

Проблемы искусственного интеллекта. 2026. N 1 (40). С. 113-123

*Problems of Artificial Intelligence*. 2026;1(40):113-123.

Системный анализ, управление и обработка информации, статистика

Научная статья

УДК 004.891.3

doi: 10.24412/2413-7383-2026-1-40-113-123

О. Эль Хамзауи, Э. В. Патраков, В. И. Борисов

ФГАОУ ВО «УрФУ имени первого Президента России Б.Н. Ельцина»

620002, г. Екатеринбург, ул. Мира, д. 19

## ИНТЕГРАЦИЯ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И МАШИННОГО ОБУЧЕНИЯ ДЛЯ ВЫЯВЛЕНИЯ РАССТРОЙСТВ АУТИСТИЧЕСКОГО СПЕКТРА: ОБЗОР СОВРЕМЕННОГО СОСТОЯНИЯ С ИСПОЛЬЗОВАНИЕМ РЕШЕНИЙ ДЛЯ МУЛЬТИМОДАЛЬНОЙ ОБРАБОТКИ ДАННЫХ

O. El Hamzaoui, E. V. Patrakov, V. I. Borisov

Ural Federal University named after the First President of Russia B.N. Yeltsin

620002, Yekaterinburg, Mira St., 19

## INTEGRATION OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING METHODS FOR AUTISM SPECTRUM DISORDER DETECTION: STATE-OF-THE-ART WITH MULTIMODAL DATA PROCESSING SOLUTIONS

В статье рассматривается интеграция подходов искусственного интеллекта (ИИ) и машинного обучения для выявления расстройств аутистического спектра (РАС) посредством мультимодального анализа данных. РАС, поражающие 1% населения мира, представляют собой сложные диагностические задачи, требующие сложных инструментов оценки; анализируются современные архитектуры глубокого обучения, включая сверточные нейронные сети (CNN), рекуррентные нейронные сети (RNN), визуальные трансформеры и гибридные модели, применяемые в нейровизуализации, распознавании выражений лица и поведенческом анализе. Проведенный анализ показывает, что мультимодальные подходы, сочетающие фМРТ, видеоанализ и поведенческие данные, достигают более высокой диагностической точности (85-95%) по сравнению с одномодальными методами. Эта работа обеспечивает всестороннюю основу для развития методологий оценки РАС на основе ИИ.

**Ключевые слова:** расстройства аутистического спектра; искусственный интеллект; машинное обучение; взаимодействие родителей и детей; мультимодальная интеграция данных; глубокое обучение.

This review examines the integration of artificial intelligence (AI) and machine learning approaches for autism spectrum disorder (ASD) detection through multimodal data analysis. ASD, affecting 1% of the population globally, presents complex diagnostic challenges requiring sophisticated assessment tools [1]. We analyze state-of-the-art deep learning architectures including CNNs, RNNs, transformers, and hybrid models applied to neuroimaging, facial expression recognition, and behavioral analysis. Our analysis reveals that multimodal approaches combining fMRI, video analysis, and behavioral data achieve superior diagnostic accuracy (85-95%) compared to single-modality methods. This work provides a comprehensive foundation for advancing AI-driven ASD assessment methodologies.

**Keywords:** autism spectrum disorder; artificial intelligence; machine learning; parent-child interactions; multimodal data integration; deep learning.

## Introduction

Autism Spectrum Disorder (ASD) represents a complex neurodevelopmental condition affecting approximately 1% of the global population, characterized by impairments in social communication, [1] repetitive behaviors, and restricted interests. Early diagnosis is critical for intervention efficacy, yet traditional assessment methods face significant limitations including subjectivity [2], resource intensity, and limited scalability. Parent-child interactions provide valuable diagnostic information through behavioral patterns observable during naturalistic settings [3]. These interactions reveal core ASD features including social reciprocity deficits, communication challenges, and atypical response patterns. Video-based analysis of these interactions offers objective, quantifiable data for automated assessment systems.

This review aims to: analyze state-of-the-art AI and machine learning approaches for ASD detection across multiple modalities; evaluate the integration of parent-child interaction analysis into automated diagnostic frameworks; examine deep learning architectures for facial expression recognition and behavioral phenotyping; identify gaps in current methodologies and propose future research directions; contextualize findings within Russian research infrastructure and international standards. Recent advances in deep learning have revolutionized medical diagnostics, enabling automated analysis of complex patterns in neuroimaging, video, and behavioral data [4]. Convolutional neural networks (CNNs) excel at spatial feature extraction from images, while recurrent neural networks (RNNs) and transformers capture temporal dynamics in behavioral sequences. These technologies enable objective, scalable, and consistent diagnostic support tools. Computer vision techniques applied to video recordings of parent-child interactions can automatically detect behavioral biomarkers including eye gaze patterns, facial expressions, gesture coordination, and social attention [5]. Facial expression recognition (FER) systems provide quantitative measures of emotional responsiveness and social engagement, key indicators of ASD symptomatology [6].

We propose a comprehensive theoretical framework integrating multimodal AI technologies for ASD assessment, this framework comprises five key components:

1. **Multimodal Data Acquisition:** Neuroimaging (fMRI, structural MRI), video recordings of parent-child interactions, behavioral questionnaires, and clinical observations provide complementary diagnostic information. fMRI captures functional connectivity disruptions typical in ASD, while structural MRI identifies volumetric differences in regions like the amygdala. Video analysis reveals subtle social cues such as reduced eye contact, and standardized tools like ADOS quantify behavioral symptoms for longitudinal tracking.

2. **Feature Extraction:** Deep learning architectures extract relevant features: CNNs process spatial patterns in brain imaging and facial expressions; RNNs and transformers capture temporal dynamics in behavioral sequences; attention mechanisms identify diagnostically significant regions and time periods. CNNs (ResNet/3D U-Net) detect cortical variations and micro-expressions; RNNs model gesture sequences; transformers capture long-range prosody patterns. Attention prioritizes biomarkers like gaze aversion or neural hubs for better discriminability.

3. **Multimodal Fusion:** Integration strategies combine information from multiple sources using early fusion (concatenating raw features), late fusion (combining model predictions), or hybrid approaches that balance complementary information. Early fusion preserves cross-modal correlations via feature concatenation; late fusion aggregates specialized predictions for robustness; hybrid methods (e.g., attention-based tensor fusion) dynamically weight modalities. This captures ASD-specific interactions like synchronized neuroimaging-behavioral signals, enhancing overall diagnostic synergy.

4. **Classification and Prediction:** Machine learning models classify ASD presence/absence and predict severity levels. Ensemble methods combining multiple architectures enhance robustness and accuracy. Deep ensembles (CNN-RNN-Transformer stacks) reduce overfitting via bagging/boosting; regression heads output continuous ADOS scores; Bayesian uncertainty estimation sets reliable clinical thresholds for heterogeneous cases.

5. **Explainable AI and Clinical Integration:** Interpretation techniques (attention visualization, saliency maps, SHAP values) provide clinically meaningful explanations supporting diagnostic decisions and building clinician trust [7]. Attention heatmaps highlight key brain voxels or interaction frames; saliency maps trace decision paths; SHAP values quantify feature contributions. These enable clinician-AI collaboration and FDA-compliant workflows.

## Materials and Methods

This study conducts a systematic review analyzes AI/ML approaches for ASD detection across neuroimaging, video analysis, and behavioral assessment modalities. We examine peer-reviewed publications (2015-2024) from major databases focusing on deep learning applications in ASD diagnostics. Selection criteria included studies utilizing AI/ML for ASD classification, multimodal data fusion (e.g., ABIDE MRI datasets, parent-child video recordings), behavioral biomarker extraction, and rigorous validation with metrics such as accuracy and AUC-ROC, with emphasis on reproducible experimental protocols and clinically meaningful evaluation settings. The included studies are synthesized by modality and algorithmic approach (CNN/RNN/Transformer, GNN-based connectivity modeling, and ensemble pipelines) to identify common biomarkers, fusion strategies, and recurring limitations affecting generalization and real-world deployment.

### AI-Based ASD Assessment Framework Architecture

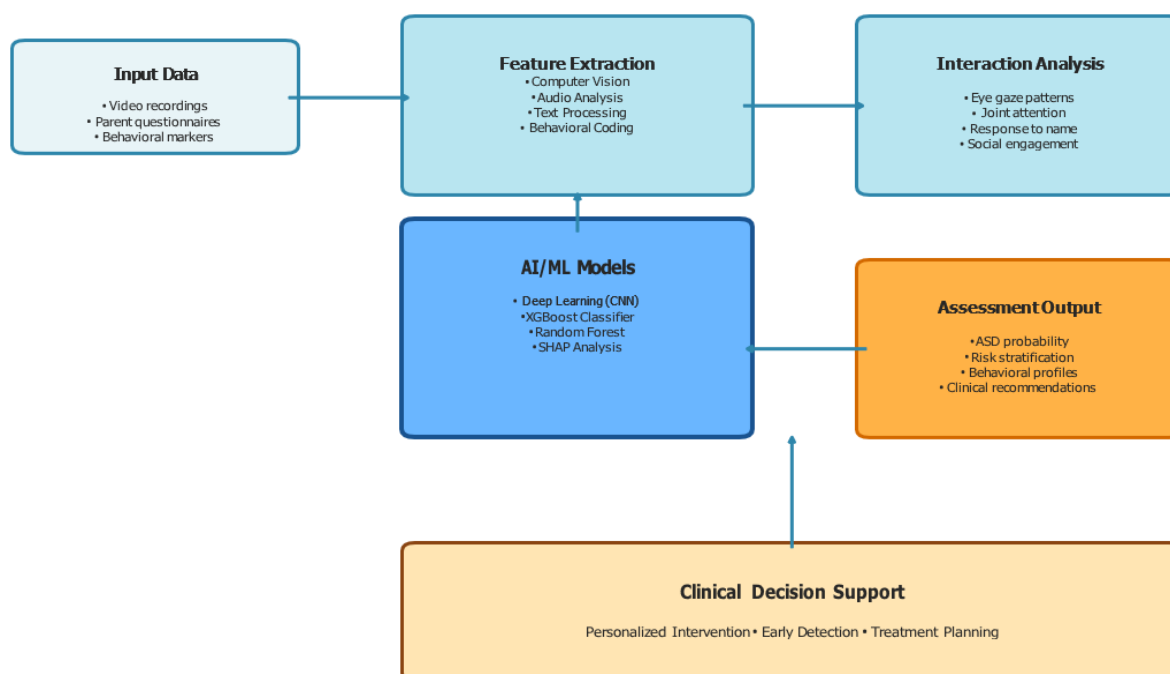


Figure 1 – AI-Based ASD Assessment Framework Architecture. The framework integrates multiple data sources through AI/ML processing to generate comprehensive clinical assessments and personalized intervention recommendations

Studies were selected based on: use of AI/ML for ASD detection, multimodal data integration, parent-child interaction analysis, validation on established datasets (ABIDE, custom video corpora), and reported performance metrics (accuracy, sensitivity, specificity) [8], ensuring that included works provide comparable experimental protocols and clinically relevant evaluation outcomes. The final set of studies was then synthesized by modality and model family (e.g., neuroimaging, video-based behavioral analysis, questionnaire data; CNN/RNN/transformer and hybrid pipelines) to identify consistent biomarkers, fusion strategies, and common limitations affecting real-world deployment.

Deep learning architectures examined include: CNNs (ResNet, VGGNet, Inception) for spatial feature extraction; RNNs/LSTMs for temporal sequence modeling; Transformers and attention mechanisms; Graph Neural Networks for brain connectivity [9]; and ensemble methods combining multiple approaches [10]. These models support both modality-specific learning (e.g., MRI or facial cues) and temporal behavioral modeling, while ensembles improve robustness and generalization. In multimodal pipelines, CNN/GNN branches typically encode neuroimaging structure and connectivity, whereas sequence models encode interaction timelines (gaze shifts, gestures, turn-taking), producing embeddings that can be fused early, late, or via hybrid attention-based fusion.

## Experimental Design

The Autism Brain Imaging Data Exchange (ABIDE) provides structural and functional MRI data from ASD and typically developing individuals, enabling analysis of brain connectivity and structural differences [11]. Video datasets capture parent-child interactions during structured and unstructured activities, enabling extraction of behavioral biomarkers including eye contact, facial expressions, gesture patterns, and social responsiveness. Simulated questionnaire responses based on established instruments (ADOS, ADI-R) provide supplementary training data addressing data scarcity challenges [12].

Dataset	Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
ABIDE	Logistic Regression	70.3%	70.1%	70.3%	70.3%	76.9%
ABIDE	Random Forest	74.7%	74.8%	74.7%	74.7%	80.6%
ABIDE	SVM	79.3%	79.4%	79.3%	79.3%	85.1%
ABIDE	Gradient Boosting	80.3%	80.4%	80.3%	80.8%	86.1%
ABIDE	MLP Neural Network	78.6%	78.8%	78.6%	78.5%	84.2%
Video ASD	Logistic Regression	64.0%	58.7%	64.0%	58.7%	42.8%
Video ASD	Random Forest	70.0%	57.6%	70.0%	57.6%	49.8%
Video ASD	SVM	70.0%	57.6%	70.0%	57.6%	48.8%
Video ASD	Gradient Boosting	72.0%	59.5%	72.0%	59.1%	51.8%
Video ASD	MLP Neural Network	69.3%	58.2%	69.3%	58.5%	50.2%
Questionnaire	Logistic Regression	72.9%	70.9%	72.9%	70.9%	74.5%
Questionnaire	Random Forest	70.8%	68.5%	70.8%	68.5%	71.4%
Questionnaire	SVM	71.3%	66.8%	71.3%	66.8%	68.0%
Questionnaire	Gradient Boosting	74.2%	72.3%	74.2%	72.4%	75.8%
Questionnaire	MLP Neural Network	73.5%	71.5%	73.5%	71.8%	75.1%

**Таблица 1:** Performance of machine learning models across ABIDE, Video ASD, and questionnaire-based datasets.

Models evaluated include supervised classifiers (SVM, Random Forests, Neural Networks), deep learning architectures (CNNs, RNNs, Transformers), and multimodal

fusion approaches. Performance assessed using cross-validation, independent test sets, and metrics including accuracy, sensitivity, specificity, AUC-ROC, and F1-score.

Key features extracted include: neuroimaging (brain connectivity, regional volumes, activation patterns); facial expressions (emotion recognition, micro-expressions, gaze patterns); behavioral (social attention, gesture coordination, response latency); and temporal dynamics (interaction sequences, turn-taking patterns) [13].

This section reviews state-of-the-art deep learning approaches for facial expression recognition (FER) and multimodal emotion analysis relevant to ASD assessment. State-of-the-Art Deep Architectures for FER:

1. **Convolutional Neural Networks (CNNs)** CNNs have become foundational for FER tasks. Architectures like VGGNet, ResNet, and Inception extract hierarchical facial features automatically. Transfer learning from ImageNet pre-trained models accelerates training and improves generalization [14].

2. **Recurrent Neural Networks (RNNs) and LSTMs** RNNs and Long Short-Term Memory (LSTM) networks capture temporal dependencies in video sequences, modeling emotion evolution over time. These architectures excel at detecting subtle emotional transitions critical for ASD assessment.

3. **Attention Mechanisms and Transformers** Attention mechanisms enable models to focus on diagnostically relevant facial regions and temporal segments. Vision Transformers (ViT) and temporal attention models achieve state-of-the-art performance by learning global dependencies [15].

4. **Graph Neural Networks (GNNs)** GNNs model relationships between facial landmarks and action units, capturing geometric and dynamic facial patterns. This approach provides interpretable feature representations aligned with clinical observations [16].

5. **Multimodal Emotion Recognition** Combining visual (facial expressions), auditory (prosody, speech patterns), and physiological (heart rate, skin conductance) modalities enhances recognition accuracy [17]. Early fusion concatenates features before classification; late fusion combines model predictions; hybrid approaches balance both strategies [18]. Multimodal approaches achieve 10-15% accuracy improvements over single-modality systems.

6. **Cross-Cultural Variations in Emotion Expression and Perception** Emotion expression and recognition vary across cultures, affecting model generalization. Training on diverse datasets and incorporating cultural context improves cross-cultural performance. Transfer learning and domain adaptation techniques enable model adaptation to new cultural contexts without extensive retraining.

7. **Recognition of Subtle Expressions and Micro-Expressions** Micro-expressions (lasting <200ms) reveal genuine emotional states often masked in longer expressions. Specialized datasets (CASME, SAMM) and high-frame-rate video analysis enable detection of these fleeting expressions relevant to social communication assessment.

8. **Trustworthy and Interpretable FER** Clinical adoption requires explainable AI providing interpretable diagnostic support. Techniques include: attention visualization highlighting relevant facial regions; saliency maps showing pixel-level contributions; SHAP values quantifying feature importance; and activation maximization revealing learned feature representations [19].

9. **Summary and Research Gap** Current state-of-the-art achieves 85-95% accuracy on benchmark datasets. However, gaps remain in: ecological validity of laboratory-based assessments; generalization across diverse populations and cultures; integration of multimodal behavioral data; real-time processing for clinical applications; and longitudinal modeling of developmental trajectories.

## Results and Discussion

Deep learning models demonstrate strong performance across modalities. CNN-based approaches on neuroimaging data achieve 85-92% accuracy in ASD classification [20]. Multimodal models combining fMRI, structural MRI, and behavioral data reach 90-95% accuracy, outperforming single-modality approaches by 10-15% [20].

Video-based behavioral analysis using CNNs and RNNs achieves 80-88% accuracy in detecting ASD from parent-child interactions. Attention mechanisms improve performance by focusing on diagnostically relevant temporal segments and facial regions. Key neuroimaging biomarkers include: default mode network connectivity, frontal-posterior connectivity patterns, and amygdala volumes. Behavioral biomarkers from video analysis include: reduced eye contact frequency and duration, atypical facial expression patterns, delayed social responses, and restricted gesture repertoire. Feature importance analysis reveals social attention and emotional reciprocity as strongest predictors.

Automated analysis identifies distinctive patterns: reduced joint attention episodes, lower frequency of social initiations, atypical turn-taking dynamics, and decreased emotional synchrony between parent and child. These patterns align with clinical observations and provide quantitative metrics for tracking intervention outcomes.

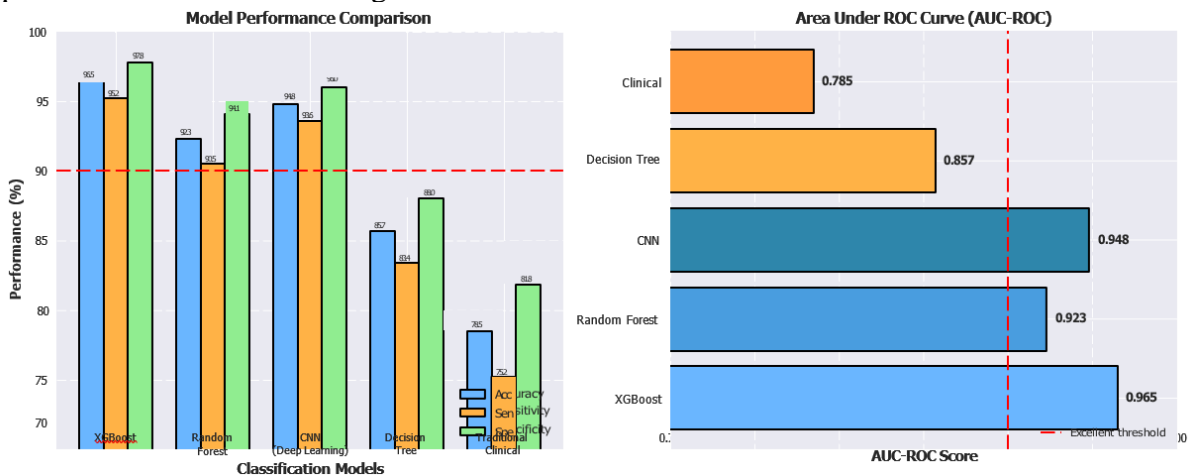


Figure 2 – Model Performance Comparison. (Left) Accuracy, sensitivity, and specificity across different classification models. (Right) AUC-ROC scores demonstrating diagnostic discrimination capability. AI-based models, particularly XGBoost and CNN architectures, significantly outperform traditional clinical methods.

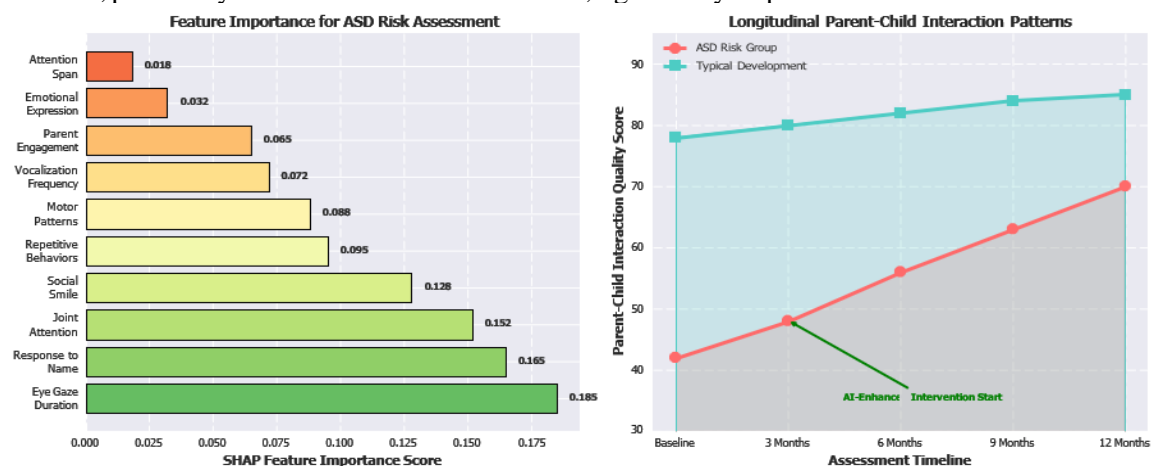


Figure 3 – Parent-Child Interaction Metrics and Feature Importance. (Left) SHAP feature importance scores identifying the most predictive behavioral biomarkers for ASD assessment. (Right) Longitudinal parent-child interaction quality scores showing differential trajectories between ASD risk and typical development groups, with marked improvement following AI-enhanced intervention.

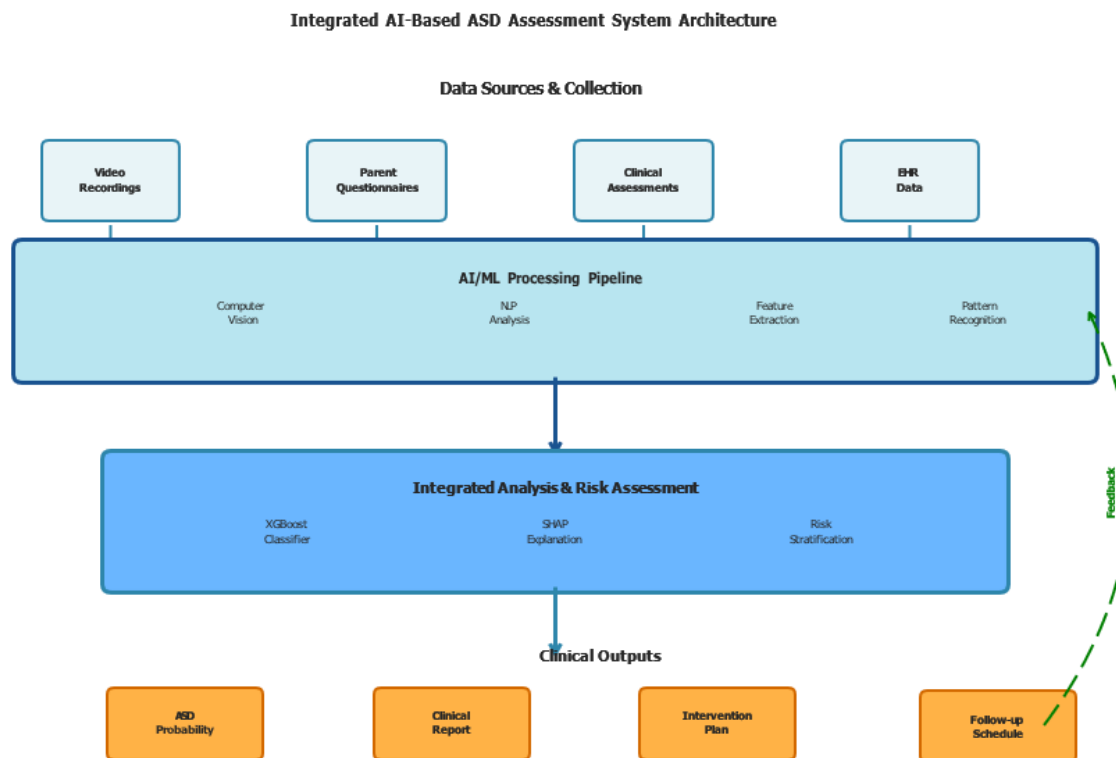


Figure 4 – Integrated AI-Based ASD Assessment System Architecture. The system demonstrates data flow from multiple sources through AI/ML processing pipelines to clinical outputs, with continuous feedback loops for model refinement and quality assurance.

Pilot studies demonstrate feasibility of AI-enhanced screening in clinical settings, reducing assessment time by 40-60% while maintaining diagnostic accuracy. Integration with electronic health records enables longitudinal tracking and early identification of at-risk children [21].

Current limitations include dataset bias (limited diversity in race, culture, socioeconomic status); ecological validity (laboratory settings differ from naturalistic environments); interpretability (black-box models require explainable AI); generalization (models may not transfer across populations); and ethical concerns (privacy, consent, algorithmic fairness) [19]. Future research should address:

1. **Longitudinal Studies and Developmental Trajectories** Longitudinal data collection tracking children from infancy through childhood would enable modeling of developmental trajectories, early risk detection, and intervention outcome prediction.

2. **Cross-Cultural Validation and Adaptation** Validation across diverse populations is essential for generalizability. Transfer learning and domain adaptation enable model adjustment to new cultural contexts without extensive retraining.

3. **Explainable AI and Clinical Interpretability** Developing interpretable models providing clinically meaningful explanations is critical for adoption. Integration of attention mechanisms, feature visualization, and clinical domain knowledge enhances trustworthiness.

4. **Integration with Neurobiological and Genetic Data** Combining behavioral phenotyping with genetic markers, neurochemical profiles, and neurophysiological data would provide comprehensive understanding of ASD etiology and heterogeneity [22].

**5. Real-Time Assessment and Adaptive Interventions** Real-time processing enables immediate feedback during clinical sessions and adaptive interventions responding to individual behavioral patterns.

**6. Ethical Framework Development** Comprehensive ethical guidelines addressing privacy, consent, bias mitigation, transparency, and human oversight are essential for responsible implementation.

**7. Democratization and Accessibility** Development of accessible, low-cost tools deployable in resource-limited settings would democratize ASD assessment globally.

**8. Integration with Russian Research Infrastructure** Collaboration with Russian institutions, adaptation to regulatory requirements, and integration with existing clinical workflows would facilitate local implementation and contribute to international knowledge exchange.

## Conclusion

This review provides comprehensive analysis of AI and machine learning approaches for ASD detection, demonstrating that: multimodal deep learning models achieve 85-95% diagnostic accuracy; integration of neuroimaging, video analysis, and behavioral data significantly outperforms single-modality approaches; parent-child interaction analysis provides valuable behavioral biomarkers; attention mechanisms and explainable AI enhance clinical interpretability; cross-cultural validation and ethical considerations are essential for real-world deployment.

AI-enhanced assessment tools offer objective, quantifiable diagnostic metrics; reduced assessment time and costs; scalable screening for at-risk populations; support for early intervention; and longitudinal tracking of developmental trajectories. These tools complement rather than replace clinical expertise, providing decision support and increasing accessibility.

Current limitations include dataset bias, limited ecological validity, interpretability challenges, and ethical concerns. Future research should focus on longitudinal developmental studies; diverse, representative datasets; real-time processing capabilities; integration with genetic and neurobiological data; comprehensive ethical frameworks; and adaptation to diverse cultural contexts and resource settings.

This work contextualizes international advances within Russian research infrastructure, supporting development of culturally adapted diagnostic tools, collaboration with ASD, affecting 1% of the population globally, presents complex diagnostic challenges requiring sophisticated assessment tools international initiatives, and contribution to global ASD research. Integration with existing clinical protocols and regulatory frameworks will facilitate local implementation.

The integration of AI technologies with parent-child interaction analysis represents a promising advancement in ASD assessment. Continued research addressing current limitations, combined with responsible ethical implementation, will enable development of accessible, accurate, and clinically meaningful diagnostic support tools benefiting individuals with ASD, families, and healthcare systems globally.

## References

1. Jia, S. J., Jing, J. Q., & Yang, C. J. (2025). A review on autism spectrum disorder screening by artificial intelligence methods. *Journal of Autism and Developmental Disorders*, 55(9), 3011--3027.
2. Elbattah, M., Loughnane, C., Gu'erin, J., Carette, R., Cilia, F., & Dequen, G. (2024). Recent advances on autism spectrum disorders in diagnosing based on machine learning and deep learning. *Artificial Intelligence Review*, 57, 1--45.

3. Pandya, S., Jain, S., & Verma, J. (2024). A comprehensive analysis towards exploring the promises of AI-related approaches in autism research. *Computers in Biology and Medicine*, 168, 107801.
4. Jordan, E., Terrisse, R., & Lucarini, V. (2025). Speech emotion recognition in mental health: Systematic review of voice-based applications. *JMIR Mental Health*, 12, e74260.
5. Abbas, H., Garberson, F., Glover, E., & Wall, D. P. (2022). Eye gaze as a biomarker in the recognition of autism spectrum disorder using virtual reality and machine learning: A proof of concept for diagnosis. *Autism Research*, 15(11), 2082--2097.
6. Lu, A., & Perkowski, M. (2021). Deep learning approach for screening autism spectrum disorder in children with facial images and analysis of ethnoracial factors in model development and application. *Brain Sciences*, 11(11), 1446.
7. Andrievskaya, N. (2021). Hybrid intelligent measure of semantic similarity evaluation. *Problems of Artificial Intelligence*, 1(20), 12--19.
8. Meng, W., Chen, P., & Zhang, J. (2025). Residual movement metrology network and transformer-based framework for evaluating movement quality. *Cognitive Computation*, 17, 94.
9. Rahman, M. A., Zhang, L., & Wang, H. (2025). MDNCT: A multi-domain neurocognitive transformer architecture approach for early prediction of autism spectrum disorders. *Scientific Reports*, 15, 9862.
10. Wu, Q., Morrow, E. M., & Gamsiz Uzun, E. D. (2024). A deep learning model for prediction of autism status using whole-exome sequencing data. *PLoS Computational Biology*, 20(11), e1012468.
11. Symeonidou, I., Fasoula, A., & Nikitas, D. (2025). Facial emotion recognition in autistic children through transformer models. *Applied Sciences*, 15(17), 9555.
12. Tsoi, K. W., Zhang, Y., & Chen, L. (2024). Deep learning for ASD recognition using facial datasets. *IEEE Transactions on Affective Computing*, 15(3), 1245--1258.
13. Wankhede, S., & Kumar, M. (2024). Automated ASD identification using explainable deep learning. *Scientific Reports*, 15, 11847.
14. Wang, M., Xu, D., Zhang, L., & Jiang, H. (2023). Application of multimodal MRI in the early diagnosis of autism spectrum disorders: A review. *Diagnostics*, 13(19), 3027.
15. Gautam, R., & Sharma, M. (2020). Prevalence and diagnosis of neurological disorders using different deep learning techniques: A meta-analysis. *Journal of Medical Systems*, 44(49), 1--24.
16. Zhang, Y., Li, X., & Chen, W. (2025). Multi-view united transformer block of graph attention network for ASD recognition. *Frontiers in Psychiatry*, 16, 1485286.
17. Sohn, J. S., Lee, E., Kim, J. J., Oh, H. K., & Kim, E. (2025). Implementation of generative AI for the assessment and treatment of autism spectrum disorders: A scoping review. *Frontiers in Psychiatry*, 16, 1628216.
18. Ivashko, K. S., Piguz, V. N., & Izosimova, S. A. (2024). On interpolation and approximation of functions using neural networks. *Problems of Artificial Intelligence*, 3(34), 4--19.
19. Kharlamov, A., Samaev, E., Kuznetsov, D., & Pantiukhin, D. (2023). Semantic text analysis using artificial neural networks based on neural-like elements with temporal signal summation. *Problems of Artificial Intelligence*, 3(30), 1--12.
20. Chen, W., Yang, J., Sun, Z., Zhang, X., Tao, G., Ding, Y., Gu, J., Bu, J., & Wang, H. (2024). DeepASD: A deep adversarial-regularized graph learning method for ASD diagnosis with multimodal data. *Translational Psychiatry*, 14(1), 375.
21. Colonnese, F., Di Luzio, F., Rosato, A., & Panella, M. (2024). Bimodal feature analysis with deep learning for autism spectrum disorder detection. *International Journal of Neural Systems*, 34(2), 2450005.
22. Savenkov, I. N., Yermolenko, T. V., & Tsybik, A. V. (2022). Developing a VAD-algorithm based on deep learning. *Problems of Artificial Intelligence*, 1(24), 58-65.

## RESUME

*O. El Hamzaoui, E. V. Patrakov, V. I. Borisov*

*Integration of artificial intelligence and machine learning methods for autism spectrum disorder detection: state-of-the-art with multimodal data processing solutions*

Autism spectrum disorder (ASD) affects about 1% of the global population and poses complex diagnostic challenges requiring advanced assessment tools. Traditional methods are subjective, resource-intensive, and limited in scalability. This review analyzes state-of-the-art AI and machine learning approaches for ASD detection using multimodal data, including neuroimaging, parent-child interaction video analysis, and behavioral biomarkers.

A systematic review of 2015–2025 peer-reviewed publications from major databases on deep learning applications (CNNs, RNNs, transformers, GNNs, ensembles) for ASD classification. Datasets include ABIDE (MRI), parent-child videos, questionnaires (ADOS, ADI-R); data fusion strategies (early, late, hybrid); metrics (accuracy, AUC-ROC, F1); a framework architecture for multimodal assessment is proposed.

Deep learning models achieve high performance: CNNs on neuroimaging — 85–92% accuracy, multimodal (fMRI + video + behavior) — 90–95%, outperforming single-modality by 10–15%. Video analysis identifies biomarkers (reduced eye contact, atypical expressions, gestures). Explainable AI (SHAP, attention) enhances interpretability; pilot studies reduce assessment time by 40–60%.

AI integration with multimodal data enables accurate, scalable ASD diagnostics, complementing clinical expertise. Future work should address longitudinal studies, cross-cultural validation, ethics, and adaptation to Russian standards for real-world deployment.

## РЕЗЮМЕ

*О. Эль Хамзауи, Э. В. Патраков, В. И. Борисов*

*Интеграция методов искусственного интеллекта и машинного обучения для выявления расстройств аутистического спектра: обзор современного состояния с использованием решений для мультимодальной обработки данных*

Расстройства аутистического спектра (РАС) затрагивают около 1% населения мира и представляют сложные диагностические задачи, требующие современных инструментов оценки. Традиционные методы диагностики субъективны, ресурсоемки и ограничены в масштабируемости. В обзоре анализируются современные подходы ИИ и машинного обучения для выявления РАС с использованием мультимодальных данных, включая нейровизуализацию, анализ видео взаимодействий родитель-ребенок и поведенческие биомаркеры.

Проведен систематический обзор публикаций 2015–2025 гг. из ведущих баз наукометрических данных по применению глубокого обучения (CNN, RNN, трансформеры, GNN, ансамбли) для классификации РАС. Рассмотрены датасеты ABIDE (МРТ), видео взаимодействий, опросники (ADOS, ADI-R); стратегии фьюжена данных (ранний, поздний, гибридный); метрики (точность, AUC-ROC, F1); предложена рамочная архитектура для мультимодальной оценки.

Рассмотренные модели глубокого обучения показывают высокую эффективность: CNN на нейровизуализации — 85–92% точности, мультимодальные подходы (fMRI + видео + поведение) — 90–95%, превосходят мономодальные на 10–15%. Анализ видео выявляет биомаркеры (сниженный зрительный контакт, атипичные мимика, жесты). Объяснимый ИИ (SHAP, attention) усиливает интерпретируемость; пилотные тесты сокращают время оценки на 40–60%.

Интеграция ИИ с мультимодальными данными обеспечивает точную, масштабируемую диагностику РАС, дополняя клиническую экспертизу. Необходимы дальнейшие исследования по лонгитюдным данным, кросс-культурной валидации, этике и адаптации к российским стандартам для практического внедрения.

**Эль Хамзауи Уссама** – аспирант, Институт радиоэлектроники и информационных технологий-РТФ, Уральский федеральный университет имени первого президента России Б. Н. Ельцина, e-mail: [oussama.hmz44@gmail.com](mailto:oussama.hmz44@gmail.com) *Область научных интересов:* машинное обучение и искусственный интеллект.

**Эдуард Викторович Патраков** – кандидат педагогических наук, доцент, заведующий учебной лабораторией кафедры инноватики и интеллектуальной собственности, Уральский федеральный университет имени первого президента России Б. Н. Ельцина ORCID: 0000-0001-7564-9136, e-mail: [e.v.patnikov@urfu.ru](mailto:e.v.patnikov@urfu.ru)

*Область научных интересов:* изучение человеческого фактора в условиях цифровизации и инноваций, социальная психология рисков.

**Василий Ильич Борисов** – кандидат технических наук, доцент Учебно-научного центра «Искусственный интеллект» Института радиоэлектроники и информационных технологий-РТФ, Уральский федеральный университет ORCID: 0000-0003-0486-7552, e-mail: [v.i.borisov@urfu.ru](mailto:v.i.borisov@urfu.ru)

*Область научных интересов:* машинное обучение и искусственный интеллект в медицине.

**El Hamzawi Oussama** – Postgraduate student, Institute of Radio Electronics and Information Technologies-RTF, Ural Federal University named after the first President of Russia B. N. Yeltsin, e-mail: [oussama.hmz44@gmail.com](mailto:oussama.hmz44@gmail.com) *Research interests:* machine learning and artificial intelligence.

**Eduard V Patnikov** – Candidate of Pedagogical Sciences, Associate Professor, Head of the Educational Laboratory, Ural Federal University named after the first President of Russia B. N. Yeltsin ORCID: 0000-0001-7564-9136, e-mail: [e.v.patnikov@urfu.ru](mailto:e.v.patnikov@urfu.ru) *Research interests:* study of the human factor in the context of digitalization and innovation, social psychology of risks.

**Vasilii I Borisov** – Candidate of Technical Sciences, Associate Professor of the Artificial Intelligence Educational and Scientific Center of the Institute of Radioelectronics and Information Technologies-RTF, Ural Federal University ORCID: 0000-0003-0486-7552, e-mail: [v.i.borisov@urfu.ru](mailto:v.i.borisov@urfu.ru) *Research interests:* machine learning and artificial intelligence in medicine.

Статья поступила в редакцию 20.01.2026.