оценить, насколько семантическая структура текста отклоняется от случайного распределения, и выявить закономерности смысловой организации предложений.

Цель исследования – разработка и тестирование вероятностного метода оценки семантической схожести предложений, использующего данные матрицы внимания и плотности распределений эмбедингов слов.

Предлагаемый метод позволяет перейти от качественной интерпретации семантических связей к их **количественной вероятностной оценке**, что соответствует вероятностной природе смысловых отношений между языковыми единицами [5].

Обзор методов построения матрицы внимания и оцифровки слов

Механизм внимания является ключевым компонентом современных нейросетевых архитектур, предназначенных для обработки текстов. Его основная идея заключается в том, что модель при анализе каждого слова учитывает влияние всех остальных слов в предложении, взвешивая их вклад в итоговое представление.

Современные подходы к построению матриц внимания включают несколько классов методов. Классический *additive attention* основан на обучении весовой функции, вычисляющей степень соответствия между состоянием декодера и входным вектором. Впоследствии Luong предложил *multiplicative attention*, обеспечивающий более эффективные вычисления за счёт скалярного произведения векторов. Метод *scaled attention*, лежащий в основе архитектуры Transformer, нормализует произведение на размерность скрытого пространства, что улучшает устойчивость при обучении [6].

Позднее были предложены модификации, направленные на сокращение вычислительной сложности и улучшение интерпретируемости – multi-head attention, sparse attention, linear attention, efficient attention, где матрица внимания оптимизируется для ускорения обработки длинных последовательностей.

Методы оцифровки слов также претерпели значительное развитие. К числу базовых относятся Word2Vec, GloVe и FastText, формирующие статические векторные представления слов. Более современные модели, такие как ELMo, BERT, RoBERTa, GPT, обеспечивают контекстно-зависимые представления, что делает возможным учёт смысла слова в конкретном контексте [7]. Для оценки семантической схожести предложений эффективными оказались модели Sentence-BERT [8] и SimCSE [9], использующие эмбединги на уровне предложений. Однако большинство из этих подходов не дают прямой вероятностной интерпретации, что делает актуальным развитие методов, основанных на плотностях распределений эмбедингов [10].

Классическая формулировка механизма внимания впервые была предложена в работах Bahdanau и Luong для машинного перевода, а затем обобщена в архитектуре трансформеров, предложенных Vaswani. В трансформерах внимание реализуется через вычисление матрицы весов [11]:

$$a_{ij} = \text{soft max} \left(\frac{\left((x_i W_q) \cdot (x_j W_k) \right)^T}{\sqrt{d_k}} \right), \quad (1)$$

где a_{ij} – элемент матрицы внимания, показывающий насколько элемент x_i «внимателен» к элементу x_j ; x_i, x_j – векторы эмбедингов слов; W_q, W_k – обучаемые матрицы весов для формирования запроса и ключа соответственно; d_k – размерность вектора ключа.

Нормализация функции *softmax* обеспечивает выполнение условия $\sum_{j=1}^n a_{i,j} = 1$,

что означает: каждая строка матрицы представляет вероятностное распределение, где сумма весов связей слова i со всеми словами предложения равна единице, и позволяет интерпретировать значения a_{ij} как вероятности зависимости между словами w_i и w_j .

Дальнейшее развитие механизма внимания привело к появлению многоголового внимания (multi-head attention), где информация о зависимостях извлекается в нескольких подпространствах признаков, что повышает точность, но затрудняет интерпретацию вероятностных весов.

Для задач вероятностного анализа текстов важно не столько высокое качество генерации, сколько интерпретируемость матрицы внимания. Поэтому в данной работе используется упрощённый вариант механизма самовнимания, основанный на нормализованном косинусном сходстве между векторами слов:

$$P(w_i | w_j) = \frac{\exp \left(\frac{\cos(x_i x_j)}{T} \right)}{\sum_{k=1}^n \exp \left(\frac{\cos(x_i x_k)}{T} \right)}, \quad (2)$$

где $P(w_i | w_j)$ – условная вероятность того, что слово w_i будет «связано» относительно слова w_j ; x_i, x_j – векторные представления слов w_i и w_j соответственно в пространстве признаков; T – параметр «температуры», регулирующий «резкость» распределения; k – порядковый номер слова в анализируемой выборке текста.

Такой способ позволяет получить матрицу вероятностей без обучения модели и сохранить её статистическую интерпретацию. Каждая строка матрицы описывает распределение вероятностей связи данного слова с остальными словами предложения. Таким образом, матрица внимания рассматривается как вероятностная матрица условных зависимостей между словами, что делает её подходящим инструментом для последующего сравнения с вероятностями, вычисленными на плотностях эмбедингов.

Методы оцифровки слов

Основой большинства моделей является гипотеза распределенного значения: слова, встречающиеся в похожих контекстах, имеют сходное значение. Наиболее известными методами векторного представления слов являются Word2Vec, GloVe и FastText. Эти методы создают точечные представления в пространстве высокой размерности, где расстояния между векторами отражают степень семантической близости слов. Современные модели, такие как ELMo, BERT, RoBERTa, формируют контекстно-зависимые представления слов: один и тот же лексемный элемент получает различные векторы в зависимости от окружения. Это повышает точность моделирования смысла, но усложняет вероятностную интерпретацию – эмбединги становятся функцией контекста, а не фиксированным распределением.

Для решения поставленной задачи целесообразно использовать вероятностные эмбединги, где каждое слово описывается гауссовым распределением [12]. Такой подход обеспечивает возможность оценивать вероятности появления отдельных слов, вычислять совместные вероятности групп слов (например, в пределах предложения) и сравнивать их с распределениями, полученными из матриц внимания. Это позволит рассматривать структуру текста как систему вероятностных зависимостей, где геометрические расстояния эмбедингов связаны с вероятностями смысловых связей.

Описание исходных данных

Для экспериментальной проверки предложенного метода использовался корпус более чем из 20 произведений А.С. Пушкина общим объёмом 3105 строк разной длины, содержащий тексты в прозе и стихах, верифицированные по морфологической базе OpenCorpora [13]. Выбор корпуса обусловлен рядом факторов:

- 1) тексты отличаются лексическим и синтаксическим разнообразием, позволяющим анализировать зависимости между словами в разных типах предложений [14];
- 2) корпус удобно структурировать на отдельные предложения и токены, что облегчает формирование выборок для анализа;
- 3) литературный язык А.С. Пушкина хорошо представлен в существующих морфологических и лексических базах данных (таких как OpenCorpora и rusvectors), что обеспечивает корректную токенизацию и морфологическую разметку.

Метод анализа

Предложенный метод основан на сопоставлении вероятностей зависимых и независимых событий, вычисляемых на основе матрицы внимания [15]. Каждое слово рассматривается как элементарное вероятностное событие, а предложение – как их совокупность. На этой основе проводится сравнение распределений, полученных для независимого (частотного) и зависимого (контекстного) случаев, что позволяет оценить степень смысловой когерентности текста [16].

Особенность метода – использование упрощённого механизма самовнимания, где вероятности зависимостей между словами определяются через нормализованные косинусные меры сходства, а эмбединги трактуются как вероятностные распределения

в многомерном пространстве. Такой метод обеспечивает интерпретируемость полученных параметров и их статистическую сопоставимость [17]. Ключевая идея состоит в том, что каждое слово рассматривается как элементарное вероятностное событие, а предложение – как совокупность этих событий. Сравнение независимых и зависимых вероятностных событий проводится по одному и тому же набору слов.

Реализация и расчёты

Пусть корпус содержит m предложений, каждое из которых состоит из n_i слов. Для предложения $S_i = \{w_1, w_2, \dots, w_{n_i}\}$ формируется матрица вероятностей внимания размерности $n_i \times n_i$:

$$A_i = \left[P(w_i | w_j) \right], \quad (3)$$

где $P(w_i | w_j)$ – вероятность связи слова w_i с w_j , определяемая по нормализованным весам внимания.

Параллельно для каждого слова w_j строится вероятностное распределение плотности $p_j(x)$, полученное из выбранной гауссовой модели эмбедингов. На основе распределений вычисляются одномерные и совместные вероятностные события.

Для иллюстрации результатов вычислений на корпусе из произведений А.С. Пушкина построены распределения вероятностных параметров и матрицы внимания.

Для анализа распределения вероятностных соотношений между зависимыми и независимыми событиями вычислялось значение натурального логарифма:

$$\ln \Delta = \ln \frac{P_{dep}}{P_{ind}}, \quad (4)$$

где P_{dep} – вероятность зависимого события, отражающая связь слова с контекстом, P_{ind} – вероятность независимого события (без учёта контекстных зависимостей). Значение $\ln \Delta$ показывает степень смысловой связанности предложений: чем больше величина, тем сильнее выражены семантические зависимости между словами (рис. 1).

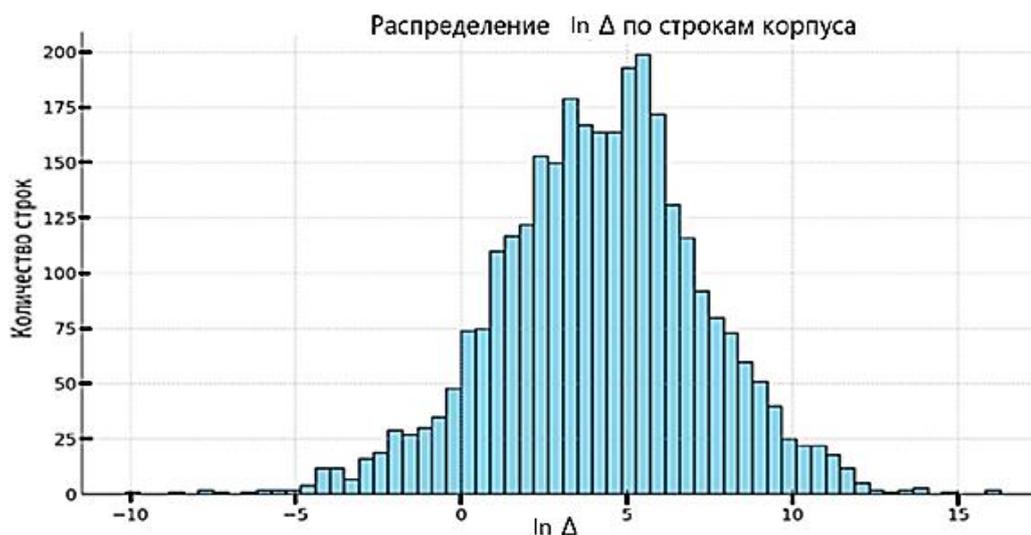


Рисунок 1 – Распределение $\ln \Delta$ по строкам корпуса произведений А.С. Пушкина

Для более детального анализа рассмотрим строку корпуса «По нивам повлечет плуг, миром изощренный», извлечённую из поэмы А.С. Пушкина «Евгений Онегин». Данный фрагмент выбран, поскольку содержит типичную для литературного текста синтаксическую структуру с явно выраженными смысловыми зависимостями между словами, что позволяет наглядно проиллюстрировать работу предлагаемого метода.

Для каждого слова w в корпусе вычислим эмпирическую вероятность $P(w)$:

$$P(w) = \frac{f(w)}{N}, \quad (5)$$

где $f(w)$ – количество вхождений слова w в корпусе, N – общее число токенов (словоформ) в корпусе.

Величина $P(w)$ отражает относительную частоту появления конкретного слова в тексте и используется при оценке независимых вероятностей. Полученные значения для слов из рассматриваемого предложения приведены в таблице 1.

Таблица 1 – Вероятности появления слов в предложении в зависимости от других слов

Позиция	Слово	$P(w)$
1	По	0,002415
2	Нивам	0,000073
3	Повлечет	0,000073
4	Плуг	0,000146
5	Миром	0,000073
6	Изощренный	0,000073

Как упоминалось ранее, матрица внимания (3) построена для выбранной строки корпуса по выражению (2). В данном случае интерес представляют условные вероятности появления слов, вычисляемые в рамках цепочного приближения.

Для каждого слова w_j определяется вероятность его появления при условии предыдущего слова w_{j-1} :

$$P(w_j | w_{j-1}) = \frac{P(w_{j-1}, w_j)}{P(w_{j-1})}, \quad (6)$$

где $P(w_{j-1}, w_j)$ – совместная вероятность появления пары слов; $P(w_{j-1})$ – эмпирическая вероятность предыдущего слова.

Такие условные вероятности позволяют учесть локальные контекстные зависимости между словами внутри предложения. Для рассматриваемой строки корпуса получены результаты, приведенные в таблице 2.

Таблица 2 – Условные вероятности для слов в заданном предложении

Позиция j	Слово w_j	$P(w_j w_{j-1})$
2	Нивам	0,000612
3	Повлечет	0,003428
4	Плуг	0,001514
5	Миром	0,009611
6	Изощренный	0,001610

Для позиции 1 («По») значения предыдущего слова отсутствует, поэтому условная вероятность не определяется и в таблицу не включается.

При предположении статистической независимости слов предложение S можно описать как произведение индивидуальных вероятностей входящих в него слов:

$$P_{ind}(S) = \prod_{j=1}^n P(w_j). \quad (7)$$

Подставляя значения из таблицы 1, получаем:

$$P_{ind}(S) = 0,002415 \cdot 0,000073 \cdot 0,000073 \cdot 0,000146 \cdot 0,000073 \cdot 0,000073 \approx 1,01 \cdot 10^{-23}.$$

В логарифмическом масштабе: $\ln P_{ind} \approx -52,944586$.

Для зависимого (цепного) случая вероятность предложения вычисляется с учётом условных вероятностей из таблицы 2:

$$P_{dep}(S) = P(w_1) \cdot \prod_{j=2}^n P(w_j | w_{j-1}). \quad (8)$$

Подставляя значения, получаем:

$$P_{dep}(S) = P(w_1) \cdot \prod_{j=2}^6 P(w_j | w_{j-1}) = 0,002415 \cdot 0,000612 \cdot 0,003428 \cdot 0,001514 \cdot 0,009611 \cdot 0,001610 \approx 1,19 \cdot 10^{-16}.$$

В логарифмическом виде: $\ln P_{dep} \approx -36,670714$.

Разница между зависимой и независимой вероятностями характеризует степень семантической связанности слов в предложении:

$$\ln \Delta = \ln P_{dep} - \ln P_{ind} = -36,67 - (-52,94) = 16,27.$$

Тогда отношение вероятностей составляет:

$$\frac{P_{dep}}{P_{ind}} \approx e^{16,273872} \approx 1,17 \cdot 10^7.$$

Таким образом, зависимая вероятность рассматриваемой строки примерно в 11,7 миллионов раз выше, чем вероятность того же набора слов при условии их статистической независимости. Полученное значение $\ln \Delta \approx 16,27$ является высоким положительным показателем, что свидетельствует о сильной смысловой связанности слов в данном предложении.

Матрица размером 6×6 (рис. 2) визуализирует нормализованные вероятности условных событий (3). В рассматриваемом примере предложение «по нивам повлечет плуг миром изощренный» содержит $n = 6$ слов, для которых матрица внимания отражает степень вероятностной зависимости каждого слова от отдельных. Ярко выраженные значения на диагонали ($P(w_j) \approx 1$) соответствуют самовниманию – максимальной значимости слова относительно себя. Низкие значения вне диагонали ($P(w_j) \ll 1$) фиксируют слабые связи между различными словами. Таким образом, матрица внимания служит вероятностной аппроксимацией смысловых отношений внутри предложения и позволяет количественно оценить распределение внутренней связи текста на уровне слов.

Для расчёта вероятностей зависимых и независимых событий, отражающих связь слова с контекстом, а также определения степени смысловой связанности предложений разработан скрипт. Для примера в таблице 3 представлены строки корпуса, для которых отношения зависимых и независимых вероятностей принимают наибольшие значения. Эти предложения характеризуются наибольшей семантической когерентностью, что подтверждается высокими положительными значениями $\ln \Delta$.

В таблице 3 приведены длины предложений (n) и рассчитанные значения $\ln(P_{ind})$, $\ln(P_{dep})$ и $\ln \Delta$. Максимальное значение $\ln \Delta = 16,27$ получено для строки «По нивам повлечет плуг миром изощренный», что соответствует примерно $1,17 \cdot 10^7$ – кратному превышению вероятности зависимой модели над независимой.

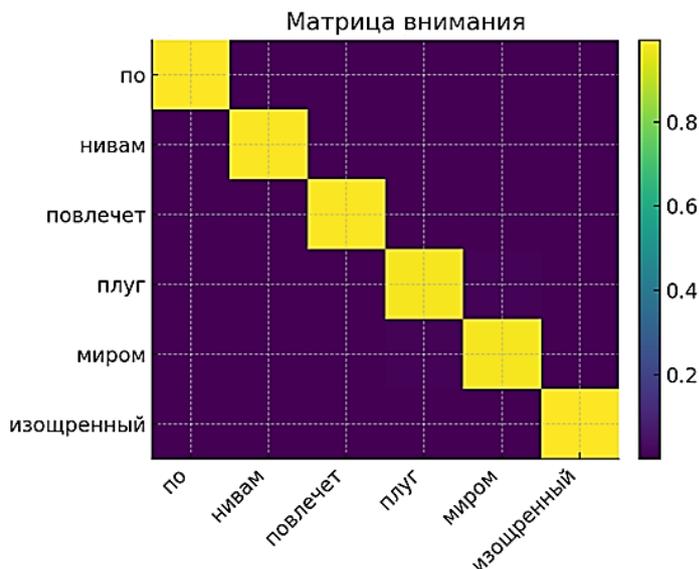


Рисунок 2 – Матрица внимания для выбранной строки

Таблица 3 – Предложения с наибольшей семантической связанностью по значению $\ln\Delta$

№	Предложение	Длина (n)	$\ln(P_{ind})$	$\ln(P_{dep})$	$\ln\Delta$
1	По нивам повлечет плуг, миром изощренный	6	-52,945	-36,671	16,274
2	И громко всем кричал: «Нашел! нашел!»	6	-44,301	-28,376	15,925
3	Держись, держись всегда прямой дороги	5	-43,230	-28,380	14,849
4	И много, много сильных пало	5	-38,468	-24,513	13,955
5	Одной, одной Осгар Мальвиною дышал	5	-42,760	-28,931	13,829
6	Из трубки пенковой дым гонит сероватый	6	-51,723	-38,060	13,663
7	В гордыне возмечтав мечом низвергнуть троны	6	-50,280	-36,902	13,378
8	На лире б возгремел гармонией небесной	6	-47,133	-33,838	13,295
9	Ни чистым золотом набиты сундуки	5	-44,973	-32,035	12,938
10	И месяц, дальних туч покинув темны сени	7	-56,596	-43,863	12,733

Представленные предложения имеют наибольшую семантическую связанность, поскольку их зависимые вероятности значительно превышают независимые, что подтверждается высокими значениями $\ln\Delta$. Во всех случаях наблюдается ярко выраженная структурная и смысловая целостность выражений, в том числе за счёт повторов и устойчивых словосочетаний. Это показывает, что предложенный метод эффективно выявляет фрагменты текста с плотными семантическими связями.

Обобщение результатов практических расчётов

Проведенные вычисления на корпусе произведений А.С. Пушкина подтвердили эффективность предложенного вероятностного метода оценки семантической связанности предложений. Сравнение независимых и зависимых событий показало, что логарифмическая разность их вероятностей распределена близко к нормальному закону, что свидетельствует о статистической устойчивости метода. Наибольшие значения данного показателя фиксируются в предложениях с выраженными смысловыми и грамматическими зависимостями, где наблюдается высокая плотность контекстуальных связей между словами. Таким образом, практические расчёты демонстрируют способность метода адекватно отражать смысловую когерентность текста на уровне предложений.

Матрица внимания позволила формализовать взаимодействие слов в пределах предложения и визуализировать вероятностные зависимости между ними, что сделало возможным количественное описание структуры контекста. В то же время применение вероятностных эмбедингов обеспечило переход от дискретных частотных характеристик к непрерывным распределениям, отражающим семантическую «плотность» слов в пространстве смыслов. Совместное использование этих инструментов дало возможность оценить как локальные связи внутри предложений, так и их глобальную смысловую организацию.

В целом результаты расчётов подтверждают, что интеграция вероятностных представлений слов и механизмов внимания обеспечивает достоверное количественное измерение семантической связанности предложений.

Выводы

В результате проведенного исследования разработан и реализован вероятностный метод оценки семантической связанности предложений, основанный на сопоставлении независимых и зависимых событий в пространстве эмбедингов. Применение матрицы внимания позволило перейти от частотного описания слов к вероятностному моделированию их взаимного влияния в контексте, что обеспечивает более глубокое понимание смысловой структуры текста.

Экспериментальная проверка метода на корпусе произведений А.С. Пушкина показала, что логарифмическая разность вероятностей зависимых и независимых событий подчиняется распределению, близкому к нормальному. Высокие значения этой величины соответствуют предложениям с выраженной смысловой когерентностью, что подтверждает корректность предложенного вероятностного метода к оценке внутренней связанности текста.

Научная новизна работы заключается в объединении вероятностного анализа эмбедингов и механизмов внимания для совместного описания семантических зависимостей. В отличие от традиционных метрик, таких как косинусная близость или частотные показатели, предложенный метод вводит интерпретируемые вероятностные параметры [18], позволяющие количественно оценить переход от независимых слов к зависимым контекстам и представить смысловую структуру текста в вероятностно-геометрических терминах.

Практическая значимость результатов определяется возможностью их применения в задачах:

- автоматической оценки когерентности текста и качества генерации в нейросетевых моделях;
- семантического поиска и кластеризации предложений по смысловой близости;
- выявления аномальных или малоинформативных фраз в корпусах;
- предварительной фильтрации текстовых данных для обучения языковых моделей.

Перспективы дальнейших исследований связаны с расширением методики для анализа масштабных корпусов и интеграцией с современными трансформерными архитектурами (BERT, GPT), что позволит уточнить вероятностные зависимости на уровне контекстуальных эмбедингов. Полученные результаты подтверждают, что вероятностные меры, основанные на матрицах внимания, способны выявлять скрытую структуру текста и формализовать понятие смысловой связанности в терминах сопоставления зависимых и независимых событий, что делает предложенный метод перспективным инструментом семантического анализа и интерпретации моделей.

Список литературы

1. Minaee S., Kalchbrenner N., Cambria E., Nikzad N., Chenaghlu M., Gao J. Deep Learning-Based Text Classification: A comprehensive review // *ACM Computing Surveys*. no 54(3), 2021: 1–40.
2. Аверин Г.В. О вероятностной природе смыслов в дискретных языковых единицах // *Системный анализ и информационные технологии в науках о природе и обществе*. №1(12)–2(13), 2017. С. 11–18.
3. Андриевская Н.К. Гибридная интеллектуальная мера оценки семантической близости // *Проблемы искусственного интеллекта*. №1(20), 2021. С. 4–17.
4. Меры семантической близости в онтологии / К.В. Крюкова, Л.А. Панкова, В.А. Пронина, В.С. Суховеров, Л.Б. Шипилина // *Проблемы управления*. Вып. 5, 2010. С. 2–14.
5. Erk K. The probabilistic turn in semantic and pragmatics // *Annu. Rev. Linguist.* 2022. 8:101–21.
6. Zheng Z., Wang Y., Huang Y., Song S., Yang M., Tang B., Xiong F., Li Z., Attention Heads of Large Language Models: A Survey // *arXiv – 2024 – arXiv:2409.03752*.
7. Собчишен А.С., Звягинцева А.В. Вероятностно-смысловые модели оцифрованных текстовых данных // *Материалы конференции «Математическое и компьютерное моделирование в экономике, страховании и управлении рисками»*. №9, 2024. – С. 55–59.
8. Reimers N., Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks // *arXiv 2019 arXiv:1908.10084*.
9. Gao T., Yao X., Chen D. SimCSE: Simple Contrastive Learning of Sentence Embeddings // *arXiv – 2021 arXiv:2104.08821*.
10. Shen L., Jiang H., Liu L., Shi S. Sen2Pro: A Probabilistic Perspective to Sentence Embedding from Pre-trained Language Models // *arXiv – 2023 – arXiv:2306.02247*.
11. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., Polosukhin I. Attention is all you need // *arXiv – 2017 – arXiv:1706.03762*.
12. Yoda S., Tsukagoshi H., Sasano R., Takeda K. Sentence Representations via Gaussian Embedding // *arXiv – 2023 – arXiv:2305.12990v2*.
13. Сайт проекта «Открытый корпус» (OpenCorpora) русского языка. – Электр. рес. – URL: <https://opencorpora.org/> (01.11.2025).
14. Chun S., Joon S., Sampaio R., Kalantidis Y., Larlus D. Probabilistic Embeddings for Cross-Modal Retrieval // *arXiv – 2021 – arXiv:2101.05068*.
15. Abdelwahab A., Landwehr N. Deep distributional sequence embeddings based on a Wasserstein loss // *arXiv – 2019 – arXiv:1912.01933*.
16. Botha J. Probabilistic modeling of morphologically rich languages // *arXiv – 2015 – arXiv:1508.04271*.
17. Бондаренко В.И., Елисеев В.О., Ермоленко Т.В. Анализ эффективности глубоких языковых моделей для задачи определения тональности русскоязычных текстов // *Проблемы искусственного интеллекта*. №1(32), 2024. – С. 51–62.
18. Никитенко К.А., Звягинцева А.В. Интерпретируемость нейросемантических моделей при их применении в прикладных областях // *Проблемы искусственного интеллекта*. №2(37), 2025. С. 79–90.

References

1. Minaee S., Kalchbrenner N., Cambria E., Nikzad N., Chenaghlu M., Gao J. Deep Learning-Based Text Classification: A comprehensive review // *ACM Computing Surveys*. no 54(3), 2021: 1–40.
2. Averin G.V. O veroyatnostnoi prirode smyslov v diskretnykh yazykovykh edinitсах [On the Probabilistic Nature of Meanings in Discrete Linguistic Units]. *Sistemnyi analiz i informatsionnye tekhnologii v nauках o prirode i obshchestve*. no 1(12)–2(13), 2017: 11–18.
3. Andrievskaya N.K. Gibridnaya intellektual'naya mera ocenki semanticheskoy blizosti // *Problemy iskusstvennogo intellekta*. №1(20), 2021: 4–17.
4. Kryukova K.V., L.A. Pankova, V.A. Pronina, V.S. Sukhoverov, L.B. Shipilina. Mery semanticheskoy blizosti v ontologii [Measures of semantic proximity in ontology] *Probl. upravl.*, 2010. Issue 5: 2–14.
5. Erk K., The probabilistic turn in semantics and pragmatics // *Annu. Rev. Linguist.* 2022. 8:101–21.
6. Zheng Z., Wang Y., Huang Y., Song S., Yang M., Tang B., Xiong F., Li Z., Attention Heads of Large Language Models: A Survey // *arXiv – 2024 – arXiv:2409.03752*.
7. Sobchishen A.S., Zviagintseva A.V. Veroyatnostno-smyslovyye modeli otsifrovannykh tekstovykh dannykh [Probabilistic-Semantic Models of Digitized Text Data]. *Materialy konferentsii “Matematicheskoe i komp'yuternoe modelirovanie v ekonomike, strakhovanii i upravlenii riskami”*. no 9, 2024: 55–59.

8. Reimers N., Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks // arXiv – 2019 – arXiv:1908.10084.
9. Gao T., Yao X., Chen D. SimCSE: Simple Contrastive Learning of Sentence Embeddings // arXiv – 2021 – arXiv:2104.08821.
10. Shen L., Jiang H., Liu L., Shi S. Sen2Pro: A Probabilistic Perspective to Sentence Embedding from Pre-trained Language Models // arXiv – 2023 – arXiv:2306.02247.
11. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., Polosukhin I. Attention is all you need // arXiv – 2017 – arXiv:1706.03762.
12. Yoda S., Tsukagoshi H., Sasano R., Takeda K. Sentence Representations via Gaussian Embedding // arXiv – 2023 – arXiv:2305.12990v2.
13. Sait proekta “Otkrytiy korpus” (OpenCorpora) russkogo yazyka [Project website “OpenCorpora” of the Russian language]. – Elektronnyy resurs. URL: <https://opencorpora.org/> (accessed November 1, 2025).
14. Chun S., Joon S., Sampaio R., Kalantidis Y., Larlus D. Probabilistic Embeddings for Cross-Modal Retrieval // arXiv – 2021 – arXiv:2101.05068.
15. Abdelwahab A., Landwehr N. Deep distributional sequence embeddings based on a Wasserstein loss // arXiv – 2019 – arXiv:1912.01933.
16. Botha J. Probabilistic modeling of morphologically rich languages // arXiv – 2015 – arXiv:1508.04271.
17. Bondarenko V.I., Eliseev V.O., Ermolenko T.V. Analiz effektivnosti glubokikh yazykovykh modelei dlya zadachi opredeleniya tonal'nosti russkoyazychnykh tekstov [Analysis of the Effectiveness of Deep Language Models for Sentiment Analysis of Russian Texts]. Problemy iskusstvennogo intellekta. no 1(32), 2025: 51–62.
18. Nikitenko K.A., Zviagintseva A.V. Interpretiruemost' neurosemanticheskikh modelei pri ikh primeneni v prikladnykh oblastiakh [Interpretability of Neurosemantic Models in Applied Domains]. Problemy iskusstvennogo intellekta. no 2(37), 2025: 79–90.

RESUME

K.A. Nikitenko, A.V. Zviagintseva

Assessment of semantic similarity of sentences using probabilistic measures based on embeddings

Modern semantic analysis in natural language processing widely relies on vector representations of words and attention-based neural architectures. However, most existing approaches describe semantic relations geometrically and do not provide explicit probabilistic interpretation of meaning. The purpose of this paper is to develop a method for evaluating semantic similarity of sentences based on the comparison of independent and context-dependent probabilistic events derived from word embeddings and attention weights.

The study introduces a probabilistic method in which words are treated as random events, and sentences are represented as distributions over independent and dependent probabilities. Probabilistic embeddings are used to model lexical meaning, while an attention-based mechanism is applied to estimate contextual dependencies between words. A comparative analysis of these probability spaces makes it possible to quantify semantic coherence at the sentence level.

The proposed method allows formalizing semantic relations between words and reveals clear distinctions between random word co-occurrence and contextually conditioned connections. The analysis shows that sentences with stronger internal semantic structure demonstrate significantly higher coherence within the probabilistic model, confirming the adequacy of the approach for interpreting meaning in text.

The developed method provides an interpretable probabilistic alternative to purely geometric similarity metrics and can be applied in semantic search, text clustering, and evaluation of generated texts. Future work includes extending the method to larger corpora and integrating contextualized embeddings from transformer models to refine probabilistic estimation of semantic relations.

РЕЗЮМЕ

К.А. Никитенко, А.В. Звягинцева

*Оценка семантической схожести предложений
с использованием вероятностных мер эмбедингов*

Современные методы обработки естественного языка активно используют рас-
пределённые представления слов и механизмы внимания, однако большинство под-
ходов описывает смысловые связи только геометрически, без явной вероятностной
интерпретации. Актуальной задачей является разработка методов, позволяющих фор-
мализовать степень смысловой связанности текста и оценивать схожесть предложений с
опорой на статистические свойства языковых моделей. Цель работы – предложить
вероятностный подход к оценке семантической схожести, основанный на сопоставлении
независимых и контекстно-зависимых событий, соответствующих словам в предложении.

В работе предложен вероятностный метод оценки семантической схожести
предложений, основанный на сопоставлении независимых и зависимых событий, со-
ответствующих словам в предложении. Метод объединяет вероятностные распределе-
ния плотностей эмбедингов и матрицы внимания, что позволяет количественно
описать степень смысловой связанности текста.

Результаты анализа показывают, что предложенный метод дает возможность
различать случайные сочетания слов и контекстно обусловленные связи, что делает
возможной формализацию семантической структуры текста в вероятностных терми-
нах. Метод демонстрирует устойчивое поведение при анализе корпуса и корректно
выделяет предложения с выраженной смысловой связностью.

Предложенный метод позволяет формализовать понятие семантической ко-
герентности текста в вероятностных терминах и применять его в задачах семанти-
ческого поиска, кластеризации предложений, оценки связности автоматически
генерируемых текстов и анализа интерпретируемости нейросетевых моделей.

Никитенко Кирилл Андреевич – аспирант кафедры компьютерных технологий
ФГБОУ ВО «ДонГУ», ул. Университетская, 24, Донецк, 283001, Россия, n1kitenkok@yandex.ru.

Область научных интересов: компьютерное зрение, машинное обучение,
нейронные сети. Число научных публикаций – 5.

Звягинцева Анна Викторовна – д.т.н., доцент, профессор кафедры
компьютерных технологий ФГБОУ ВО «ДонГУ», ул. Университетская, 24, Донецк,
283001, Россия, zviagintsevaav@gmail.com. *Область научных интересов:* системный
анализ, событийная и комплексная оценка; безопасность и управление социально-
экономическими и техногенными системами; информационно-аналитические
системы; обработка и анализ данных. Число научных публикаций – более 150.

Статья поступила в редакцию 29.08.2025

Д.А. Филипишин, С.А. Зори
Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий национальный технический университет», г. Донецк, 283001, г. Донецк, ул. Артема, 58

МЕТОД РЕЗОЛЮЦИИ КАК СРЕДСТВО РАСШИРЕНИЯ МАШИНЫ ВЫВОДА ОНТОЛОГИЙ*

D.A. Filipishin, S.A. Zori
Federal State Budgetary Educational Institution of Higher Education "Donetsk National Technical University", 283001, Donetsk, Artema str, 58

THE RESOLUTION METHOD AS A MEANS OF EXTENDING THE ONTOLOGY INFERENCE MACHINE

В статье рассматривается применение решателя языка Prolog как средства расширения машины вывода (ризонера) для построения онтологий предметных областей. Предложен новый подход к решению продукционных задач в рамках редактора онтологий, рассматриваемого как интегрированная среда разработки (IDE). Проанализирована перспектива использования решателя языка Prolog, основанного на языке предикатов математической логики дизъюнктов Хорна, семантическими связями и аксиоматикой бинарных отношений, свойственных онтологиям. Предложено решение имеющихся недостатков в обеих системах с целью создания новой онтологической системы, включающей возможности обоих подходов.

Ключевые слова: онтология, Prolog, IDE, reasoner, решатель, продукции, загадка Эйнштейна, игра Холмс.

The article discusses the use of the Prolog language solver as a means of extending the inference machine (reasoner) to build domain ontologies. A new approach to solving production tasks within the framework of the ontology editor, considered as an integrated development environment (IDE), is proposed. The perspective of using a Prolog language solver based on the predicate language of the mathematical logic of Horn disjuncts, semantic connections and axiomatics of binary relations inherent in ontologies is analyzed. A solution to the existing shortcomings in both systems is proposed in order to create a new ontological system that includes the capabilities of both approaches.

Key words: ontology, Prolog, IDE, reasoner, solver, products, Einstein's Riddle, the Holmes game.

* Работа выполнена в рамках молодежной научной лаборатории «Искусственный интеллект» (Код НИОКТР FRFR-2024-0010)

Введение

Любой язык программирования всегда ориентирован на определенный круг задач, при решении которых он наиболее эффективен. Для языка Prolog типичными являются проекты, связанные с разработкой систем искусственного интеллекта – это различные экспертные системы, системы планирования, программы-переводчики, интеллектуальные игры и т.п. Prolog обладает достаточно мощными средствами, позволяющими извлекать информацию из баз данных и знаний. При этом его методы поиска принципиально отличаются от «традиционных». Кроме этого, Prolog часто используют в задачах, связанных с манипулированием на естественном языке [1].

Существует не меньше десятка зарубежных инструментов онтологического инжиниринга, поддерживающих формализмы для описания знаний и использующих машины вывода из этих знаний. Для анализа возможностей инструмента онтологии был выбран редактор онтологий Protégé. Этот редактор имеет исчерпывающую документацию, поддерживается значительным сообществом, состоящим из разработчиков и ученых, правительственных и корпоративных пользователей, использующих его для решения задач, связанных со знаниями. Protégé доступен для свободного скачивания с официального сайта вместе с плагинами и онтологиями [2].

Protégé – локальная, свободно распространяемая Java-программа, предназначенная для создания, просмотра и редактирования онтологий явных моделей предметной области с целью включения этих моделей в программный код. Редактор онтологий Protégé позволяет разворачивать иерархическую структуру абстрактных или конкретных классов и слотов. Структура онтологии аналогична иерархической структуре каталога. На основе онтологии Protégé может генерировать формы получения знаний для введения экземпляров классов и подклассов. Имеет графический интерфейс. Поддерживает использование языка OWL. Генерирует HTML-документы, отображающие структуру онтологий. Использует фреймовую модель представления знаний ОКВС – позволяет редактировать модели предметных областей, представленных в форматах UML, XML, SHOE, DAML+OIL, RDF / RDFS и т.п. (а не только в не в OWL) [3].

Таким образом, Prolog – это язык и система логического программирования, основанные на языке предикатов математической логики дизъюнктов Хорна [4]. Хорновский дизъюнкт [4] в контексте математической логики — это дизъюнктивный одночлен с не более чем одним положительным литералом. Литерал в теории булевых функций и логике высказываний — булева формула, имеющая вид x (истина) или \bar{x} (ложь) для некоторой переменной x . Хорновские дизъюнкты могут быть пропозициональными формулами, либо формулами первого порядка, в зависимости от того, рассматриваются ли пропозициональные литералы или литералы первого порядка.

Резолюция цели с определённым дизъюнктом для получения новой цели является основной для правила вывода в SLD-резолюции (*Selective Linear Definite resolution*), используемого для реализации логического программирования и языка программирования Пролог. В логическом программировании определённый дизъюнкт используется как процедура редукции цели. Например, дизъюнкт $\neg p \vee \neg q \vee \dots \vee \neg t \vee u$ работает как процедура: «чтобы показать u , показать p, q, \dots и t ».

Решатель (программа для решения задач) в языке Prolog устроен на основе фактов, правил и запросов. Prolog автоматически управляет решением задачи, стараясь найти все возможные наборы значений, удовлетворяющие запросу. Для это используется метод поиска с возвратом. Рассмотрим особенности и недостатки языка Prolog.

Prolog не алгоритмический, программа в нем не содержит явных алгоритмических конструкций типа условных или циклических операторов. Не выполняет команды по порядку, а строит выводы на основе правил, условий или функций без состояния. Это требует другой модели мышления. Один и тот же код может вести себя по-разному из-за порядка строк, типов или связи с платформой, что усложняет отладку. Механизм вывода (решатель) пытается найти решение «в лоб», чем часто вызывает критические ошибки исполнения.

Онтологическая система представляет собой инструмент поиска скрытых знаний на основе семантических связей и множества аксиом, заданных на отношениях. Редактор онтологий Protégé, являющийся наиболее перспективным представителем программной системы, основанной на дескриптивном описании онтологии (OWL), не имеет в своем арсенале средств решения продукционных задач.

Таким редактором выступает система Ontorion Fluent Editor, от польских разработчиков, которая позволяет управлять онтологией скриптовым описанием с использованием реального английского языка, а также использовать «if... then...» для фактов, в соответствии с рисунком 1.

```
If an airplane has-value-mcr lower-or-equal-to 0.7 and a wing-profile has-value-c equal-to 15 then the wing-profile useds-on the airplane .
```

Рисунок 1 - Пример использования «if then» и союза «and» в качестве конъюнкции [5]

Ризонер (машина вывода в онтологии) на текущий момент не умеет обрабатывать множественные решения и, несмотря на концепцию об открытости мира, требует значительного уточнения для своей работы. Таким образом, является актуальной задача разработки инструмента построения дескриптивных онтологий (OWL) с использованием решателя языка Prolog, методом резолюций.

Цель работы и постановка задачи

Целью работы является анализ практического расширения машины вывода ризонера в редакторе онтологий, представленного как интегрированная среда разработки (IDE), методом резолюции, т.е. средствами решателя Prolog. Пример решения продукционной задачи продемонстрирован разбором решения общеизвестной задачи-игры «HOLMES», в соответствии с рисунком 2, которая по сути является усложненной «загадкой Эйнштейна», в соответствии с рисунком 3, имеющей дополнительные ограничения на расстановку.

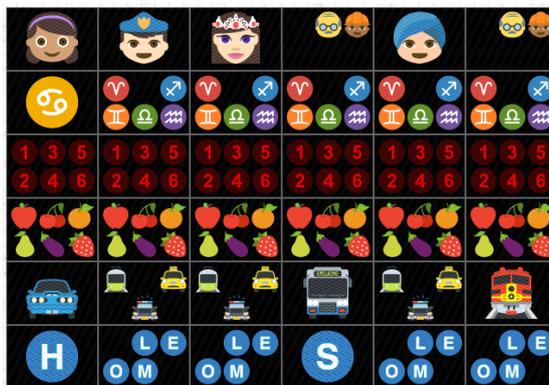


Рисунок 2 - Пример продукционной задачи - игра Шерлок в режиме сложности «легко»

Задача Эйнштейна (загадка Эйнштейна, головоломка о зебре) – логическая задача, по легенде созданная Альбертом Эйнштейном в годы его детства. Также бытует мнение, что она использовалась Эйнштейном для проверки кандидатов в ассистенты на способность к логическому мышлению. Доказательств авторства задачи нет. Иногда автором головоломки называют Льюиса Кэрролла, но в условии упоминаются марки сигарет, которые не существовали при жизни Кэрролла и во времена детства Эйнштейна.

```
rdfs:comment [language: ru]
На улице стоят пять домов.
Англичанин живёт в красном доме.
У испанца есть собака.
В зелёном доме пьют кофе.
Украинец пьёт чай.
Зелёный дом стоит сразу справа от белого дома.
Тот, кто курит Old Gold, разводит улиток.
В жёлтом доме курят Коол.
В центральном доме пьют молоко.
Норвежец живёт в первом доме.
Сосед того, кто курит Chesterfield, держит лису.
В доме по соседству с тем, в котором держат лошадь, курят Коол.
Тот, кто курит Lucky Strike, пьёт апельсиновый сок.
Японец курит Parliament.
Норвежец живёт рядом с синим домом.
```

Кто пьёт воду? Кто держит зебру?

Рисунок 3 - Пример «загадки Эйнштейна»

Для решения «загадки Эйнштейна» необходимо создать 5 классов по 5 индивидов к каждому из них, задать 6 бинарных отношений со свойствами «functional» и «inverse functional», связывающих классы между собой, описать перечисления индивидов для каждого из класса, указать что сущности не пересекаются между собой («different» и «disjoint»), задать аксиоматику и указать факты какое «животное» у кого имеется и т.п., а также добавить посылки для исключения вариантов бездействия индивидов. Предлагаемая система модификаций редактора онтологий в интегрированную среду проектирования (IDE) позволяет устранить ряд недостатков: каждый индивид создается как экземпляр класса (не нужно назначать класс после его создания), исключает необходимость указывать перечисления индивидов для классов, вся основная логика семантических связей переносится на аксиоматику отношений (свойства domain и range), посылкой является сам запрос, но при наличии достаточных оснований для рассуждения (фактов). Такой эффект достигается методом «сужения» слишком открытого мира в Protégé, с доработкой существующих инструментов построения онтологий и машины вывода.

Формальная модель онтологии в рамках предлагаемой интегрированной среды разработки (IDE) представляет собой дополненный кортеж формальной модели онтологии в моделях, основанных на дескриптивной логике (OWL):

$$O = \langle C, R, H_C, H_R, I, A, H_A \rangle, \quad (1)$$

где C – множество классов; $R: C \times C$ – множество свойств классов; $H_C: C \times C$ – иерархия классов; $H_R: R \times R$ – иерархия отношений; I – множество индивидов; A – множество аксиом над классами и ограничений отношений; H_A – множество ограничений аксиом отношений. Отличие формальных моделей OWL [6] и предлагаемой (в рамках IDE) в наличии (H_A) множества ограничений не только для отношений, а и аксиом, т.е. наличия множества условий для семантической связи одного отношения для разных пересечений классов, что на сегодняшний день нет ни в одном из редакторов онтологий.

Таким образом, при обработке запроса «Аня имеет флешку» или вопрос на естественном языке «Какую флешку имеет девочка Аня?», при описании онтологии

класс-«девочка»/индивид-«Света» → предикат-«имеет» (domain «девочка» - range «флешка» - свойство «not functional» или «один к одному») → класс-«флешка» и наличия фактов наличия флешек у девочек «Аня» имеет НЕ-«красная», «Варя» имеет НЕ-«синяя» и «Света» имеет НЕ-«красная» и НЕ-«белая», обратится к имеющемуся множеству индивидов «флешка» и методом жадного алгоритма получит ответ «белая».

Индивиды для класса «девочка»: Аня, Варя, Света. Индивиды для класса «флешка»: синяя, белая, красная. Согласно аксиоматике отношения «имеет» в онтологии сказано: одна «флешка» принадлежит одной «девочке». Ответ получен так: у Светы «синяя», у Ани НЕ-«красная», а значит «синяя» или «белая», «синяя» у Светы, значит у Ани «белая». Из примера видно, что посылку (аксиому) «девочка» имеет «флешка» или «ЛЮБАЯ девочка имеет хотя бы одну флешку» более задавать не нужно. Также система работает с тем набором данных, который добавлен в множество классов-концептов.

Такой подход практически делает невозможным применение существующего инструмента онтологической системы, а именно хранение онтологии в домене и подключение метаонтологии, как например DBpedia [7]. Но значительно упрощает процесс построения онтологий и снижает порог вхождения для обычного пользователя. Рассмотрим краткое решение задачи-игры «HOLMES», для анализа необходимых средств при решении подобных продукционных задач (Рис.2).

Решатель языка Prolog и предлагаемое решение задачи-игры «HOLMES»

Рассмотрим формальный подход методу резолюции. Базовой конструкцией теории является схема отношения, выступающая некоторым аналогом классического предиката [8].

Выражение вида [8]:

$$f: a_1, \dots, a_n \rightarrow a_0, \quad (2)$$

где $a_i, i = \overline{1, n}$ – имена величин, называется функциональной связью (ФС). В записи (2) f – это имя ФС, $a_i, i = \overline{1, n}$ – аргументы, a_0 – результат ФС.

При задании интерпретации [3] структура вида:

$$\begin{aligned} S(r) = (r)(S_{01}(a_{01}), \dots, S_{0N0}(a_{0N0}), \\ \text{if } p_1 \Rightarrow S_{11}(a_{11}), \dots, S_{1N1}(a_{1N1}) \dots \\ p_k \Rightarrow S_{k1}(a_{k1}), \dots, S_{kNk}(a_{kNk}) \text{ fi} | \text{filter} \end{aligned} \quad (3)$$

фиксирует некоторое подмножество размеченного декартова произведения [8], определяемое ФС из набора *filter*. Подразумевается, что в подмножество попадают только те кортежи (наборы) атрибутов схемы, значения которых удовлетворяют всем ФС этой схемы. Неформально, мы можем ставить два типа вопросов: (i) удовлетворяет ли набор конкретных значений атрибутов схеме, и (ii) какие наборы удовлетворяют схеме, если зафиксированы значения некоторых из атрибутов. Принимаются явные соглашения, связанные с использованием предопределенных отношений (порядка и т.д.) для первичных схем [4].

Таким образом, иерархия типов строится в виде совокупности объявлений в форме [4], близкой к фиксации башни функционалов, когда каждый элемент явно дефинируется предками:

$$\begin{aligned} S_0(r_0) \leftarrow (r_0). (S_{01}(r_{01}), \dots, S_{0k}(r_{0k})); \\ (r_0). S_{01}(r_{01}) \leftarrow (r_0. r_1). (S_{11}(r_{11}), \dots); (r_0). S_{0k}(r_{0k}) \leftarrow \dots \end{aligned} \quad (4)$$

При этом полагают, что в терминальных конструкциях все $S_{ij} \in E$ («построение от базиса»); считается также, что в записи $S(r)$, синтаксически r представляет собой место индивидуальной подстановки, а семантически форма сообщает, что конкретизация r удовлетворяет отношению S (набор *filter* индуцирует функцию, принадлежности [9]). Важно отметить, что применение для спецификации программы логического языка даже первого порядка неизбежно приводит нас к проблеме, связанной с результатом Геделя. На практике это вызывает необходимость искусственного торможения резолютивного вывода (управления последним) с помощью внелогических операций, подобных операции усечения («cut» или «!») в языке Пролог. Очевидная эклектичность вызывает не менее очевидное недоверие к инструменту.

Предлагается следующая трактовка [4] стандартной клаузы языка Пролог (интерпретация фактов и целевого утверждения остается стандартной). Пусть имеет место:

$$S(r): -S_1(r_1), \dots, S_m(r_m), Q_1(t_1), Q_N(t_N), \quad (5)$$

где $S_i(r_i), i = \overline{1, M}$, удовлетворяет (3), т.е. выступают «аналогами» классических предикатов, $Q_j(t_j), j = \overline{1, N}$ имеют форму ФС, причем для любого $j = \overline{1, N}, t_j \subseteq \cup r_i (i = \overline{1, M})$. Приведенные соглашения практически полностью покрывают синтаксис любой стандартной Пролог-системы.

Таким образом, продемонстрирована возможность сведения стандартных предложений языка Пролог к конструкциям языка теории С-моделей. В [12] показана полнота этой теории, приведена схема алгоритма логического вывода, а также установлено, что сложность такого алгоритма не превосходит $O(|N|^3)$, где N – константа, соответствующая длине исходной спецификации. Более того, используемый алгоритм реализует обратный вывод в стиле С.Ю. Маслова (см. [12], Приложение 1).

Как уже говорилось выше, действия решателя языка Prolog напрямую зависят от порядка описания фактов, чего нет у ризонера, т.к. механизм вывода онтологий основан на аксиоматике отношений и порядок фактов не играет существенной роли, определяя лишь порядок загрузки фактов в базу знаний (справедливо только для Ontorion Fluent Editor, в Protégé нет скриптового управления и факты добавляются вручную каждому индивиду). В современных программных системах с элементами искусственного интеллекта [4] обязательной составляющей является подсистема объяснения полученного вывода, а отсутствие в алгоритме бэк-трекинга делает процесс объяснения более простым и естественным по сравнению с традиционным подходом.

Вернемся к продукционной задаче-игре «Холмс». Эта игра была создана по мотивам олдскульной игры "Шерлок", первоначально написанной Эвереттом Кейзером для MSDOS более 20 лет назад. Шерлок – это компьютеризированная версия логических головоломок, в которой вам предлагается ряд подсказок, которые помогут вам определить точное расположение всех изображений на игровом поле. Каждый ряд игрового поля содержит однотипные изображения (лица, дома, цифры, фрукты, уличные знаки, буквы и т.д.). Компьютер скремблирует расположение элементов в каждом ряду (не показывая вам их расположения), а затем представляет вам набор графических подсказок, описывающих расположение различных изображений. Вы используете подсказки, чтобы определить, где чего быть не может (и где что должно быть), пока не узнаете, где находятся все изображения.

Данная задача описана матрицей 6 на 6, в каждой ячейке которой по итогу должен находиться только 1 элемент. Расстановка элементов осуществляется с помощью продукционных правил сопоставления («в одном столбце» или НЕ-«в одном столбце») и расстановки (элемент «рядом» и НЕ-«рядом», в соответствии с рисунком 4б, «в одном»

или НЕ-«в одном» столбце (Рис.4а), а также описание смежности в соответствии с рисунком 4б, т.е. расстановки 3х элементов рядом друг с другом, которое может быть прямым и обратным, и правило следования «символ предшествует символу»).

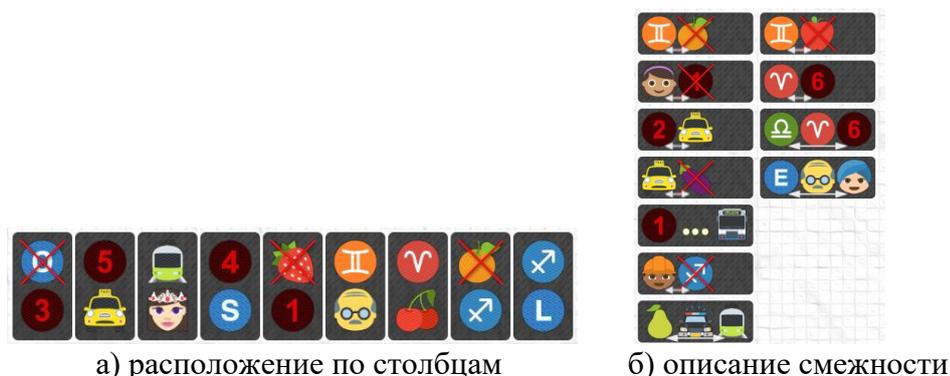


Рисунок 4 - Продукционные правила для решения задачи-игры Шерлок

В «загадке Эйнштейна», отношение «в одном столбце» описано как «мужчина курит сигареты» или «мужчина держит животное», отношение смежности и «рядом» описано также, но конкретно для каждого индивида или класса. В целом можно считать, эти две загадки равноценными задачами построения онтологии, за исключением применения правил исключения вариантов и потребности разрешения противоречивых вариантов для игры «шерлок» даже в режиме сложности «легко».

Приведенные на рисунке 4 ограничения легче описать в виде предикатов или фактов, нежели разбирать на аксиомы и искусственно ограничивать пространство решения для ризонера в онтологии. Для решения задачи «Шерлок» в режиме «сложно», в котором изначально не заполнена ни одна из ячеек, жизненно необходимо «пробовать» подставлять условия в виду наличия достаточного количества противоречивых подстановок, переключая возможность «ошибаться» на машину вывода. Потому как описать множественное решение согласно приведенным в игре ограничениям имеющимися инструментами редактора онтологий Protégé или Ontorion Fluent Editor не является возможным. Для этого следует применять механизм обратного вывода в стиле С.Ю. Маслова (см. [12], Приложение 1), на основе решателя языка Prolog. Но не просто скопировать решатель у языка Prolog, а расширить имеющуюся машину вывода включающей возможности обеих систем (ризонер в Protégé и решатель в Prolog) продукционными правилами (if...then...) как в редакторе онтологий Ontorion Fluent Editor.

Таким образом, полученная система вывода позволит не просто перебирать множество предикатов согласно описанным аксиомам онтологии, а и реагировать на дополнительные ограничения применяя механизм обратного вывода, свойственно Prolog. И как уже было сказано ранее, вводимый механизм также потребует решения проблемы, связанной с результатом Геделя. А именно необходимостью продумать систему управления машиной перебора, например, аналогично языку Prolog, условием отсечения:

Name=diana, #условие отсечения
!. #отсечение

Рассмотрим одно из возможных решений задачи-игры «Шерлок», в соответствии с рисунком 2. Перед началом разбора, рассмотрим предложенную матрицу на предмет подготовленных «ответов» (единичных значений в представленных ячейках). Как

видим, в соответствии с рисунком 2, в нашем случае, уже единично определены: «девочка», «полицейский», «царевна», «индус», «рак», «машина», «автобус», «поезд», «Н» и «S». Эти элементы будут служить опорными точками для применения указанных, в соответствии с рисунком 4 ограничений.

Далее, согласно правилу расположения «по середине», т.е. элементы, которые указаны в последовательностях в середине, можно убрать с краев матрицы. Для представленного примера это «Овен» и «Мужчина». Правила расположения, в соответствии с рисунком 4а, являются взаимозаменяемыми, например, «О» не может находиться в одном столбце с «З», что означает и обратное отношение: «З» не может находиться в одном столбце с «О». Правила расстановки в соответствии с рисунком 4б, также являются взаимозаменяемыми. Например, рядом с «Близнецами» НЕ-может находиться «апельсин», что также означает: рядом с «апельсином» НЕ-может находиться «Близнецы» (т.е. слева или справа).

Далее проверим соответствия по столбцам: «трамвай – царевна», «4 - S» и «Близнецы - Мужчина». После чего можно снова приступить к правилам расстановки. Рядом с «Близнецы» уберем «апельсины» и «яблоки». Также нам известно, где находится «девочка», поэтому уберем рядом «1».

Следуя правилам расстановки, предположим, что мы не знаем расположение «2», «такси», «патиссон». «Автобус» занимает свое единичное место, поэтому можно четко определить, что «1» справа от него быть не должно, в соответствии с рисунком 5. Рядом с «шахтер» не может быть «стрелец», но под ним «стрелец» находиться может. Правило «груша – полицейская машина - трамвай» также можно однозначно определить, в виду того, что «трамвай» единично определен и рядом с ним «полицейская машина» может находиться только слева. Отметим «груша» и «полицейская машина».



Рисунок 5 - Правило расстановки-следования: «1» предшествует «автобус»

Таким образом, методом исключения нам стало известное расположение «такси». Отметим «2» справа, т.к. слева находится «4» и уберем по бокам «патиссон». Также в столбце с «такси» отметим «5» согласно правилу расположения столбцов «5 - такси». Далее, согласно правилу «вишня - овен», можно удалить «вишня» там, где нет «Овен». А именно, там, где расположен «Близнецы». Также можно убрать «стрелец», там, где нет «L» и наоборот. Выполним правило расстановки «Е – мужчина - индус», т.к. мы знаем расположение «мужчины» и «индуса». Здесь же можно убрать «стрелец», т.к. в этом столбце больше нет «L».



Рисунок 6 - Расстановка оставшихся элементов

Далее, исходя из правила «Овен» рядом с «6», уберем «Овен», который рядом с «4» и «2». Также уберем «вишня» там, где нет «Овен». Согласно правилу «весы – овен - 6», можно убрать «Весы», которые находятся справа от «Близнецы», т.к. этот знак должен быть рядом с «Овен».

Теперь мы знаем расположение «Стрелец». Отметим «L» и уберем «апельсин». Далее необходимо совместить два правила «овен - б» и «весы – овен - б», чтобы оба правила подходили (Рис. 6). Согласно описанным правилам, мы можем расположить только в обратном порядке «б – овен - весы», т.к. в прямом нет возможности отметить «б». После чего, отметим «вишня» в столбце с «Овен», уберем «клубника» в столбце с «1» и, по итогу, уберем «О» в столбце с «3». Загадка решена (Рис. 7). Буквы внизу составляют слово «HOLMES», т.е. название игры.

Также можно ускорить процесс решателя добавив правило исключения, а именно убирать первый символ, который должен быть «в том же столбце» в тех местах, где уже нет второго символа сверху или снизу, и наоборот. Правила можно объединять, для случаев наличия смежных расстановок, через какой-то общий символ, т.е. если в подмножестве невозможна какая-то пара, то следует исключить всю цепочку из этого столбца.

Сюда же относится правило исключения подстановок, когда система может не применять правила смежности по ряду или столбцу, а просматривать количество смежных ячеек, в которых можно применить рассматриваемое правило смежности. Если количество этих ячеек равно 1, то следует применить это правило для этой ячейки как единственно подходящий вариант.

Для правил расстановки (3х символов) можно применять правило, не задумываясь в том случае, если один из крайних символов является единично определенным и находится в крайнем столбце. Тогда остальные символы просто больше некуда записывать.

Решение производственных задач в рамках предлагаемого подхода сводится к добавлению продукции «if...then...» для фактов в онтологии и дальнейшего перебора условий с возвратом. Также необходимо помечать примененные условия, с целью уменьшения множества оставшихся не выполненных условий. Дополнительно следует реализовать алгоритм, проверяющий достоверность полученных результатов, чтобы в каждой «ячейке» был единичным только тот символ, которого нет в других ячейках в этом же ряду.

Приемы решения логических задач базируются на выдвижении некоторой гипотезы и ее подтверждении (или опровержении) согласно некоторой методике. Таким образом соответствуют вычислению ответа на основе связанной многократной работы двух модулей: построение некоторого состояния, проверка полученного состояния. Подход вычисления ответа, а точнее – построения ответа согласуется с решением задач на языке Prolog [17].

Машинный способ решения описанной задачи заключается в циклической проверке правил до тех пор, пока хотя бы одно правило было применено из перечисленных множеств, в нашем случае это 2 множества правил. Тип множества правил последовательно переключается с целью поиска подходящего решения. В случае, если в ходе перебора правил система не может найти решения ни в одном из множеств правил, следует применить принудительную подстановку из имеющихся (не выполненных) правил расстановки или ином доступном, если все правила расстановки уже были задействованы. Далее проверить правила сходимости относительно других правил и, в случае обнаружения несовместимости, отменить принятое ранее решение и ряд изменений, после чего применить следующее правило расстановки и повторить проверку совместимости до достижения желаемого результата. Такой способ не лишен недостатков и требует значительной доработки, но позволяет решить большую часть поставленных производственных задач в рамках семантической модели.

Использование языка Пролог в качестве базы позволяет использовать его в качестве интеллектуального («разумного») инструмента решения задач, в виду доступности синтаксиса и простоты интерпретации семантических конструкций. Работы в этом направлении ведутся со времени основания логического языка, и на сегодняшний день разработано множество программ (баз знаний фактов и правил), которые развиваются вместе с прикладными научными направлениями [17], [19-22].

Чтобы закодировать в терминах языка Пролог процедуру проверки на соответствие (согласование), нужно описать правилами все требования, предъявляемые к искомому состоянию. Примеры таких процедур фактов и правил рассмотрены в различных прикладных работах по логическим языкам [18], [19].

Для перебора элементов внутри множеств был предложен несколько иной подход к использованию конструкции циклического перебора на базе существующей конструкции «*for..in*», записываемый ключевым словом «*loop*» в правилах семантики внутреннего скриптового языка.

Рассмотрим пример предлагаемой конструкции цикла:

```
bool k = false;
loop(rule in _rules; k == true; k = false){
    if( is_acceptable(rule) ) k = true;
}
```

где *rule* - имя переменной, которая является указателем на объект внутри массива или кортежа; *_rules* – массив или кортеж объектов для перебора; «*k==true*» - условие последующего выполнения, которое выполняется аналогично конструкции «*do while*»; «*k = false*» - инструкция, которая выполняется в конце каждой итерации.

Главным отличием от стандартной конструкции *for* является механизм перебора всех элементов внутри переданного массива (или кортежа) за каждую итерацию. Проверяется наличие элементов внутри массива (или кортежа), выполняется инструкция шага, выполняется тело циклической конструкции, проверяется условие продолжения и цикл повторяется до тех пор, пока выполняется условие продолжения или не будет вызван оператор «*break*». Данный тип циклической конструкции позволяет значительно эффективнее обрабатывать множество правил целиком, нежели существующие виды циклов. Эта особенность в значительной мере позволяет упростить логику работы с элементами множеств, но также добавляет ряд недостатков, которые могут быть связаны с необходимостью последовательной обработки каждого элемента, аналогично стандартному циклу «*for..in*» или «*for..of*», что вероятно потребует ввода ряда дополнительных переменных.

Таким образом, можно подвести итог, что рассмотренный пример решения задачи-игры «Холмс», которая является несколько усложненным вариантом «загадки Эйнштейна» (общеизвестным примером для обучения построения онтологий OWL), наглядно демонстрирует потребность внедрения в машину вывода онтологии (ризор) механизм перебора с возвратом, аналогичный решателю языка Prolog. Это позволит вывести процесс построения онтологий предметных областей на новый уровень и решать более сложные задачи искусственного интеллекта, нежели на сегодняшний день способен OWL и OWL2.

Заключение

В статье рассмотрена проблема отсутствия необходимых инструментов в онтологии для решения продукционных задач. В качестве решения описанной проблемы предложено расширить машину вывода редактора онтологий, представленного как интегрированная среда разработки (IDE), с помощью метода резолюций, аналогичного решателю языка Prolog.

К перспективам развития работы можно отнести внедрение в онтологии не рассмотренные в данной статье инструменты языка Prolog, позволяющие проектировать экспертные системы, с целью реализации инструментов принятия решений и дополнения инструмента построения онтологий предметных областей с «собственным поведением».

Список литературы

1. Хабаров С.П. Интеллектуальные информационные системы. PROLOG-язык разработки интеллектуальных и экспертных систем: учебное пособие. СПб. СПбГЛТУ, 2013. 138 С.
2. Цуканова, Н.И. Онтологическая модель представления и организации знаний: учебное пособие для вузов. М.: Горячая линия – Телеком, 2015. 272 с.
3. Гаркуша, Д.А. Функциональные особенности реализованных онтологических платформ // Проблемы искусственного интеллекта. 2023. № 4 (31). 4-11. <http://search.rads-doi.org/project/14374/object/210537> doi: 10.34757/2413-7383.2023.31.4.001.
4. Новосельцев В.Б. Эффективный нерезолютивный вывод для ограниченного исчисления хорновских дизъюнктов // Известия Томского политехнического университета, 2008. Т. 312 №5. С. 94 – 97.
5. Боргест Н.М., Орлова А.А. Онтологический редактор Fluent Editor: учебно-методическое пособие к лабораторным работам / сост.: Н.М. Боргест, А.А. Орлова. Самара: Изд-во Самарского университета, 2017. 44 С.
6. Mendes P. N. DBpedia: a multilingual cross-domain knowledge base / P. N. Mendes, M. Jakob, and C. Bizer // LREC. 2012. Pp. 1813–1817.
7. Цуканова, Н.И. Онтологическая модель представления и организации знаний: учебное пособие для вузов. М.: Горячая линия – Телеком, 2015. 272 С.
8. Адаменко А.Н., Кучуков А.М. Логическое программирование и Visual Prolog. СПб.: БХВ-Петербург, 2003. 992 с., ил. ISBN 5-94157-156-9
9. Вирт Н. Структурное программирование. М.: Мир, 1975. 189 С.
10. Клини С. Математическая логика. М.: Мир, 1973. 480 С.
11. Новосельцев В.Б. Теория структурных функциональных моделей // Сибирский математический журнал. 2006. Т. 47. - № 5. С. 1014 1030.
12. Косарев, Н.И. Продукционная модель представления знаний в системах поддержки принятия решения // Вестник Сибирского юридического института ФСКН России №2 (13), 2013.
13. Дорохина Г. В. Требования к информационной технологии цифрового сбора, обработки и анализа данных. Проблемы искусственного интеллекта. 2020. № 4 (19). С. 4–9.
14. Bruno Borlini Duarte An Ontology-based Reference Model for the Software Systems Domain with a focus on Requirements Traceability/ Bruno Borlini Duarte. Vitória, ES, 2022- 149 p.: il.; 30 cm. (дата обращения: 04.12.2025).
15. Гаркуша, Д. А. Анализ онтологических платформ. Искусственный интеллект: теоретические аспекты, практическое применение: материалы Донецкого международного научного круглого стола. Донецк: ФГБНУ «ИПИИ», 2023. 252 с. С. 37–40.
16. Загорулько Ю.А. Современные средства формализации семантики областей знаний на основе онтологий. Информационные и математические технологии в науке и управлении. 2018. № 3 (11). С. 27 36. DOI:10.25729/2413-0133-2018-3-03.
17. Половикова О.Н. Особенности программной реализации логических задач на языке Prolog / О.Н. Половикова, В.В. Ширяев, Н.М. Оскорбин, Л.Л. Смолякова // Известия АлтГУ. Математика и механика, 2021, №1 (117).
18. Лорьер Ж.Л. Системы искусственного интеллекта. URL: http://lib.alnam.ru/book_sii.php (дата обращения: 04.12.2025).
19. Математическая логика и логическое программирование // Математический форум MathHelpPlanet. URL: <http://mathhelpplanet.com/static.php?p=matematicheskaya-logika-i-logicheskoye-programmirovaniye>. (дата обращения: 04.12.2025).
20. Решение логических задач на Prolog // Блог программиста: программирование и алгоритмы (версия от 28.05.2018). URL: <https://pro-prof.com/archives/1299>. (дата обращения: 04.12.2025).
21. Santos Costa V. On Just in Time Indexing of Dynamic Predicates in Prolog // Lecture Notes in Computer Science. Berlin, Heidelberg, 2009. DOI:10.1007/978-3-642-04686-5.
22. Повторение и рекурсия. Откат // Проект: project:prolog:povtorenie_i_rekursija. URL: http://verim.org/project/prolog/povtorenie_i_rekursija. (дата обращения: 04.12.2025).

References

1. Khabarov S.P. Intelligent information systems. PROLOG is a language for developing intelligent and expert systems: a textbook / S.P. Khabarov. - St. Petersburg State Technical University, 2013. – 138P.
2. Tsukanova, N.I. The ontological model of knowledge representation and organization: a textbook for universities / N.I. Tsukanova, Moscow: Hotline – Telecom, 2015. – 272P.
3. Garkusha, D.A. Functional features of implemented ontological platforms // D.A. Garkusha // Problems of artificial intelligence. - 2023. № 4 (31). - 4-11. - <http://search.rads-doi.org/project/14374/object/210537> doi: 10.34757/2413-7383.2023.31.4.001.
4. Novoseltsev V.B. Effective nonresolutionary derivation for the limited calculus of Horn disjuncts // Izvestiya Tomsk Polytechnic University, 2008. – Т. 312 №5. – Pp. 94 – 97.
5. Borgest N.M., Orlova A.A. Ontological editor Fluent Editor: an educational and methodological guide to laboratory work / comp.: N.M. Borgest, A.A. Orlova. Samara: Publishing House of Samara University, 2017. – 44P.
6. Mendes P. N. DBpedia: a multilingual cross-domain knowledge base [Текст] / P. N. Mendes, M. Jakob, and C. Bizer // LREC. – 2012. – Pp. 1813–1817.
7. Tsukanova, N.I. The ontological model of knowledge representation and organization: a textbook for universities / N.I. Tsukanova, Moscow: Hotline – Telecom, 2015. – 272P.
8. Adamenko A.N., Kuchukov A.M. Logical programming and Visual Prolog. - St. Petersburg: BHV-Petersburg, 2003. – 992P. ISBN 5-94157-156-9
9. Wirth N. Structural programming. Moscow: Mir, 1975. – 189P.
10. Klini S. Mathematical Logic, Moscow: Mir, 1973. – 480P.
11. Novoseltsev V.B. Theory of structural functional models // Siberian Mathematical Journal. – 2006. – Т. 47. - № 5. – Pp. 1014 – 1030.
12. Kosarev, N.I. The production model of knowledge representation in decision support systems / N.I. Kosarev // Bulletin of the Siberian Law Institute of the Federal Drug Control Service of Russia №2 (13), 2013.
13. Dorokhina G. V. Information technology requirements for digital data collection, processing and analysis. Problems of artificial intelligence. 2020. № 4 (19). Pp. 4–9.
14. Bruno Borlini Duarte an Ontology-based Reference Model for the Software Systems Domain with a focus on Requirements Traceability/ Bruno Borlini Duarte. – Vitória, ES, 2022- 149 p.: il.; 30 cm. (date of request: 04.12.2025).
15. Garkusha, D. A. Analysis of ontological platforms. Artificial intelligence: theoretical aspects, practical application: proceedings of the Donetsk International Scientific Round Table. Donetsk: FGBNU "IPII", 2023. 252P. Pp. 37–40.
16. Zagorulko Yu.A. Modern means of formalizing the semantics of knowledge domains based on ontologies. Information and mathematical technologies in science and management. 2018. № 3 (11). С. 27 36. DOI:10.25729/2413-0133-2018-3-03.
17. Polovikova O.N. Features of software implementation of logical tasks in the Prolog language / O.N. Polovikova, V.V. Shiryayev, N.M. Offending, L.L. Smolyakova // News of AltSU. Mathematics and Mechanics, 2021, №1 (117).
18. Laurier J.L. Artificial intelligence systems. URL: http://lib.alm.ru/book_sii.php (date of request: 04.12.2025).
19. Mathematical logic and logical programming // Mathematical Forum MathHelpPlanet. URL: <http://mathhelpplanet.com/static.php?p=matematicheskaya-logika-i-logicheskoye-programmirovaniye>. (date of request: 04.12.2025).
20. Solving logical problems in Prolog // Programmer's blog: programming and algorithms (version from 28.05.2018). URL: <https://pro-prof.com/archives/1299>. (Date of request: 04.12.2025).
21. Santos Costa V. On Just in Time Indexing of Dynamic Predicates in Prolog // Lecture Notes in Computer Science. Berlin, Heidelberg, 2009. DOI:10.1007/978-3-642-04686-5.
22. Repetition and recursion. Rollback // Project: project:prolog:povtorenie_i_rekursija. URL: http://verim.org/project/prolog/povtorenie_i_rekursija. (date of request: 04.12.2025).

RESUME

D.A. Filipishin, S.A. Zori

The resolution method as a means of extending the ontology inference machine

At the moment, the existing description of domain ontology construction approaches does not have the tools to solve production problems using the reverse inference method. Each task, the purpose of which is to create and implement an ontology in information systems, software complexes and other software products, requires analysis and identification of the purpose of building an ontology, as well as development or involvement of third-party tools using the ontology as a knowledge base.

To work effectively within the framework of the ethnographic approach (end-user orientation), it is necessary to use the developed ontology comprehensively and as limited as possible from third-party software products. The work involves a number of modifications of the ontological platform for solving programming problems.

РЕЗЮМЕ

Д.А. Филипишин, С.А. Зори

Метод резолюции как средство расширения машины вывода онтологий

На данный момент существующее описание подходов построения онтологий предметных областей не имеет инструментов решения производственных задач методом обратного вывода. Каждая задача целью которой является создание и внедрение онтологии в информационные системы, программные комплексы и другие программные продукты, вызывает необходимость проводить анализ и выделять цель построения онтологии, а также разрабатывать или привлекать сторонние средства, использующие онтологию как базу знаний.

Для эффективной работы в рамках этнографического подхода (ориентирования на конечного пользователя) необходимо использовать разработанную онтологию комплексно и максимально ограничено от сторонних программных продуктов. Работа предполагает ряд модификаций онтологической платформы для решения задач программирования, в том числе ряд заимствований возможностей у инструментов языка логического программирования Пролог.

Филипишин Дмитрий Александрович – ассистент кафедры программной инженерии им. Л.П. Фельдмана ФГБОУ ВО «Донецкий национальный технический университет», научный сотрудник молодежной научной лаборатории «Искусственный интеллект». *Область научных интересов:* онтологии, системы искусственного интеллекта; эл. почта domaco@mail.ru, адрес: 283001, г. Донецк, ул. Артема, д. 58, телефон +7 949 334 91 49.

Зори Сергей Анатольевич – доктор технических наук, доцент, заведующий кафедрой программной инженерии им. Л.П. Фельдмана ФГБОУ ВО «Донецкий национальный технический университет». *Область научных интересов:* системы искусственного интеллекта; системный анализ; эл. почта ik.ivt.rec@mail.ru, адрес: 283001, г. Донецк, ул. Артема, д. 58, телефон +7 949 315 43 43.

Статья поступила в редакцию 11.10.2025

УДК 004.932.4

DOI 10.24412/2413-7383-2025-4-39-254-265

И. В. Чернядьев, Т. В. Ермоленко
Федеральное государственное бюджетное научное учреждение
«Институт проблем искусственного интеллекта» г. Донецк,
283048, г. Донецк, ул. Артема, 118б

УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ С ПРИМЕНЕНИЕМ ГЛУБОКОГО ОБУЧЕНИЯ: АНАЛИЗ И ПЕРСПЕКТИВНЫЕ РАЗРАБОТКИ*

I. V. Chernyadev, T. V. Yermolenko
Federal State Budgetary Scientific Institution
"Institute of Artificial Intelligence Problems" Donetsk,
283048, Donetsk, Artyoma str., 118b

IMAGE IMPROVEMENT USING DEEP LEARNING: ANALYSIS AND PROSPECTIVE DEVELOPMENTS

В статье рассматривается актуальная проблема улучшения качества изображений с использованием методов глубокого обучения, с акцентом на super-resolution, шумоподавление и устранение артефактов. Анализируются ключевые архитектуры на основе Transformer и сверточных нейронных сетей, их преимущества и недостатки, такие как высокая ресурсоемкость и узкая специализация. Предлагаются улучшения в виде рандомизированной аугментации данных для синтеза обучающей выборки и замены стандартной свертки на быструю свертку Фурье для захвата глобального контекста. Целью работы является повышение качества восстановления изображений без высоких вычислительных затрат.

Ключевые слова: super-resolution, нейронные сети, компьютерное зрение, глубокое обучение, обработка изображений.

The article discusses the current problem of improving image quality using deep learning methods, with a focus on super-resolution, noise reduction, and artifact removal. It analyzes key architectures based on transformers and convolutional neural networks, their advantages and disadvantages, such as high resource consumption and narrow specialization. The article proposes improvements in the form of randomized data augmentation for synthesizing the training set and replacing the standard convolution with Fast Fourier Convolution to capture global context. The goal of the work is to improve the quality of image restoration without high computational costs.

Keywords: super-resolution, neural networks, computer vision, deep learning, image processing.

* Исследование проводится при поддержке гранта РНФ (проект № 25-21-20602).

Введение

На сегодняшний день существует потребность в автоматизации технических процессов, связанных с улучшением качества изображений, в том числе, с повышением разрешения и сохранением детализации. Изображения низкого разрешения часто страдают от таких явлений, как артефакты, шум, размытие, что ограничивает их практическую ценность.

Задача повышения качества изображений, включая super-resolution, приобретает все большую актуальность ввиду высокого роста объемов цифрового контента и требований к его качеству в различных отраслях. В спутниковой съемке методы повышения разрешения востребованы для мониторинга сельского хозяйства, реагирования на стихийные бедствия, где низкокачественные изображения могут быть преобразованы в высокодетализированные для более эффективного анализа. В индустрии видеоигр super-resolution оптимизирует производительность графических процессоров, позволяя рендерить кадры в низком разрешении с последующим улучшением до высокого качества.

В отличие от классических алгоритмов, нейронные сети обучаются адаптироваться к разнообразным искажениям, демонстрируя превосходство в задачах super-resolution, шумоподавления, восстановления утраченных деталей. Однако, несмотря на прогресс, остаются нерешенные проблемы. Во-первых, большинство моделей требуют огромных вычислительных ресурсов и объемов данных для обучения. Во-вторых, существующие алгоритмы часто демонстрируют нестабильность при работе с изображениями, содержащими сложные комбинации искажений, например, шум, размытие, низкая освещенность. В-третьих, вопросы обеспечения устойчивости моделей к артефактам генерации остаются открытыми. Вышеперечисленное дает простор для новых исследований и улучшения существующих прикладных решений.

Целью данной работы является определение основных проблем в задаче улучшения качества изображений, а также в выявлении ключевых особенностей применения глубокого обучения в этой предметной области. В работе были поставлены следующие задачи:

1. Провести анализ основных подходов к улучшению качества изображений.
2. Выявить основные проблемы в данной задаче.
3. Выделить ключевые особенности применения глубокого обучения к улучшению качества изображений и предложить методы устранения выявленных недостатков.

Постановка задачи улучшения качества изображений

Обозначим $I_{LR} \in \mathbb{R}^{h \times w \times c}$ – входное изображение низкого разрешения (LR) с высотой h , шириной w и количеством каналов c ; $I_{HR} \in \mathbb{R}^{H \times W \times c}$ – выходное изображение высокого разрешения (HR) с высотой H и шириной W .

В качестве входных данных для задач улучшения качества изображений, как правило, используются пары изображений I_{HR} и I_{LR} для обучения моделей в режиме с учителем. Выходные данные представлены в виде изображений высокого разрешения с восстановленными деталями, сниженным шумом и минимизированными артефактами, обычно в формате с увеличенным количеством пикселей. Задачи в улучшении качества изображений условно можно разделить на два типа:

1. Super-resolution с сохранением деталей. Принцип работы состоит в том, чтобы на основе LR-изображения сгенерировать HR-версию, максимально близкую к реальному высококачественному аналогу, с акцентом на восстановление текстур, границ и мелких элементов.

2. Комплексное устранение искажений. По входному LR-изображению с комбинацией шумов, размытия и других артефактов модель стремится восстановить чистую HR-версию, учитывая разнообразные сценарии реальных искажений.

1 Архитектуры глубоких сетей, используемые в super-resolution

1.1 Анализ глубоких архитектур для задачи повышения дискретизации изображений

Повышение разрешения является распространенной процедурой в цифровой обработке изображений и видео. Она выполняется всякий раз, когда осуществляется масштабирование изображения на экране мобильного устройства или воспроизведение видеопотока с исходным разрешением 480p на дисплее с разрешением 1080p. Традиционные методы повышения дискретизации не ориентированы на получение высокоточных предсказаний внешнего вида изображения при увеличенном разрешении. Их основное назначение заключается в обеспечении корректного отображения визуальных данных после масштабирования. Подобные алгоритмы применяются, например, при необходимости адаптации изображения к размерам экрана монитора или при подготовке его к печати, когда требуется изменение пространственного разрешения. Несмотря на то, что такое повышение разрешения технически увеличивает количество пикселей, это по-прежнему картинка с низкой детализацией, которая просто просматривается в более высоком формате с некоторым сглаживанием. В данной работе основное внимание уделяется анализу эффективности методов повышения дискретизации изображений с использованием моделей глубоких нейронных сетей.

Среди существующих архитектурных подходов особое внимание привлекают модели SRCNN, SRResNet, SRGAN, а также решения, основанные на архитектуре Transformer. Каждая из них характеризуется специфическими особенностями и принципами построения, определяющими эффективность восстановления высокочастотных деталей изображения. Несмотря на хорошие результаты, эти подходы сталкиваются с различными ограничениями: от проблем масштабируемости и вычислительной сложности до артефактов генерации и зависимости от больших объемов обучающих данных. Ниже приведено описание технологических особенностей существующих решений, а также основные недостатки этих архитектур:

– SRCNN (*Super-Resolution Convolutional Neural Network*). Ключевым преимуществом SRCNN является ее простота – сеть содержит всего три сверточных слоя, что обеспечивает быстроту обработки. Эксперименты показывают превосходство над бикубической интерполяцией [1]. Из недостатков стоит отметить, что структура модели не позволяет захватывать сложные иерархические особенности изображений, что проявляется в размытии текстур при масштабировании [2].

– SRResNet (*Super-Resolution Residual Network*) представляет собой глубокую нейронную сеть, специально разработанную для сверхвысокого разрешения. Она включает в свою архитектуру остаточные блоки и пропускные соединения из архитектуры ResNet, позволяющие эффективно обучать глубокие нейронные сети, избегая проблемы затухающих градиентов [3]. К недостаткам нейросети можно отнести однородную обработку всех регионов, независимо от их сложности, приводит к недостаточной детализации текстур в областях с высокой вариативностью, тогда как однородные зоны восстанавливаются лучше. В результате изображения, особенно в местах с высокой детализацией, теряют четкость и выглядят размытыми.

– SRGAN (*Super-Resolution Generative Adversarial Network*) является инновационной архитектурой, сочетающей super-resolution и генеративные состязательные сети. Генератор пытается создавать изображения с увеличенным разрешением, а дискриминатор обучается отличать реальные от сгенерированных [4]. Модель способна создавать изображения с более естественными текстурами и деталями, что делает результаты более приятными для человеческого глаза по сравнению с предыдущими методами [5]. К недостаткам можно отнести: 1) чрезмерный акцент на текстурах создает ложные элементы на изображении, которых могло не быть в оригинале; 2) нестабильность: если одна из сетей начинает доминировать, это может привести к дисбалансу в обучении. Если дискриминатор становится слишком эффективным, он подавляет генератор, что приводит к замедлению или даже полному прекращению его обучения. Если же доминирует генератор, дискриминатор теряет способность направлять обучение, и качество выходных изображений резко падает.

– TTSR (*Texture Transformer Super-Resolution*) представляет собой архитектуру, основанную на Transformer. TTSR предлагает альтернативный подход к задачам компьютерного зрения, отказываясь от сверточных нейронных сетей (*Convolutional Neural Network, CNN*) в пользу Transformer, которые обычно используются в задачах обработки естественного языка. Использование механизмов внимания позволяет глобально учитывать взаимосвязи между различными частями изображения, что полезно для восстановления текстур и деталей [6]. Однако, TTSR требует значительных вычислительных ресурсов. К тому же, данная нейросеть, по сравнению с CNN, где свёрточные слои изначально ориентированы на локальные особенности, хуже восстанавливает такие компоненты, как тонкие линии, текстуры, волосы.

Исходя из вышеперечисленных методов, их достоинств и недостатков, можно сделать вывод, что каждая архитектура решает конкретные проблемы, но приносит новые ограничения, что подчеркивает узкую специализацию подходов. Например, SRGAN, отлично справляется с визуальной реалистичностью, но при этом порождает артефакты из-за стремления к максимальной детализации. Для улучшения качества изображения необходимо одновременно улавливать мелкие локальные детали и сохранять целостность изображения, учитывая глобальный контекст. CNN сильны в первом, Transformer во втором, GAN в имитации реализма, но ни один подход не охватывает оба аспекта полноценно. Каждый из методов имеет свои сильные и слабые стороны, что подчеркивает необходимость разработки новых подходов сочетающих достоинства различных архитектур для достижения оптимального баланса между качеством и вычислительной эффективностью.

Архитектуры для реконструкции, основанные на Transformer, обычно делят входное изображение на фрагменты фиксированного размера (патчи) и обрабатывают каждый патч независимо. Такой метод имеет один недостаток: граничные пиксели не могут использовать соседние пиксели, находящиеся вне патча для восстановления изображения, в следствии чего восстановленное изображение может содержать артефакты по краям вокруг каждого фрагмента. Эту проблему можно решить путем перекрытия патчей, но это приведет к дополнительной вычислительной нагрузке. Архитектура SwinIR, основанная на Swin Transformer [7] на данный момент является перспективным решением, поскольку объединяет в себе преимущества CNN и Transformer.

1.2 Аналитический обзор архитектуры SwinIR

SwinIR состоит из трёх модулей [8]: 1) извлечения поверхностных признаков, 2) извлечения глубоких признаков, 3) реконструкции изображений (рисунок 1).



Рисунок 1 – Модули архитектуры SwinIR

На вход нейросети подается изображение в низком разрешении LQ. Для извлечения поверхностных признаков F_{SF} используется сверточный слой L_{SF} :

$$F_{SF} = L_{SF}(LQ)$$

Затем извлекаются глубокие признаки F_{DF} :

$$F_{DF} = L_{DF}(F_{SF}),$$

где L_{DF} – модуль извлечения глубоких признаков, который содержит N блоков RSTB и сверточный слой. Промежуточные признаки T_1, \dots, T_N и выходной глубокий признак F_{DF} извлекаются по блокам:

$$F_{DF} = L_{conv} \left(L_{RSTB_i}(T_{i-1}) \right), \quad i = 1, 2, \dots, N,$$

где L_{RSTB_i} означает i-тый блок RSTB, а L_{conv} – сверточный слой в конце.

Поверхностные и глубокие признаки объединяются и подаются в модуль реконструкции F_{RHQ} :

$$F_{RHQ} = L_{REC}(F_{SF} + F_{DF}),$$

где L_{REC} – функция модуля реконструкции. Для реализации модуля реконструкции используется субпиксельный сверточный слой для повышения дискретизации изображения.

Как показано на рисунке 2, блок RSTB состоит из нескольких слоев *Swin Transformer Layer* (STL) и сверткой в конце.

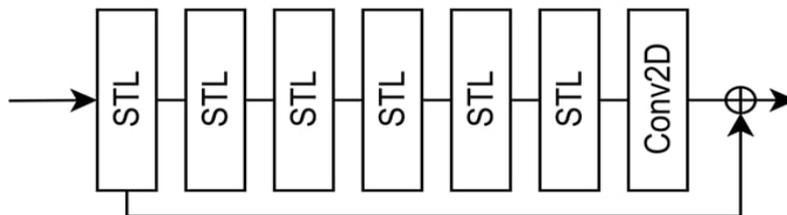


Рисунок 2 – Строение блока RSTB

Учитывая входной признак $T_{i,0}$ i-го RSTB блока, сначала извлекаются промежуточные признаки $T_{i,0}, T_{i,2}, \dots, T_{i,K}$ с помощью K слоёв STL. Выход RSTB выглядит следующим образом:

$$T_{i,out} = F_{conv_i} \left(F_{STL_{i,j}}(T_{i,j-1}) \right) + T_{i,0}, \quad j = 0, 1, \dots, K,$$

где $F_{STL_{i,j}}$ – это j-й слой STL в i-том блоке RSTB, F_{conv_i} – сверточный слой в i-том RSTB. STL основан на стандартном механизме внимания из оригинального Transformer [9]. Изображения сначала разбиваются на неперекрывающиеся локальные окна, после чего вычисляется внутреннее внимание отдельно для каждого окна. Далее границы окон сдвигаются ровно на половину размера окна. Сдвинутое окно включает в себя части других, это создает связь между соседними областями, позволяя модели учитывать глобальный контекст для восстановления изображения, избегая при этом артефактов на границах.

К числу ограничений архитектуры SwinIR относятся недостаточная способность к реалистичному моделированию искажений, возникающих в реальных условиях формирования низкокачественных изображений, а также ограниченная эффективность в учёте глобального контекста сцены. Это обусловлено особенностями синтеза пар изображений с низким и высоким разрешением для обучения (LR-HR), которые не всегда отражают реальные виды искажений, встречающихся на практике. Кроме того, механизмы обработки информации в SwinIR преимущественно ориентированы на локальные области, что может привести к потере глобальных взаимосвязей и снижению достоверности восстановления сложных структур на изображении. Ниже изложены перспективы развития этой архитектуры и процедуры обучения для задачи super-resolution.

2 Актуальные подходы к совершенствованию архитектуры и процесса обучения

2.1 Методы синтеза изображений для формирования обучающих выборок

Для обучения модели используется набор данных, состоящий из пар изображений высокого разрешения (HR) и соответствующих низкоразрешённых (LR) вариантов, которые обычно синтезируются путём последовательного понижения дискретизации и добавления шума к исходным изображениям. Такой подход формирует однородно искажённый тренировочный набор, что не всегда адекватно отражает разнообразие повреждений, встречающихся в реальных условиях съёмки. Вследствие этого эффективность модели снижается, если реальные искажения не соответствуют тем, на которых она обучалась [10], [11]. Для решения этой проблемы в перспективном представляется рандомизированный подход к аугментации, заключающийся в случайном перемешивании различных видов искажений в процессе синтеза LR изображений (перемешиваемые размытия, интерполяции и шумы). В данном контексте следует отметить работу [12], в которой впервые введена концепция «mixture of augmentations» для задачи super-resolution. Авторы показали, что комбинирование и случайный выбор нескольких аугментаций приводит к устойчивому улучшению качества восстановления. Данный подход закрепил идею случайного смешения аугментаций как эффективной стратегии повышения обобщающей способности моделей super-resolution.

Методы данного подхода для генерации пар изображений с низким и высоким разрешением основаны на трёхступенчатой рандомизированной аугментации. Подобный рандомизированный синтез позволяет моделировать широкий спектр реальных искажений, что способствует лучшей обобщающей способности обучаемой модели. Предусмотрено последовательное применение (в случайном порядке и с рандомизацией параметров) следующих этапов искусственного искажения исходного изображения:

1. Комбинация изотропного и анизотропного размытия. Изотропное размытие имитирует расфокусировку, а анизотропное размытие движение камеры или объекта [13]. Их комбинация даёт широкий спектр реальных искажений.
2. Случайный выбор типа интерполяционного алгоритма (метода понижения дискретизации):
 - Метод К-ближайших соседей.
 - Билинейная интерполяция.
 - Бикубическая интерполяция.

Сначала изображение сильно уменьшается, а затем увеличивается до стандартного размера LR. Этот метод позволяет смоделировать артефакты, возникающие при многократном пережатии изображения [14].

3. Добавление шума со случайной вероятностью с вариацией его свойств и параметров:

- Гауссов шум. Уровень шума выбирается случайным образом.
- Шум JPEG-сжатия. Изображение сжимается с помощью JPEG со случайным коэффициентом качества.
- Шум Пуассона. Этот тип шума актуален для изображений, снятых в условиях низкой освещенности [15]. Уровень шума варьируется случайным образом, чтобы имитировать реальные фотографические сценарии.

Иллюстрация синтеза обучающих данных с помощью трёхступенчатой рандомизированной аугментации представлена на рисунке 3. Приведённые результаты демонстрируют внешний вид синтетических изображений LR, формируемых из исходных HR посредством случайных комбинаций размытия, интерполяции и добавления шума. Подобная визуализация позволяет наглядно оценить реализуемое разнообразие искусственных искажений, приближённых к реальным условиям формирования низкокачественных данных.

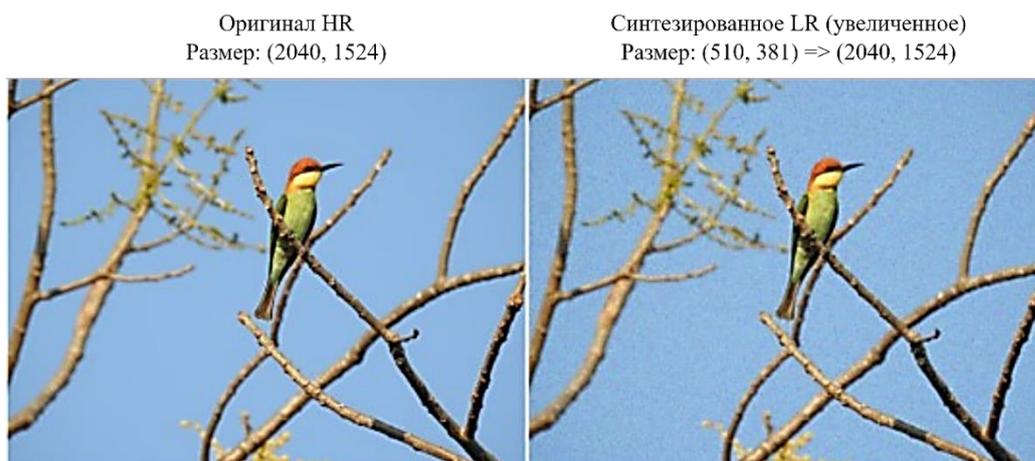


Рисунок 3 – Пример синтетического LR-изображения, полученного в результате трёхэтапной аугментации (справа) из исходного HR-изображения (слева)

Дальнейшим развитием подхода рандомизированного синтеза искажений является внедрение формализованной трёхэтапной схемы с вероятностной композицией и параметризацией операций. Такая организация процесса синтеза позволяет моделировать широкий спектр возможных искажений, характерных для реальных условий формирования низкоразрешённых изображений, и способствует повышению обобщающих свойств обучаемой модели по сравнению с традиционными методами генерации обучающих данных.

2.2 Интеграция блока быстрого преобразования Фурье в архитектуру SwinIR

В задачах восстановления изображений и повышения их качества с использованием архитектур на основе Transformer основное внимание уделяется преодолению разрыва между извлечением локальных признаков и интеграцией глобального

контекста. В частности, в модели SwinIR каждый *Swin Transformer Block* (RSTB) завершается стандартной сверточной операцией, что обеспечивает выделение локальных признаков благодаря ограниченному рецептивному полю. Однако в задаче super-resolution глобальный контекст, включая распределение цветов, повторяющиеся текстурные структуры и прочие интегративные характеристики, имеет существенное значение для получения изображений высокого качества без артефактов. Следует отметить, что стандартные сверточные операции, обладая ограниченным рецептивным полем, недостаточно эффективны для захвата глобального контекста, а попытки увеличить эффективное поле за счёт глубины сети ведут к росту вычислительной сложности [16].

Внедрение блока *Fast Fourier Convolution* (FFC) позволяет организовать обработку входных признаков в двух параллельных ветвях: локальной и глобальной. Локальная ветвь использует стандартные сверточные операции для анализа деталей в малых областях изображения. Глобальная ветвь применяет быстрое преобразование Фурье, переводя признаки в частотную область, где изображение описывается спектром коэффициентов, отражающих вклад различных пространственных частот. Низкочастотные компоненты характеризуют такие свойства, как общий цветовой фон и крупные объекты, а высокочастотные – детализацию, границы и текстурные элементы. Точечная свертка обновляет коэффициенты по отдельности, что влияет на всё изображение как единое целое. После обратного преобразования Фурье восстанавливается пространственное представление признаков.

Объединение обработанных признаков из локальной и глобальной ветвей формирует обогащённое представление, включающее как локальные детали, так и глобальный контекст [17]. Это объединение происходит перед передачей признаков на вход следующего блока RSTB или модуля реконструкции изображения. Измененный выход RSTB выглядит следующим образом:

$$T'_{i,out} = F_{FFC_i} \left(F_{STL_{i,j}}(T_{i,j-1}) \right) + T_{i,0}, \quad j = 0, 1, \dots, K,$$

где F_{FFC_i} – это быстрая свертка Фурье в i -том блоке RSTB, которая состоит из двух ветвей:

$$F_{FFC_i} = L_{conv} \left(F_{STL_{i,j}}(T_{i,j-1}) \right) + L_{global} \left(F_{STL_{i,j}}(T_{i,j-1}) \right), \quad j = 0, 1, \dots, K,$$

где L_{global} – глобальная ветвь, L_{conv} – локальная.

Исследования в смежных областях, таких как классификация [18], сегментация [19] и детекция объектов [20], [21], показывают, что использование FFC приводит к улучшению значений количественных метрик (PSNR, SSIM) на 5–10% по сравнению с моделями, в которых используются исключительно стандартные сверточные операции. На основании полученных результатов можно ожидать сопоставимого прироста качества и в задаче super-resolution.

Заключение

В настоящей работе проведён сравнительный анализ современных архитектур, применяемых для повышения качества изображений в задаче super-resolution, включая SRCNN, SRResNet, SRGAN, TTSR и SwinIR. На основании проведённого анализа особенностей архитектур можно сделать вывод, что наиболее перспективным направлением дальнейшего развития является использование архитектуры SwinIR, обладающей высокой эффективностью и гибкостью.

Ключевые возможности для повышения эффективности такой архитектуры включают использование рандомизированной аугментации при формировании обучающих данных, что обеспечивает более реалистичное моделирование искажений, а также интеграцию блока FFC, позволяющую заменить традиционные сверточные слои и эффективно учитывать глобальный контекст изображения. Совокупное применение указанных подходов способствует улучшению качества реконструкции изображений при оптимальных вычислительных затратах, определяя потенциал дальнейшего совершенствования методов глубокого обучения в сфере повышения разрешения визуальных данных. Такие разработки не только повысят точность анализа изображений, но и расширят область их применимости в критически важных сферах, требующих анализа изображений высокого качества.

Список литературы

1. Сай С.В. Способ оценки четкости фотореалистичных изображений с высоким разрешением // Компьютерная оптика. 2022. Т. 46, № 1. С. 121-129. ISSN 2412-6179. DOI: 10.18287/2412-6179-CO-899.
2. Коновалов, В.Ф. Единый нейросетевой метод повышения разрешения для разнородных цифровых изображений дистанционного зондирования Земли / В.Ф. Коновалов, В.В. Мясников, В.В. Сергеев // Компьютерная оптика. 2024. Т. 48, № 6. С. 944-955. DOI: 10.18287/2412-6179-CO-I610.
3. Singh P., Ganotra D. Brain MRT image super resolution using discrete cosine transform and convolutional neural network. Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2023, vol. 23, no. 4, pp. 734-742. doi: 10.17586/2226-1494-2023-23-4-734-742.
4. Сай С.В., Кудяшов А.А. Оценка качества изображений природного ландшафта с повышенным разрешением на основе GAN // Вестник ТОГУ. 2024. № 1 (72). С. 45-52. УДК 004.896.
5. Шелковникова, Т.Е. Применение генеративно-сопоставительных нейросетей для формирования баз данных в сканирующей туннельной микроскопии / Т.Е. Шелковникова, С.Ф. Егоров, П.В. Гуляев // Компьютерная оптика. 2023. Т. 47, № 2. С. 314-322. DOI: 10.18287/2412-6179-CO-1144.
6. Yang F., Yang H., Fu J., Lu H., Guo B. Learning Texture Transformer Network for Image Super-Resolution. 2020. P. 5790-5799. DOI: 10.1109/CVPR42600.2020.00583.
7. Liu, Z. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows [Text] / Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo // arxiv preprint. 2021. arxiv:2103.14030.
8. Liang, J. SwinIR: Image Restoration Using Swin Transformer / J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, R. Timofte // arxiv preprint. 2021. arxiv:2108.10257.
9. Бережнов, Н.И. Совершенствование механизмов внимания для архитектуры трансформер в задачах повышения качества изображений / Н.И. Бережнов, А.А. Сирота // Компьютерная оптика. 2024. Т. 48, № 5. С. 726-733. DOI: 10.18287/2412-6179-CO-1393.
10. Ермоленко Т. В., Самородский И. Е. Анализ эффективности архитектур глубоких нейросетей для классификации // Журнал «Проблемы искусственного интеллекта». 2022. № 1(24). С. 54-64. ISSN 2413-7383.
11. Денисов А.К., Быковский С.В., Кустарев П.В. Метод увеличения разрешения изображения с использованием референсных изображений на основе диффузионной модели // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 2. С. 321-327. doi: 10.17586/2226-1494-2025-25-2-321-327.
12. Yoo, J.S., Kim, M., & Kim, S.J. (2020). Rethinking Data Augmentation for Image Super-Resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA, pp. 10735-10744. DOI: 10.1109/CVPR42600.2020.01075.
13. Чаганова О.Б., Григорьев А.С., Николаев Д.П., Николаев И.П. Прикладные аспекты современных методов неослепительного восстановления изображений [Текст] // Компьютерная оптика. – 2024. Т. 48, № 4. С. 562–572. ISSN 2412-6179. DOI: 10.18287/2412-6179-CO-1409.
14. Zhang X., Wu X. Dual-layer Image Compression via Adaptive Downsampling and Spatially Varying Upsampling. 2023. P. 1-10. DOI: arXiv:2302.06096.
15. Котенко, И. В. Метод противодействия состязательным атакам на системы классификации изображений [Текст] / И. В. Котенко, И. Б. Саенко, О. С. Лаута, Н. А. Васильев, В. Е. Садовников // Вопросы кибербезопасности. 2025. № 2(66). С. 114-123. ISSN 2311-3456. DOI: 10.21681/2311-3456-2025-2-114-123.

16. Пикалёв Я.С., Ермоленко Т.В. О нейронных архитектурах извлечения признаков для задачи распознавания объектов на устройствах с ограниченной вычислительной мощностью // Журнал «Проблемы искусственного интеллекта». 2023. № 3. С. 44-54. ISSN 2413-7383. DOI 10.34757/2413-7383.2023.30.3.004.
17. Чернядьев, И. В. Системный анализ программного обеспечения концептуальной нейросетевой модели системы классификации эмоций на изображениях / И. В. Чернядьев В. В. Бондарчук // Евразийский Союз Ученых. Серия: технические и физико-математические науки. 2024. № 08(123). С. 6-15. ISSN 2413-9335. DOI: 10.31618/ESU.2413-9335.
18. Rao, Y. Global Filter Networks for Image Classification / Y. Rao, W. Zhao, Z. Zhu, J. Lu, J. Zhou // arxiv preprint. 2021. arxiv:2107.00645.
19. Quattrini, F. Volumetric Fast Fourier Convolution for Detecting Ink on the Carbonized Herculaneum Papyri [Text] / F. Quattrini, V. Pippi, S. Cascianelli, R. Cucchiara // arxiv preprint. 2023. arxiv:2308.05070.
20. Gu, L. Fast Fourier Convolution Based Remote Sensor Image Object Detection for Earth Observation [Text] / L. Gu, R. Qiao, X. Chen // arxiv preprint. 2022. arxiv:2209.00551.
21. Miao W., Shen J., Xu Q., Hamalainen T., Xu Y., Cong F. SpikingYOLOX: Improved YOLOX Object Detection with Fast Fourier Convolution and Spiking Neural Networks / W. Miao, J. Shen, Q. Xu, T. Hamalainen, Y. Xu, F. Cong. 2025. Vol. 39, iss. 2. P. 1465-1473. DOI: 10.1609/aaai.v39i2.32137.

References

1. Sai S.V. A method for assessing photorealistic image quality with high resolution [Text] // Computer Optics. – 2022. – Vol. 46, No. 1. – P. 121-129. – ISSN 2412-6179. – DOI: 10.18287/2412-6179-CO-899.
2. Konovalov, V.F. Unified Neural Network Method for Enhancing Resolution for Heterogeneous Digital Images of Earth Remote Sensing / V.F. Konovalov, V.V. Myasnikov, and V.V. Sergeev // Computer Optics. - 2024. - Vol. 48, No. 6. - Pp. 944-955. - DOI: 10.18287/2412-6179-CO-I610.
3. Singh P., Ganotra D. Brain MRT image super resolution using discrete cosine transform and convolutional neural network. Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2023, vol. 23, no. 4, pp. 734-742. doi: 10.17586/2226-1494-2023-23-4-734-742.
4. Sai S.V., Kudyashov A.A. Assessment of the quality of natural landscape images with increased resolution based on GAN [Text] // Bulletin of TOGU. – 2024. – No. 1 (72). – P. 45-52. – UDC 004.896.
5. Shelkovnikova TE, Egorov SF, Gulyaev PV. Application of generative adversarial neural networks for the formation of databases in scanning tunneling microscopy. Computer Optics 2023; 47(2): 314-322. DOI: 10.18287/2412-6179-CO-1144.
6. Yang F., Yang H., Fu J., Lu H., Guo B. Learning Texture Transformer Network for Image Super-Resolution / F. Yang, H. Yang, J. Fu, H. Lu, B. Guo. – 2020. – P. 5790-5799. – DOI: 10.1109/CVPR42600.2020.00583.
7. Liu, Z. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows [Text] / Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo // arxiv preprint. – 2021. – arxiv:2103.14030.
8. Liang, J. SwinIR: Image Restoration Using Swin Transformer [Text] / J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, R. Timofte // arxiv preprint. – 2021. – arxiv:2108.10257.
9. Berezhnov NI, Sirota AA. Improving attention mechanisms in transformer architecture in image restoration. Computer Optics 2024; 48(5): 726-733. DOI: 10.18287/2412-6179-CO-1393.
10. Yermolenko T.V., Samorodsky I.Ye. Analysis of the efficiency of deep neural networks architectures for the classification of products images [Text]//Yermolenko T.V., Samorodsky I.Ye. Journal "Problems of Artificial Intelligence". – 2022. – № 1(24). – С. 54-64. – ISSN 2413-7383.
11. Denisov A.K., Bykovskii S.V., Kustarev P.V. Reference-based diffusion model for super-resolution. Scientific and Technical Journal of Information Technologies, Mechanics and Optics, 2025, vol. 25, no. 2, pp. 321-327 (in Russian). doi: 10.17586/2226-1494-2025-25-2-321-327.
12. Yoo, J.S., Kim, M., & Kim, S.J. (2020). Rethinking Data Augmentation for Image Super-Resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA, pp. 10735-10744. DOI: 10.1109/CVPR42600.2020.01075.
13. Chaganova O.B., Grigoryev A.S., Nikolaev D.P., Nikolaev I.P. Applied aspects of modern non-blind image deconvolution methods [Text] // Computer Optics. – 2024. – Vol. 48, no. 4. – P. 562–572. – ISSN 2412-6179. – DOI: 10.18287/2412-6179-CO-1409.
14. Zhang X., Wu X. Dual-layer Image Compression via Adaptive Downsampling and Spatially Varying Upconversion / X. Zhang, X. Wu. – 2023. – P. 1-10. – DOI: arXiv:2302.06096.

15. Kotenko, I. V. A Method for Countering Competitive Attacks on Image Classification Systems [Text] / I. V. Kotenko, I. B. Saenko, O. S. Lauta, N. A. Vasilyev, and V. E. Sadovnikov // *Cybersecurity Issues*. – 2025. – No. 2(66). – Pp. 114-123. – ISSN 2311-3456. – DOI: 10.21681/2311-3456-2025-2-114-123.
16. Pikalyov Ya.S., Yermolenko T.V. About Neural Architectures of Feature Extraction for the Problem of Object Recognition on Devices with Limited Computing Power [Text] // *Journal "Problems of Artificial Intelligence"*. – 2023. – № 3. – С. 44-54. – ISSN 2413-7383. – DOI 10.34757/2413-7383.2023.30.3.004.
17. Chernyadev, I. V. System Analysis of the Software of the Conceptual Neural Network Model of the Emotion Classification System in Images [Text] / I. V. Chernyadyev, V. V. Bondarchuk // *Eurasian Union of Scientists. Series: Technical and Physical and Mathematical Sciences*. – 2024. – No. 08(123). – Pp. 6-15. – ISSN 2413-9335. – DOI: 10.31618/ESU.2413-9335.
18. Rao, Y. Global Filter Networks for Image Classification [Text] / Y. Rao, W. Zhao, Z. Zhu, J. Lu, J. Zhou // *arxiv preprint*. – 2021. – arxiv:2107.00645.
19. Quattrini, F. Volumetric Fast Fourier Convolution for Detecting Ink on the Carbonized Herculaneum Papyri [Text] / F. Quattrini, V. Pippi, S. Cascianelli, R. Cucchiara // *arxiv preprint*. – 2023. – arxiv:2308.05070.
20. Gu, L. Fast Fourier Convolution Based Remote Sensor Image Object Detection for Earth Observation [Text] / L. Gu, R. Qiao, X. Chen // *arxiv preprint*. – 2022. – arxiv:2209.00551.
21. Miao W., Shen J., Xu Q., Hamalainen T., Xu Y., Cong F. SpikingYOLOX: Improved YOLOX Object Detection with Fast Fourier Convolution and Spiking Neural Networks / W. Miao, J. Shen, Q. Xu, T. Hamalainen, Y. Xu, F. Cong. – 2025. – Vol. 39, iss. 2. – P. 1465-1473. – DOI: 10.1609/aaai.v39i2.32137.

RESUME

I. V. Chernyadev, T. V. Yermolenko
Image improvement using deep learning:
analysis and prospective developments

The article is devoted to the analysis and development of image quality improvement methods based on deep learning aimed at eliminating the disadvantages of existing solutions, such as artifacts, high resource intensity.

Improvements are proposed in the work: 1) randomized augmentation to create LR-HR pairs: a combination of isotropic and anisotropic blurring, random selection of interpolation; adding noise with varying levels and sequences. This simulates real-world distortions, increasing the generalization of the model; 2) replacing convolution with FFC, which divides features into local and global branches, which enhances the global context, improving the quality of image recovery. An increase in metrics is expected, as in related tasks.

As a result, the following conclusions can be drawn: 1) during neural network training, special attention should be paid to LR image synthesis methods, because they significantly affect the quality of the model; 2) Transformer-based architectures are a promising basis for further work on image quality improvement; 3) the proposed modifications will increase the efficiency and applicability of the neural network.

РЕЗЮМЕ

И. В. Чернядьев, Т. В. Ермоленко
Улучшение изображений с применением
глубокого обучения: анализ и перспективные разработки

Статья посвящена анализу и разработке методов улучшения качества изображений на основе глубокого обучения, направленных на устранение недостатков существующих решений, таких как артефакты и высокая ресурсоемкость.

В работе предлагаются улучшения: 1) рандомизированная аугментация для создания пар LR-HR: комбинация изотропного и анизотропного размытия, случайный выбор понижения дискредитации, добавление шумов с вариацией уровней и

последовательности. Это моделирует реальные искажения, повышая обобщение модели. 2) замена свертки на FFC, разделяющий признаки на локальную и глобальную ветви, что усиливает глобальный контекст, повышая качество восстановления изображений. Ожидается прирост в метриках, как в смежных задачах.

Выводы: 1) во время обучения нейросети необходимо уделять особое внимание на методы синтеза LR изображений, потому что они значительно влияют на качество работы модели; 2) архитектуры, основанные на Transformer, являются перспективной основой для дальнейшей работы над повышением качества изображений; 3) предложенные модификации повысят эффективность и применимость нейросети.

Чернядьев Иван Валерьевич – инженер-исследователь ФГБНУ "Институт проблем искусственного интеллекта", бакалавр физико-технического факультета ФГБОУ ВО "ДонГУ". *Область научных интересов:* компьютерное зрение, машинное обучение, нейронные сети.

Ермоленко Татьяна Владимировна – к.т.н., научный сотрудник отдела системного анализа и интеллектуальных интерфейсов ФГБНУ "Институт проблем искусственного интеллекта", кандидат технических наук, доцент. *Область научных интересов:* машинное обучение, обработка естественного языка, распознавание образов, интеллектуальный анализ данных.

Статья поступила в редакцию 03.10.2025.

Авторы номера

Абдулвадулов М. Л. инженер-программист	Научно-производственный центр «Электронные вычислительно-информационные системы», Россия, г. Зеленоград, abdulvadudovmagomed@mail.ru
Близно М. В.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк
Бондаренко В. И., к.т.н., доцент	Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий государственный университет», Донецкая Народная Республика, г. Донецк
Бондарчук В. В., к.т.н.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк
Васильев П. В.	Институт механики им. Р. Р. Мавлютова УФИЦ РАН, г. Уфа
Вахидова К. Л., доцент	Грозненский государственный нефтяной технический университет имени академика М.Д. Миллионщикова, Чеченская Республика, г. Грозный
Водолазская Н. В., к.т.н., доцент	МИРЭА — Российский технологический университет, институт искусственного интеллекта, Россия, г. Москва, vny26@bk.ru .
Володина А. М.	Институт искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, volodina@mirea.ru
Воронов А. Э.,	Федеральное государственное бюджетное образовательное учреждение высшего образования «Луганский государственный университет имени Владимира Даля», Луганская Народная Республика, г. Луганск, osooler@ya.ru
Вотинцев М. К.	МИРЭА — Российский технологический университет, институт искусственного интеллекта, Россия, г. Москва, votintsev.mikhail.k@mail.ru
Григорьев А.В., к.т.н., доцент	Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий национальный технический университет», Донецкая Народная Республика, г. Донецк, grigorievalvl@gmail.com
Гурин А.Г.	Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий национальный технический университет», Донецкая Народная Республика, г. Донецк, gurin.sasha.1996@gmail.com
Данилов В. В., д.т.н., проф.	Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий государственный университет», Донецкая Народная Республика, г. Донецк, ut5iv@mail.ru
Даринцев О. В., д.т.н., доцент	ФГБОУ ВО «Уфимский университет науки и технологий», г. Уфа
Дзебоев Д. И.	Национальный исследовательский университет «Высшая школа экономики», Москва, dzeboev.daniil@gmail.com
Доу Л.	Московский государственный технический университет имени Н.Э. Баумана, г. Москва, Россия

Авторы номера

Елисеев В. О.	Федеральное государственное бюджетное научное учреждение «Институт прикладной механики и математики», Донецкая Народная Республика, г. Донецк
Звягинцева А. В., д.т.н., доцент	ФГБОУ ВО «Донецкий государственный университет», Донецкая Народная Республика, г. Донецк
Зори С. А., д.т.н., доцент	Государственное образовательное учреждение высшего профессионального образования «Донецкий национальный технический университет», Донецкая Народная Республика, г. Донецк
Зуев В. М.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк, zvm.ipai@mail.ru
Ивашко К. С.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк
Изосимова С. А.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк
Кириллов С. К.	Институт искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, kirillov.s342@gmail.com
Копытова Е. В.	МИРЭА — Российский технологический университет, институт искусственного интеллекта, Россия, г. Москва, kopytova@mirea.ru
Ковалева К.В.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк
Кравченко Н. М., к.т.н.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк
Курган Е. Г., к.э.н., доцент	Федеральное государственное бюджетное научное учреждение «Институт научно-технической информации», Донецкая Народная Республика, г. Донецк
Курнасов Е. В., к.т.н., доцент	Институт искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, kurnasov@mirea.ru
Логунова Д. М.	Институт искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, dm.logunova@gmail.com
Лямцев О. А.	Федеральное государственное бюджетное научное учреждение «Институт прикладной механики и математики»; ФГБОУ ВО «ДонГУ», Донецкая Народная Республика, г. Донецк, gelo2003@yandex.ru
Максименко И. И., к.ф.-м.н.	Федеральное государственное бюджетное научное учреждение «Институт прикладной механики и математики»; ФГБОУ ВО «ДонГУ», Донецкая Народная Республика, г. Донецк, igor.maksimenko_1967@mail.ru.
Максимова А. Ю., к.т.н., доцент	Федеральное государственное бюджетное научное учреждение «Институт прикладной механики и математики», Донецкая Народная Республика, г. Донецк
Мороз Е. С.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк, azidaan.moroz@yandex.com

Никитенко К. А.	ФГБОУ ВО «Донецкий государственный университет», Донецкая Народная Республика, г. Донецк
Мунасыпов Р. А., д.т.н.	ФГБОУ ВО «Уфимский университет науки и технологий», г. Уфа
Ногамирзаев С. М.	Грозненский государственный нефтяной технический университет имени академика М.Д. Миллионщикова, kafedra2020@mail.ru
Остапущенко Д. Л., к.т.н.	Федеральное государственное бюджетное образовательное учреждение высшего образования «Луганский государственный университет имени Владимира Даля», Луганская Народная Республика, г. Луганск, ostapuschenko_dmitriy@mail.ru
Павленко Б. В.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк, bogdanpavl2000@mail.ru
Павлюченко К. А.	Белорусский государственный университет информатики и радиоэлектроники, Минск
Парамонов А. И., к.т.н., доцент	Белорусский государственный университет информатики и радиоэлектроники, Минск
Петренко М. О.	Федеральное государственное бюджетное образовательное учреждение высшего образования «Луганский государственный университет имени Владимира Даля», Луганская Народная Республика, г. Луганск, max_petrenko_2002@mail.ru
Петрова К. С., инженер-программист	Федеральное государственное бюджетное научное учреждение «Институт научно-технической информации», Донецкая Народная Республика, г. Донецк
Пигуз В. Н.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк
Пикалёв Я. С., к.т.н.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк, i@pikaliov.ru
Рущечников Я. И.	Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий государственный университет», Донецкая Народная Республика, г. Донецк, ya.rushechnikov@donnu.ru
Собчишен А. С., инженер-программист	Федеральное государственное бюджетное научное учреждение «Институт научно-технической информации», Донецкая Народная Республика, г. Донецк
Солод В. С., к.т.н.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк, vs.solod@mail.ru
Тонконоженко Н. Д.	Институт искусственного интеллекта РТУ МИРЭА, г. Москва, Россия, tonkononikita@gmail.com
Третьяков И. А. к.т.н., доцент	Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий государственный университет», Донецкая Народная Республика, г. Донецк, i.tretiakov@mail.ru

Авторы номера

Устенко В. Ю.	Федеральное государственное бюджетное научное учреждение «Институт проблем искусственного интеллекта», Донецкая Народная Республика, г. Донецк, ustenko.vova@ya.ru
Хиценко А. И., к.т.н., доцент	Государственный университет просвещения, Россия, г. Москва, ganna_x@mail.ru
Черных В. В., к.т.н.	Федеральное государственное бюджетное образовательное учреждение высшего образования «Луганский государственный университет имени Владимира Даля», Луганская Народная Республика, г. Луганск, viktoria.chernykh@mail.ru
Филипишин Д.А.	Государственное образовательное учреждение высшего профессионального образования «Донецкий национальный технический университет», Донецкая Народная Республика, г. Донецк

Рецензенты номера

Анцыферов С. С., д.т.н., профессор
Андриевская Н.К., к.т.н., доцент
Балдин А.В. д.т.н., профессор
Беловодский В. Н., к.т.н., доцент
Болбаков Р. Г., к.т.н., доцент
Варламов О. О., д.т.н., профессор
Гудаев О.О.
Зори С.А., д.т.н., профессор
Зуев В.М., н.с.
Кузнецов Д. Н., к.т.н., доцент
Максимова А.Ю., к.т.н.
Максименко И.И., к.ф.-м.н., доцент
Ниценко А. В., к.т.н.
Павлыш В. Н., д.т.н., профессор
Рычка О.В., к.т.н.
Сальников И. С., к.т.н.
Солод В. С., к.т.н.
Толстых В.К., д.ф.-м.н., профессор
Шелепов В. Ю., д.ф.-м.н., профессор

Авторский указатель

А

Абдулвадулов М. Л., 163

Б

Бливно М. В., 84

Бондаренко В. И., 36

Бондарчук В. В., 206

В

Васильев П.В., 183

Вахидова К. Л., 163

Водолазская Н. В., 125

Володина А. М., 137

Воронов А. Э., 99

Вотинцев М. К., 125

Г

Григорьев А.В., 15

Гурин А.Г., 15

Д

Данилов В. В., 150

Даринцев О. В., 183

Дзэбоев Д. И., 27

Доу Л., 194

Е

Елисеев В. О., 36

Ермоленко Т. В., 254

З

Звягинцева Г. В., 229

Зори С. А., 241

Зуев В. М., 173

И

Ивашко К. С., 206

Изосимова С. А., 206

К

Кириллов С. К., 137

Ковалева К. В., 206

Копытова Е. В., 125

Кравченко Н. М., 206

Курган Е. Г., 219

Курнасов Е. В., 137

Л

Логунова Д. М., 137

Лямцев О. А., 49

М

Максименко И. И., 49

Максимова А. Ю., 36

Мороз Е. С., 60

Мунасыпов Р. А., 183

Н

Никитенко К. А., 229

Ногамирзаев С. М., 163

О

Остапущенко Д. Л., 99

П

Павленко Б. В., 71

Павлюченко К. А., 4

Парамонов А. И., 4

Петренко М. О., 99

Петрова К. С., 219

Пигуз В. Н., 206

Пикалёв Я. С., 60

Р

Рушечников Я. И., 150

С

Солод В. С., 112

Собчишен А. С., 219

Т

Тонконоженко Н. Д., 137

Третьяков И. А., 150

У

Устенко В. Ю., 84

Ф

Филипишин Д. А., 241

Х

Хиценко А. И., 125

Ч

Черных В. В., 99

Чернядьев И. В., 254

Т

СОДЕРЖАНИЕ	CONTENTS
Раздел 1 ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ	Section 1 ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING
Парамонов А. И., Павлюченко К. А. Категоризация текстовых документов на основе многоаспектного анализа эмоционального контекста (англ.) 4	Paramonov A. I., Pavluchenko K. A. Multi-Aspect Analysis Of Emotional Context For Text Document Categorization..... 4
Гурин А.Г., Григорьев А.В. Сравнительный анализ методов повышения эффективности производственных систем для работы с большими данными. комбинация двух подходов 15	Gurin A.G., Grigoriev A.V. Comparative Analysis of Methods for Enhancing the Efficiency of Production Systems for Big Data Processing..... 15
Дзедобоев Д. И. Поиск единиц в тензориальной алгебре 27	Dzeboev D. I. Search Of Units In Tensorial Algebra..... 27
Елисеев В. О., Бондаренко В. И., Максимова А. Ю. Разработка интеллектуального ассистента для образовательной организации 36	Eliseev V. O., Bondarenko V. I., Maksimova A. Y. Development of an intelligent assistant for an educational institution 36
Лямцев О. А., Максименко И. И. Обзор и проблемы использования глубокого обучения для трехмерной оценки позы человека по одному изображению 49	Lyamtsev O. A., Maksimenko I. I. Overview and challenges of using deep learning for three-dimensional estimation of a person's posture from a single image.... 49
Мороз Е. С., Пикалёв Я. С. Анализ Sota-архитектур для семантической аугментации 60	Moroz Ye. S., Pikaliyov Y.S. Analysis of Sota-Architectures For semantic augmentation 60
Павленко Б. В. Подход к мультимодальному объединению данных в задачах регрессии и классификации на снимках с БПЛА 71	Pavlenko B. V. An approach to multimodal data fusion In tasks of regression and classification In drone images 71
Устенко В.Ю., Близно М.В. Методика расширения данных для сегментации ключевых объектов на аэрофотоснимках БПЛА 84	Ustenko V.Y. Blizno M.V. Data Augmentation Methodology For Segmenting Key Objects In UAV Aerial Imagery 84

Раздел 2 МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ, ЧИСЛЕННЫЕ
МЕТОДЫ И КОМПЛЕКСЫ ПРОГРАММ

**Остапущенко Д. Л., Воронов А. Э.,
Черных В. В., Петренко М. О.**

Численное решение задачи аэродинамики
при моделировании полета беспилотных
летательных аппаратов с интеллектуальными
системами управления..... 99

Солод В. С.

Алгоритм и программа оптимизации режима
ускоренного охлаждения арматурных профилей
в термоустановке конструкции «Thermex» 112

Раздел 3 АВТОМАТИЗАЦИЯ
И УПРАВЛЕНИЕ ТЕХНОЛОГИЧЕСКИМИ
ПРОЦЕССАМИ И ПРОИЗВОДСТВАМИ

**Водолазская Н. В., Вотинцев М.К.,
Копытова Е. В., Хиценко А.И.**

Информационно-управляющая система
контроля технологического процесса
изготовления комплексных минеральных
удобрений..... 125

**Логунова Д.М., Володина А.М., Курнасов Е.В.,
Кириллов С.К., Тонконоженко Н.Д.**

Применение свёрточных нейронных сетей
на производстве гитар для выявления
дефектов продукции и автоматизированного
подбора пар заготовок 137

Раздел 4 МЕТОДЫ И СИСТЕМЫ
ЗАЩИТЫ ИНФОРМАЦИИ,
ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

**Рущечников Я. И., Третьяков И. А.,
Данилов В. В.**

Адаптивная автоматизированная система
подавления побочных электромагнитных
излучений и наводок 150

Раздел 5 РОБОТЫ, МЕХАТРОНИКА.
РОБОТОТЕХНИЧЕСКИЕ СИСТЕМЫ

**Абдулвадулов М. Л., Вахидова К.Л.,
Ногамирзаев С.М.**

Применение методов повышения эффективности
нейросетевого анализа гетерогенных систем
на кристалле в задачах робототехники 163

Зуев В. М.

Погрешность прогнозирования положения
следующего кадра съёмки 173

Section 2 MATHEMATICAL MODELING,
NUMERICAL METHODS
AND SOFTWARE PACKAGES

**Ostapushchenko D. L., Voronov A. E.,
Chernykh V. V., Petrenko M. O.**

Numerical Solution Of The Aerodynamics
Problem In Modeling The Flight
Of Unmanned Aerial Vehicles
With Intelligent Control Systems 99

Solod V. S.

Algorithm and program for optimizing
The accelerated cooling operation mode
Of reinforcement bars in the thermal unit
Of the "thermex" design 112

Section 3 AUTOMATION AND CONTROL
OF TECHNOLOGICAL PROCESSES
AND PRODUCTION

**Vodolazskaya N. V., Votintsev M. K.,
Kopytova E. V., Khitsenko A. I.**

Information And Control System
For Monitoring The Technological Process
Of Manufacturing Complex Mineral
Fertilizers 125

**Logunova D.M., Volodina A.M., Kurnasov E.V.,
Kirillov S.K., Tonkonozhenko N.D.**

The use of convolutional neural networks
in the manufacture of guitars to identify
product defects and automated selection
Of pairs of blanks 137

Section 4 METHODS AND SYSTEMS
INFORMATION PROTECTION,
INFORMATION SECURITY

**Rushechnikov Ya. I., Tretiakov I. A.,
Danilov V. V.**

Adaptive Automated System
For Suppression Of Electromagnetic
Side Effects And Interference 150

Section 5 ROBOTS, MECHATRONICS.
ROBOTIC SYSTEMS

**Abdulvadudov M. L., Vakhidova K. L.,
Nogamirzaev S. M.**

Application of methods for increasing
The efficiency of neural network analysis
Of heterogeneous systems-on-chip
In robotics tasks 163

Zuev V. M.

Error in predicting the position
Of the next film frame 173

Раздел 6 СИСТЕМНЫЙ АНАЛИЗ,
УПРАВЛЕНИЕ И ОБРАБОТКА
ИНФОРМАЦИИ, СТАТИСТИКА

**Васильев П.В., Даринцев О.В.,
Мунасыпов Р.А.**

Применение ассоциативного регулятора в
контуре управления мехатронным модулем 183

Доу Л.

Метод обработки информации и изменения
пороговых параметров правил в миварной
экспертной системе с использованием
больших языковых моделей..... 194

**Ковалева К. В., Кравченко Н. М., Бондарчук В.В.,
Пигуз В. Н., Изосимова С. А., Ивашко К. С.**

Результаты опытной эксплуатации
компьютерной интеллектуальной системы
диагностики и коррекции стрессовых
состояний 206

Курган Е. Г., Петрова К. С., Собчишен А. С.

Разработка электронной базы данных
как инструмент цифровизации центральной
научно-технической библиотеки 219

Никитенко К. А., Звягинцева А. В.

Оценка семантической схожести предложений
с использованием вероятностных мер
эмбедингов 229

Филипишин Д.А., Зори С.А.

Метод резолюции как средство расширения
машины вывода онтологий 241

Чернядьев И. В., Ермоленко Т. В.

Улучшение изображений
с применением глубокого обучения:
анализ и перспективные разработки..... 254

АВТОРЫ НОМЕРА 266

РЕЦЕНЗЕНТЫ НОМЕРА..... 270

АВТОРСКИЙ УКАЗАТЕЛЬ..... 271

ПОРЯДОК НАПРАВЛЕНИЯ
В РЕДАКЦИЮ, РЕЦЕНЗИРОВАНИЯ
И ОПУБЛИКОВАНИЯ СТАТЕЙ 275

Section 6 SYSTEM ANALYSIS,
MANAGEMENT AND INFORMATION
PROCESSING, STATISTICS

**Vasilev P.V., Darintsev O.V.,
Munasyrov R.A.**

Application of an Associative Controller in the
Control Circuit Of A Mechatron Module..... 183

Dou L.

Method for Information Processing
and Modification of Threshold Parameters
of Rules in a Mivar Expert System Using Large
Language Models 194

**Kovaleva K. V., Kravchenko N. M.,
Bondarchuk V. V., Piguz V. N.,
Izosimova S. A., Ivashko K. S.**

The Results Of The Pilot Operation Of An
Intelligent Computer System For The Diagnosis
And Correction Of Stress Status 206

Kurgan E. G., Petrova K. S., Sobchishen A. S.

Development Of An Electronic Database
As A Tool For Digitalization Of The Central
Scientific And Technical Library 219

Nikitenko K. A., Zviagintseva A. V.

Assessment of semantic similarity of sentences
using probabilistic meausres based
on embeddings..... 229

Filipishin D.A., Zori S.A.

The Resolution Method As A Means Of
Extending The Ontology Inference Machine .. 241

Chernyadev I. V., Yermolenko T. V.

Image Improvement Using Deep Learning:
Analysis And Prospective Developments ... 254

AUTHORS OF THE ISSUE 266

REVIEWERS OF THE ISSUE 268

AUTHOR INDEX..... 269

THE PROCEDURE FOR SENDING
TO THE EDITORIAL, REVIEWING
AND PUBLISHING ARTICLES 273

Подписной индекс журнала
«Проблемы искусственного
интеллекта»:
Министерство связи ДНР
ГП «Почта Донбасса»
28037

Subscription index journal
«Problems of Artificial Intelligence»:
Ministry of Communications
of the Donetsk People's Republic
State enterprise "Post of Donbass"
28037

Научное издание

Международный рецензируемый
научно-теоретический журнал
«Проблемы искусственного интеллекта» /
«Problems of Artificial Intelligence»
ISSN 2413-7383
выходит 4 раза в год

Министерство науки и высшего образования
Российской Федерации
Федеральное государственное бюджетное учреждение
«Институт проблем искусственного интеллекта»

Адрес редакции:
283048, г. Донецк, ул. Артема, 118-б
тел. +38 (062) 311-72-01
e-mail
redakcija_intellekt@mail.ru
Адрес сайта журнала : <http://paijournal.guiaidn.ru/>

Ответственный редактор С. Б. Иванова
Технический редактор В. Н. Пигуз
Компьютерная верстка и корректура
К. С. Ивашко, С. А. Изосимова
Корректура английских текстов А. Л. Червяков
Дизайн обложки С. А. Большакова

**Свидетельство
о регистрации средства массовой информации
Серия ААА № 000114 от 14.03.2017**

Подписано в печать 21.11.2025
Формат 70×108/16
Уч.-изд. л. 24,84. Тираж 100 экз.
Заказ
Бесплатно

Оригинал-макет изготовлен
в ФГБНУ «Институт проблем искусственного интеллекта»
Адрес редакции, издателя, типографии: 283048, г. Донецк,
ул. Артема, дом 118-б, тел. +7 (856) 311-72-01
e-mail: *redakcija_intellekt@mail.ru*,
pvn_redakcija_intellekt@mail.ru

Scientific publication

International peer-reviewed
scientific journal
«Problemy iskusstvennogo intellekta» /
«Problems of Artificial Intelligence»
ISSN 2413-7383
is published four times a year

Ministry of science and higher education
of the Russian Federation
The Federal State Budgetary
Scientific Institution «Institute of Problems
of Artificial Intelligence»

283048, Donetsk, Artema str., 118-b
tel. +38 (062) 311-72-01
e-mail: *iai_sb_ivanova@mail.ru*
maxpvn77@gmail.com
redakcija_intellekt@mail.ru
Journal site address: <http://paijournal.guiaidn.ru/>

Managing editor S. B. Ivanova
Technical editors V. N. Piguz
Computer making-up and proof-reading
K. S. Ivashko, S. A. Izosimova
English proof-reading A. L. Chervyakov
Cover page design S. A. Bolshakova

**Mass Media Registration Certificate
Series AAA № 000114 dated by 14.03.2017**

Passed for printing 21.11.2025.
Format 70×108/16.
Publisher's signature 24,84. Circulation 100 copies.
Order
Free of charge

The original mock-up was made
at the Institute of Artificial Intelligence Problems
The address of the editorial office, publisher, printing house:
283048, Donetsk, Artyom str., 118-b, tel. +7 (856) 311-72-01
e-mail: *redakcija_intellekt@mail.ru*,
pvn_redakcija_intellekt@mail.ru

