

Проблемы искусственного интеллекта. 2026. N 2 (41). С. 26-39

Problems of Artificial Intelligence. 2026;2 (41):26-39.

Математическое моделирование, численные методы и комплексы программ

Научная статья

УДК 004.021+004.853

DOI 10.24412/2413-7383-2026-2-41-26-39

Е. А. Хрюкин, Т. В. Мартыненко, Т. А. Васяева

Федеральное государственное бюджетное образовательное учреждение высшего образования «Донецкий национальный технический университет»

283001, г. Донецк, ул. Артёма, д. 58

РАЗРАБОТКА МОДИФИКАЦИИ МЕТОДА МЕШКА СЛОВ ДЛЯ АГРЕГАЦИИ РЕЗУЛЬТАТОВ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ БОЛЬШОГО РАЗМЕРА

E. A. Khryukin, T. V. Martynenko, T. A. Vasyaeva

Federal State Budgetary Educational Institution of Higher Education

“Donetsk National Technical University”

283001, Donetsk, Artema str, 58

DEVELOPMENT OF A MODIFICATION OF THE BAG-OF- WORDS METHOD FOR AGGREGATING CLASSIFICATION RESULTS OF LARGE-SCALE IMAGES

В работе рассматривается задача агрегации результатов классификации фрагментов полнослайдовых медицинских изображений. Предложена модификация метода мешка слов – Weighted Bag-of-Words (WBoW), предназначенная для формирования информативного признакового представления на основе вероятностных предсказаний модели первичной классификации. Проведены экспериментальные исследования с использованием полнослайдовых изображений биопсии предстательной железы. Подтверждена эффективность предложенной модификации метода.

Ключевые слова: полнослайдовые изображения; классификация изображений; агрегация признаков; мешок слов; машинное обучение.

This paper addresses the problem of aggregating classification results of fragments of whole slide medical images. A modification of the Bag-of-Words approach, called Weighted Bag-of-Words (WBoW), is proposed to form an informative feature representation based on probabilistic predictions of the primary classification model. Experimental studies were conducted using histological images of prostate biopsy. The results demonstrate that the proposed WBoW method improves the final classification accuracy compared to the classical Bag-of-Words approach.

Keywords: whole slide images; digital pathology; image classification; feature aggregation; Bag-of-Words; machine learning.

Введение

Современная медицинская диагностика всё активнее опирается на анализ цифровых изображений, получаемых с помощью высокоточного оборудования. Одним из наиболее информативных типов таких данных являются полнослайдовые медицинские изображения (Whole Slide Images, WSI) или высокодетализированные медицинские изображения (ВДМИ), формируемые при сканировании гистологических препаратов с использованием цифровых микроскопов высокого разрешения. Подобные изображения содержат детализированное представление тканей, клеточных структур и патологических изменений на микроуровне, что делает их важным источником информации для клинической диагностики и научных исследований [1-3].

Характерной особенностью таких изображений является их чрезвычайно высокий объём. Размер одного ВДМИ может достигать нескольких гигабайт, что обусловлено высоким пространственным разрешением и необходимостью сохранения мельчайших морфологических деталей. Такие характеристики создают существенные сложности при хранении, передаче и особенно при автоматизированной обработке данных. Прямой анализ полного изображения с использованием алгоритмов машинного обучения часто оказывается вычислительно неэффективным или практически невозможным из-за ограничений по памяти и вычислительным ресурсам [4].

В связи с этим при разработке систем автоматического анализа ВДМИ широко применяется методика, основанная на разбиении исходного изображения на множество фрагментов (тайлов) меньшего размера (в соответствии с рисунком 1). Каждый фрагмент может быть обработан независимо, например с использованием моделей глубокого обучения, предназначенных для решения задач классификации или детекции патологических структур. Это позволяет значительно снизить требования к вычислительным ресурсам и упростить процесс обработки данных [5], [6].

Такая методика приводит к возникновению новой задачи – необходимости интеграции результатов обработки отдельных фрагментов для получения итогового заключения по всему изображению. При этом различные фрагменты могут содержать неодинаковую диагностическую информацию: часть из них может включать патологические структуры, тогда как другие содержат лишь фоновую ткань или артефакты. В результате простые методы объединения результатов, такие как усреднение вероятностей или большинство голосов, не всегда обеспечивают достаточную точность итоговой классификации [7].

Таким образом, актуальной научной задачей является разработка методов агрегации результатов классификации фрагментов полнослайдовых медицинских изображений, обеспечивающих более точное представление структуры изображения и повышение эффективности автоматизированной диагностики.

Постановка задачи

Пусть дано высокодетализированное медицинское изображение (ВДМИ)

$$D = \{d_1, d_2, \dots, d_N\},$$

где N – количество фрагментов (тайлов), на которые разбивается ВДМИ, и которые содержат значимую информацию (прошли пороговую фильтрацию), d_i – i -й тайл (размером 1024×1024 пикселей).

По каждому тайлу d_n модель первичной классификации формирует предсказание S_n , включающее информацию о пяти наиболее вероятных классах изображения $\{t_{1n} \dots t_{5n}\}$ и соответствующих им вероятностях $\{tc_{1n} \dots tc_{5n}\}$. Классы $\{t_{1n} \dots t_{5n}\}$ отсортированы в порядке убывания вероятностей, так, что t_{1n} – это наиболее вероятный класс,

а t_{5n} – наименее вероятный. Требуется на основании всех предсказаний $S = \{S_1, S_2, \dots, S_N\}$ сформировать вектор признаков, пригодный для классификации алгоритмами машинного обучения.

Одним из возможных направлений решения вышеуказанной задачи является использование методов статистического представления данных. В частности, интерес представляет модель «мешка слов» (Bag-of-Words, BoW), широко применяемая в задачах обработки текстовой информации. В рамках данного алгоритма текст представляется в виде набора слов без учёта их порядка, а его содержание описывается распределением частот встречаемости отдельных элементов словаря.

Несмотря на то, что классическая модель мешка слов изначально была разработана для анализа текстовых данных, её концепция может быть адаптирована и для решения задач анализа изображений [8]. В этом случае роль «слов» могут играть локальные элементы изображения или результаты классификации его отдельных фрагментов. Такое представление позволяет формировать обобщённое описание сложного объекта на основе статистического распределения его локальных компонентов [9], [10].

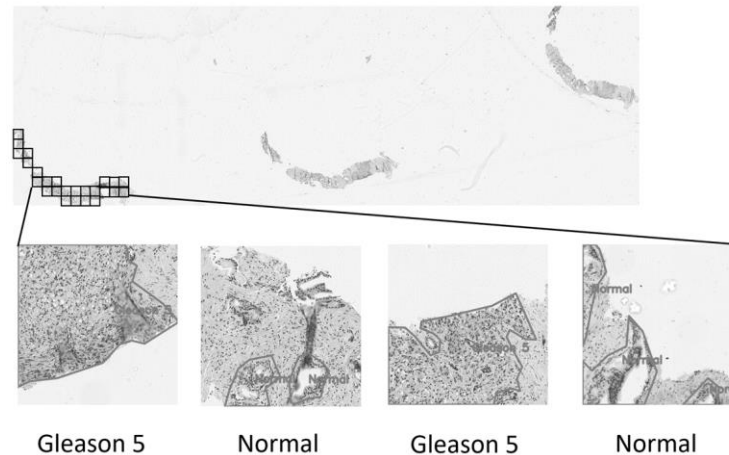


Рисунок 1 – Разбиение изображения на фрагменты с последующей классификацией и сегментацией каждого фрагмента

Применительно к задаче анализа полнослайдовых медицинских изображений подобная методика может быть использована для агрегирования результатов классификации отдельных фрагментов (тайлов) и формирования интегрального представления всего изображения.

Метод BoW в рамках рассматриваемой задачи может быть представлен в следующем виде. ВДМИ D представляется в виде вектора признаков C :

$$C = \{c_1, c_2, \dots, c_M, c_n, c_b\} \in \mathbb{R}^{M+2},$$

где M – количество классов патологий, $c_1, c_2, \dots, c_M$ – множество всех классов патологий, c_n – норма (отсутствие патологии), c_b – фон (нет полезной информации).

Значения элементов вектора C формируются на основе агрегирования результатов классификации фрагментов изображения. Для каждого класса m значение c_m вычисляется как сумма вероятностей tc_{mi} его предсказания по всем фрагментам изображения:

$$c_m = \sum_{i=1}^N tc_{mi},$$

где N – количество фрагментов изображения, tc_{mi} – вероятность принадлежности i го фрагмента классу m .

Полученный вектор C имеет фиксированный размер $M+2$, что делает его пригодным для использования в алгоритмах машинного обучения, например, в методе опорных векторов (Support Vector Machine, SVM).

Требуется разработать модификацию метода BoW, обеспечивающую повышение точности классификации при использовании алгоритмов классического машинного обучения по сравнению с базовой реализацией метода.

Модификация метода мешка слов

В классическом методе BoW агрегация результатов классификации выполняется без учёта взаимного расположения фрагментов, что приводит к потере важной контекстной информации. В настоящей работе предлагается модификация метода мешка слов – взвешенный мешок слов (Weighted Bag-of-Words, WBoW), в рамках которой агрегация результатов классификации фрагментов выполняется с учётом весовых коэффициентов, отражающих вклад отдельных фрагментов в результат классификации.

Следует отметить, что при параллельной обработке фрагментов изображения порядок следования элементов в множестве предсказаний S не соответствует их исходному расположению на изображении. В связи с этим возникает необходимость восстановления пространственной структуры данных. Для решения данной задачи использовался алгоритм кластеризации Density-Based Spatial Clustering of Applications with Noise (DBSCAN), позволяющий выделять группы пространственно близких фрагментов без предварительного задания числа кластеров, а также эффективно обрабатывать шумовые элементы [11].

Кластеризация выполняется на основе координат фрагментов изображения, в результате чего множество предсказаний S разбивается на $\mathcal{L}$ кластеров, каждый из которых соответствует локальной области изображения (в соответствии с рисунком 2). При этом каждому элементу S_i сопоставляется метка кластера, к которому он принадлежит. Дополнительно возможна сортировка элементов вектора S в соответствии с принадлежностью к кластерам.

Использование кластерной структуры позволяет учитывать локальную концентрацию диагностически значимых фрагментов. Для каждого кластера k определяется его размер $\mathcal{K}_k$, равный количеству входящих в него элементов. При формировании итогового признакового вектора предлагается усиливать вклад фрагментов, принадлежащих крупным кластерам, поскольку такие кластеры с большей вероятностью соответствуют значимым патологическим областям.

С учётом вышесказанного, значения вероятностей tc_{mki} , соответствующие элементам i -го кластера, масштабируются с использованием весового множителя, зависящего от количества элементов в кластере. Итоговое выражение для вычисления элементов вектора C принимает вид:

$$C_m = \sum_{k=1}^{\mathcal{L}} \left(\sum_{i=1}^{\mathcal{K}_k} tc_{mki} * \mathcal{K}_k * \mathcal{F}_k \right),$$

где $\mathcal{L}$ – количество кластеров, $\mathcal{K}_k$ – количество элементов в k -м кластере, $\mathcal{F}_k$ – коэффициент кластера, определяемый экспериментально.

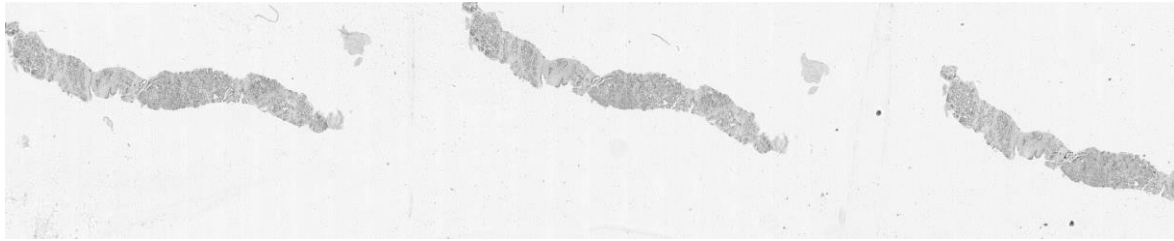


Рисунок 2 – Кластеризация результатов классификации фрагментов изображения методом DBSCAN

Следующим этапом модификации является нормализация агрегированных значений вероятностей, формирующих вектор признаков C . Целью данного шага является учёт частоты появления каждого класса в качестве наиболее вероятного при классификации фрагментов изображения.

Для этого вводится вспомогательный вектор $\mathcal{V}$:

$$\mathcal{V} = \{v_1, v_2, \dots, v_M, v_n, v_b\} \in \mathbb{R}^{M+2},$$

где каждый элемент v_m отражает количество случаев, в которых соответствующий класс был предсказан как наиболее вероятный для фрагментов изображения. Значения элементов вектора $\mathcal{V}$ вычисляются следующим образом:

$$v_m = \sum_{i=1}^N \phi(S_i), \quad \phi(S_i) = \begin{cases} 1, & t_{1i} = m \\ 0, & t_{1i} \neq m \end{cases},$$

где t_{1i} – наиболее вероятный класс для i -го фрагмента.

Далее выполняется нормализация вектора $\mathcal{V}$, позволяющая привести значения его элементов к единой шкале:

$$v_m = \frac{v_m}{\text{sum}(\mathcal{V})}.$$

Полученные значения используются для нормализации элементов вектора признаков C . Итоговые значения вычисляются по формуле:

$$C_m = \frac{c_m}{v_m + 1}.$$

Добавление единицы в знаменатель позволяет избежать деления на ноль в случае, если соответствующий класс ни разу не был предсказан как наиболее вероятный.

Третьим этапом модификации является снижение влияния предсказаний класса нормы в случае наличия патологических фрагментов изображения. Необходимость данного преобразования обусловлена тем, что при анализе ВДМИ значительная часть фрагментов может соответствовать нормальной ткани, что приводит к доминированию данного класса в агрегированном представлении и снижению чувствительности к патологическим изменениям.

Для решения данной проблемы предлагается ввод дополнительного корректирующего преобразования для элемента c_n , соответствующего классу нормы. В случае, если среди наиболее вероятных предсказаний присутствуют патологические классы,

вклад класса нормы уменьшается с использованием специального коэффициента. Преобразование имеет следующий вид:

$$c_n = \frac{c_n}{\varphi(\mathcal{V})}, \quad \varphi(\mathcal{V}) = \begin{cases} \mathcal{F}_N, & \exists v \in \{v_1, v_2, \dots, v_M\} > 1 \\ 1, & \nexists v \in \{v_1, v_2, \dots, v_M\} > 1 \end{cases},$$

где $\mathcal{F}_N$ – определяющий степень подавления вклада класса нормы и подбираемый экспериментально.

Наборы данных

Для проведения экспериментальных исследований был подготовлен набор полнослайдовых изображений биопсийного материала предстательной железы. Исходная выборка включала 114 ВДМИ, распределённых патологоанатомом по диагностическим классам [12], [13] следующим образом: 40 изображений с патологией Gleason 3, 35 изображений с патологией Gleason 4, 11 изображений Gleason 5 и 28 изображений, соответствующих нормальной ткани.

Все изображения были предварительно аннотированы специалистом-патологоанатомом, что позволило выделить области, соответствующие различным типам патологических изменений. Полученные аннотации использовались для формирования обучающих данных.

Поскольку современные модели глубокого обучения, применяемые для анализа изображений, как правило имеют ограничения на размер входных данных (порядка 1024×1024 пикселей), исходные ВДМИ были подвергнуты процедуре фрагментации. Каждое изображение разбивалось на тайлы фиксированного размера. При этом выполнялся автоматический пересчёт координат аннотированных областей и определение принадлежности каждого фрагмента к соответствующему классу патологии [14].

В результате выполненной фрагментации был сформирован набор тайлов, который далее был разделён на обучающую и валидационную выборки. Основные характеристики полученных наборов данных представлены в таблице 1.

Таблица 1 – Характеристики полученного набора данных

Класс изображения	Тренировочная выборка, количество изображений	Валидационная выборка, количество изображений
Gleason 3	4913	722
Gleason 4	6373	690
Gleason 5	1870	278
None	6535	3012
Normal	8948	1249

Модели машинного обучения

Для решения задачи первичной классификации фрагментов изображений было выполнено трансферное обучение (transfer learning) модели семейства YOLO11 [15], а именно архитектуры YOLO11s-cls, ориентированной на эффективную работу при ограниченных вычислительных ресурсах [16–18]. Основные характеристики используемой предобученной модели приведены в таблице 2. Значения ключевых гиперпараметров [19], использованных в процессе обучения модели, приведены в таблице 3.

Таблица 2 – Характеристики предобученной модели

Характеристика	YOLO11s-cls
Размер изображения (пиксели)	224
top-1, %	75.4
top-5, %	92.7
Скорость (CPU, ONNX, мс)	7.9 ± 0.2
Скорость (T4, TensorRT 10, мс)	1.3 ± 0.0
Кол-во параметров (M)	5.5
FLOPs (G)	1.6

Таблица 3 – Гиперпараметры, используемые при обучении модели

Параметр	Назначение	YOLO11s-cls
degrees	Максимальный угол поворота изображения	45
mosaic	Вероятность применения мозаичной аугментации	0.0
erasing	Вероятность случайного "стирания" части изображения	0.0
scale	Коэффициент масштабирования изображения	0.1
batch size	Размер пакета	128

В результате обучения модели первичной классификации была достигнута точность Top-1, равная 0.9136. Полученное значение свидетельствует о высокой способности модели корректно определять наиболее вероятный класс для фрагментов изображений и позволяет использовать её результаты в дальнейшем этапе агрегации предсказаний при анализе полнослайдовых медицинских изображений [20].

Полученные в результате работы модели первичной классификации предсказания для отдельных фрагментов изображения используются на следующем этапе обработки данных. Для формирования интегрального представления полнослайдового изображения применяется предложенная модификация метода мешка слов WBoW. Данный алгоритм выполняет агрегацию вероятностных предсказаний классификатора для всех тайлов изображения и преобразует их в вектор признаков фиксированной размерности [21].

Сформированные таким образом признаковые векторы используются в качестве входных данных для алгоритмов классического машинного обучения, выполняющих итоговую классификацию полнослайдового изображения. В рамках экспериментального исследования были рассмотрены различные модели машинного обучения, включая методы линейной классификации и ансамблевые алгоритмы.

Результаты проведенных экспериментов показали, что наилучшее качество классификации достигается при использовании метода опорных векторов (Support Vector Machine, SVM) с линейным ядром (linear kernel) [22], [23].

Экспериментальные исследования

Для оценки эффективности предложенного метода WBoW была проведена серия экспериментальных исследований. В качестве программной платформы использовался язык программирования Python версии 3.13, а также библиотека scikit-learn (sklearn), предоставляющая реализацию алгоритмов классического машинного обучения и инструменты для статистической оценки моделей.

Для всех полнослайдовых изображений из подготовленного набора данных были получены предсказания обученной модели первичной классификации YOLO11s-cls. Для каждого фрагмента сохранялись результаты классификации, включающие наиболее вероятные классы и соответствующие им вероятности. Полученные данные агрегировались и сохранялись в виде массива, содержащего результаты предсказаний по всем фрагментам каждого изображения.

В качестве метки класса для всего изображения использовался диагностический результат, установленный специалистом-патологоанатомом при анализе соответствующего полнослайдового изображения. Таким образом, для каждого изображения формировалась пара данных: вектор признаков, полученный на основе результатов классификации фрагментов, и соответствующая ему метка класса (норма или патология).

Сформированные массивы векторов признаков и меток классов сохранялись в отдельные файлы и использовались для последующего обучения и тестирования моделей машинного обучения. Для итоговой классификации был выбран метод опорных векторов (Support Vector Machine, SVM) с линейным ядром (linear kernel). Данный алгоритм был выбран на основании предварительных экспериментов, в ходе которых он продемонстрировал наилучшие результаты по сравнению с другими рассмотренными моделями. Оценка качества классификации проводилась с использованием процедуры перекрёстной проверки (cross-validation) с применением функции `cross_val_score` библиотеки `scikit-learn`. В качестве основных метрик качества использовались *ассигасу* и F1-мера.

В качестве базового варианта был рассмотрен наивный алгоритм принятия решения на уровне всего изображения, в рамках которого этап вторичной классификации исключается. Предполагается, что если модель первичной классификации выявляет патологию хотя бы на одном из фрагментов изображения, то результат всего исследования классифицируется как патологический. Экспериментальная оценка показала, что применение такого алгоритма обеспечивает значение метрики *ассигасу*, равное 0.77876, и значение F1-меры, равное 0.50551. Полученные результаты свидетельствуют о том, что простое объединение результатов классификации фрагментов не позволяет достичь высокой точности диагностики и требует применения более информативных методов агрегации.

В ходе экспериментов была проведена оценка классического метода BoW. Полученные результаты показали, что при его использовании средняя точность классификации составила 0.98221, а среднее значение F1-меры — 0.97406.

Далее было исследовано влияние коэффициентов кластера $\mathcal{F}_{\mathcal{K}}$ и нормы $\mathcal{F}_{\mathcal{N}}$ на качество классификации. На первом этапе проводился поиск в широком диапазоне значений от 1 до 20 с шагом 1. Результаты представлены на рисунке 3 в виде трёхмерного графика для значения *ассигасу* и на рисунке 4 в виде тепловой карты для F1-меры.

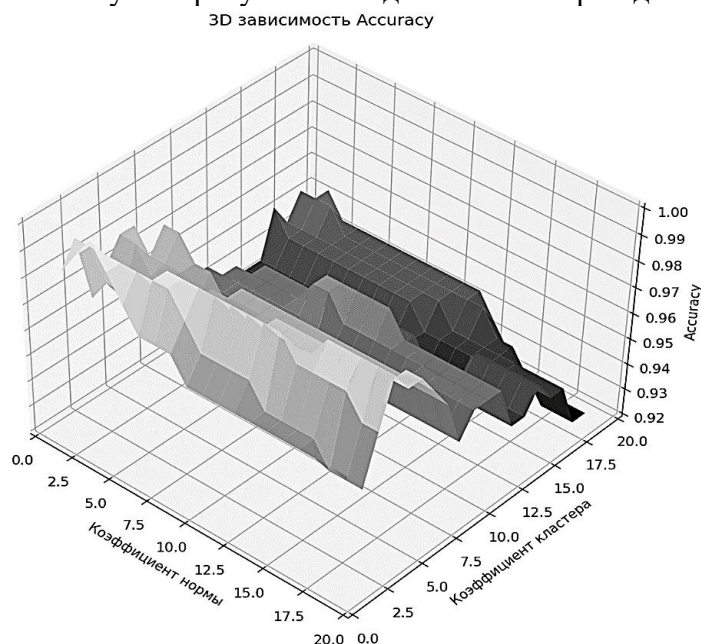


Рисунок 3 – Исследование влияния коэффициентов нормы и кластера на *ассигасу* в диапазоне значений 0-20

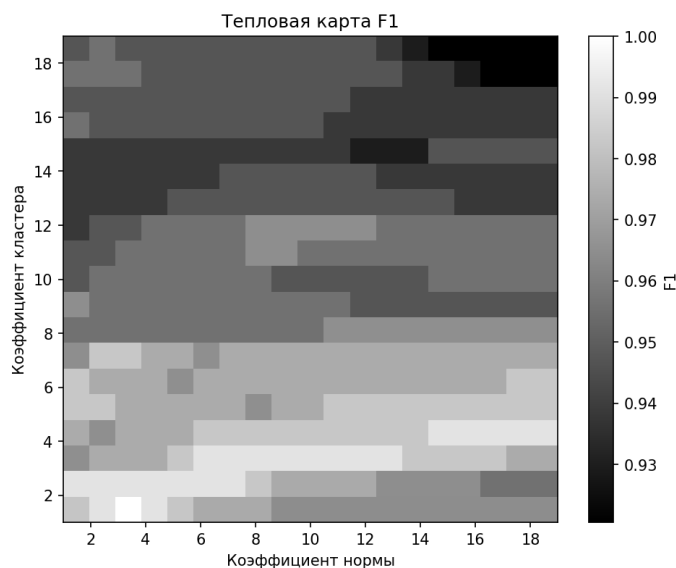


Рисунок 4 – Исследование влияния коэффициентов нормы и кластера на F1-меру в диапазоне значений 0-20

Анализ полученных зависимостей показывает, что наилучшие значения метрик достигаются в области малых значений коэффициентов, близких к 2.

На втором этапе был проведён более детальный анализ в узком диапазоне значений от 0.1 до 3 с шагом 0.05. Результаты представлены на рисунках 5 и 6, где также приведены тепловая карта значений F1-меры и трёхмерная поверхность для ассигасы. Анализ данных визуализаций позволил определить оптимальные значения коэффициентов: $\mathcal{F}_{\mathcal{K}} = 1$ и $\mathcal{F}_{\mathcal{N}} = 1.5$. При указанных параметрах достигается максимальное качество классификации: значение ассигасы равно 0.999, а F1-мера также достигает 0.999.

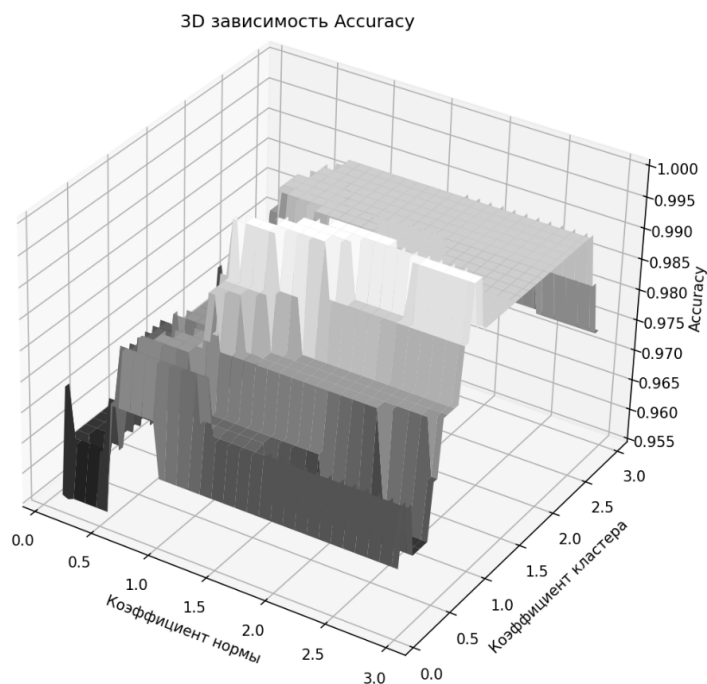


Рисунок 5 – Исследование влияния коэффициентов нормы и кластера на ассигасы в диапазоне значений 0.1-3

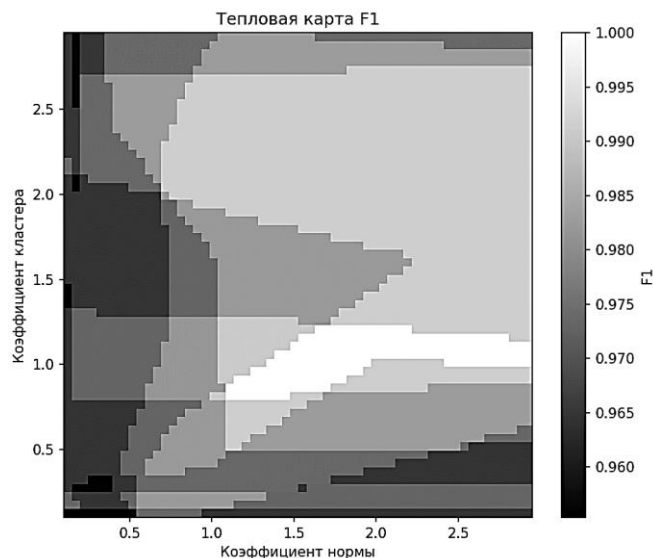


Рисунок 6 – Исследование влияния коэффициентов нормы и кластера на F1-меру в диапазоне значений 0.1-3

Полученные результаты свидетельствуют о том, что использование предложенной взвешенной модификации метода мешка слов (WBoW) позволяет более эффективно агрегировать результаты классификации фрагментов изображения и обеспечивает существенное повышение качества итоговой классификации по сравнению с базовым методом.

Заключение

В работе предложена модификация метода мешка слов – Weighted Bag-of-Words (WBoW), предназначенная для агрегации результатов классификации фрагментов полнослайдовых изображений. Проведённые экспериментальные исследования показали, что использование предложенной модификации позволяет формировать более информативное признаковое представление изображения по сравнению с классическим методом Bag-of-Words.

Экспериментальная оценка продемонстрировала, что применение метода WBoW в сочетании с алгоритмами классического машинного обучения обеспечивает высокое качество итоговой классификации. Наилучшие результаты были получены при использовании метода опорных векторов с линейным ядром: средняя точность классификации и значения F1-меры составили 0.999, что превышает показатели классического метода BoW. Также показано, что предложенная модификация метода существенно превосходит наивный алгоритм принятия решения на уровне всего изображения.

Полученные результаты подтверждают эффективность предложенной модификации метода мешка слов для задачи агрегации результатов классификации фрагментов изображений большого размера.

Список литературы

1. Shafi S., Parwani A.V. Artificial intelligence in diagnostic pathology. *Diagn Pathol.* 2023;18(1).109. DOI: 10.1186/s13000-023-01375-z.
2. Wei J.W., Tafe L.J., Linnik Y.A., Vaickus L.J., Tomita N., Hassanpour S. Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. *Sci Rep.* 2019;9(1).3358 DOI: 10.1038/s41598-019-40041-7.

3. Применение искусственного интеллекта в патологической анатомии / Д. С. Швороб, Т. А. Васяева, Е. А. Хрюкин, А. В. Папакина // Сибирский журнал клинической и экспериментальной медицины. 2025. Т. 40, № 2. С. 211–217. DOI 10.29001/2073-8552-2025-40-2-211-217. EDN VCDXNM.
4. Litjens, G., Kooi, T., Bejnordi, B. E., et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, 42. P. 60–88. DOI: 10.1016/j.media.2017.07.005.
5. Wong S.T. Is pathology prepared for the adoption of artificial intelligence? *Cancer Cytopathol.* 2018;126(6). P. 373–375. DOI: 10.1002/cncy.21994.
6. Искусственный интеллект в диагностике рака предстательной железы / Попов Г.В., Чуб А.А., Лернер Ю.В., Цой Л.В., Дубинина А.В. Варшавский В.А. // Архив патологии. 2021. №83(2). P. 38–45. DOI: 10.17116/patol20218302138.
7. Gertych A., Swiderska-Chadaj Z., Ma Z., Ing N., Markiewicz T., Cierniak S. et al. Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. *Sci Rep.* 2019. № 9(1). P. 1483 DOI: 10.1038/s41598-018-37638-9.
8. Abdullah MMA Al-Omari и Abdullah Noman, «Bag of Visual Words Techniques for Content- Based Image Retrieval and the Role Using in Computer Vision Major: A Survey», сен. 2021, doi: 10.5281/zenodo.5448763.
9. Разработка комплексной методики для аннотирования изображений биопсии / Е. А. Хрюкин, Т. В. Мартыненко, Т. А. Васяева, Д. С. Швороб // Информатика и кибернетика. 2025. № 1(39). С. 33–37. EDN YHQBUX.
10. Хрюкин Е.А. Разработка комплекса программ для анализа медицинских изображений на основе математического моделирования / Хрюкин Е.А., Васяева Т.А., Швороб Д.С. // Материалы IX Международной научно-технической конференции «Современные информационные технологии в образовании и научных исследованиях» «СИТОНИ-2025» (Донецк, 20 ноября 2025 г.) – Донецк: Изд-во ДонНТУ, 2025. Т. 1. С. 152–157.
11. Dubey Dhanraj Shevendrakumar. A STUDY ON DENSITY BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE. *ScienceOpen Posters*. 2023. DOI: 10.14293/P2199-8442.1.SOP-PQUWPP.v1
12. Nguyen TH, Sridharan S, Macias V, Kajdacsy-Balla A, Melamed J, Do MN et al. Automatic Gleason grading of prostate cancer using quantitative phase imaging and machine learning. *J Biomed Opt.* 2017. 22(3). P. 36015. DOI: 10.1117/1.JBO.22.3.036015.
13. Применение искусственного интеллекта в молекулярной и генетической диагностике рака простаты / Морозов А.О., Базаркин А.К., Вовденко С.В., Тараткин М.С., Балашова М.С., Еникеев Д.В. // Вестник урологии. 2024;12(1). P. 117–130. DOI: 10.21886/2308-6424-2024-12-1-117-130.
14. Бабичева, М. В. Автоматизация процедуры распознавания фальшивых изображений посредством нейронных сетей / М. В. Бабичева, И. А. Третьяков // Проблемы искусственного интеллекта. 2025. № 1(36). С. 94–105. DOI 10.24412/2413-7383-94-105. EDN GNSJQU.
15. Jocher G., Qiu J. Ultralytics YOLO11 [Электронный ресурс] / G. Jocher, J. Qiu. Version 11.0.0. Ultralytics, 2024. Режим доступа: <https://docs.ultralytics.com/models/yolo11/>
16. Швороб, Д. С. Автоматизированная система диагностики аденокарциномы предстательной железы на базе искусственного интеллекта / Д. С. Швороб, Е. А. Хрюкин // Клиническая и экспериментальная морфология. 2025. Т. 14, № 6. С. 71–78. DOI 10.31088/CEM2025.14.6.71-78. EDN DOENOG.
17. OpenCV Documentation. Open Source Computer Vision Library. Интернет-ресурс. Режим доступа : [www/](http://www.opencv.org/) URL: <https://docs.opencv.org/>.
18. Large-Image Documentation. Kitware, Inc. <https://doi.org/10.5281/zenodo.14624225> Интернет-ресурс. - Режим доступа : [www/](http://www.girder.github.io/large_image/) URL: https://girder.github.io/large_image/
19. Зуев, В. М. Ортогонализация весовых коэффициентов нейросети / В. М. Зуев // Проблемы искусственного интеллекта. 2025. № 1(36). С. 135–140. DOI 10.24412/2413-7383-135-140. EDN CRADNZ.
20. Bulten W, Kartasalo K, Chen PC, Ström P, Pinckaers H, Nagpal K et al. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat Med.* 2022. 28(1). P. 154–63. DOI: 10.1038/s41591-021-01620-2.
21. Никитенко, К. А. Интерпретируемость нейросемантических моделей при их применении в прикладных областях / К. А. Никитенко, А. В. Звягинцева // Проблемы искусственного интеллекта. 2025. № 2(37). С. 79–90. DOI 10.24412/2413-7383-2025-2-37-79-90. EDN VFRDWH.
22. Каспаров, Ю. А. Обучение моделей компьютерного зрения для задачи детекции и сегментации стволовых клеток / Ю. А. Каспаров, В. М. Зуев // Проблемы искусственного интеллекта. 2025. № 3(38). С. 183–189. DOI 10.24412/2413-7383-2025-3-38-183-189. EDN KQRFEV.
23. Применение ансамблей нейронных сетей типа Lenet-5 для диагностирования пневмонии на рентгеновских изображениях легких / Д. Н. Тумаков, А. Р. Бердигулов, Д. И. Хайруллина, И. И. Бариев // Проблемы искусственного интеллекта. 2025. № 3(38). С. 34–50. DOI 10.24412/2413-7383-2025-3-38-34-50. EDN URVNAV.

References

1. Shafi S., Parwani A.V. Artificial intelligence in diagnostic pathology. *Diagn Pathol.* – 2023. – № 18(1). – 109 p. – DOI: 10.1186/s13000-023-01375-z.
2. Wei J.W., Tafe L.J., Linnik Y.A., Vaickus L.J., Tomita N., Hassanpour S. Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks // *Sci Rep.* – 2019. – № 9(1). – 3358 p. – DOI: 10.1038/s41598-019-40041-7.
3. Shvorob, D. S. Automated system for prostate adenocarcinoma diagnosis based on artificial intelligence / D. S. Shvorob, E. A. Khryukin // *Clinical and Experimental Morphology.* – 2025. – Vol. 14, No. 6. – Pp. 71–78. – DOI: 10.31088/CEM2025.14.6.71-78. – EDN DOEHOG.
4. Litjens, G., Kooi, T., Bejnordi, B. E., et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, 42: 60–88. DOI: 10.1016/j.media.2017.07.005.
5. Wong S.T. Is pathology prepared for the adoption of artificial intelligence? *Cancer Cytopathol.* 2018;126(6):373–375. DOI: 10.1002/cncy.21994.
6. Popov G. V., Chub A. A., Lerner Yu. V., Tsoy L. V., Dubinina A. V., Varshavsky V. A. Artificial intelligence in the diagnosis of prostate cancer // *Archive of Pathology.* – 2021. – Vol. 83, No. 2. – Pp. 38–45. – DOI: 10.17116/patol20218302138.
7. Gertych A., Swiderska-Chadaj Z., Ma Z., Ing N., Markiewicz T., Cierniak S. et al. Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides // *Sci Rep.* – 2019. – № 9(1). – 1483 p. – DOI: 10.1038/s41598-018-37638-9.
8. Abdullah MMA Al-Omari и Abdullah Noman, «Bag of Visual Words Techniques for Content- Based Image Retrieval and the Role Using in Computer Vision Major: A Survey». – Сен. 2021. – doi: 10.5281/zenodo.5448763.
9. Development of a comprehensive methodology for biopsy image annotation / E. A. Khryukin, T. V. Martynenko, T. A. Vasyaeva, D. S. Shvorob // *Informatics and Cybernetics.* – 2025. – No. 1(39). – Pp. 33–37. – EDN YHQUBX.
10. Khryukin E. A. Development of a software complex for medical image analysis based on mathematical modeling / E. A. Khryukin, T. A. Vasyaeva, D. S. Shvorob // *Proceedings of the IX International Scientific and Technical Conference Modern Information Technologies in Education and Scientific Research (SITONI-2025) (Donetsk, November 20, 2025).* – Donetsk: DonNTU Publishing House, 2025. Vol. 1. Pp. 152–157.
11. Dubey Dhanraj Shevendrakumar. A STUDY ON DENSITY BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE. *ScienceOpen Posters*. 2023. DOI: 10.14293/P2199-8442.1.SOP-PQUWPP.v1
12. Nguyen TH, Sridharan S, Macias V, Kajdacsy-Balla A, Melamed J, Do MN et al. Automatic Gleason grading of prostate cancer using quantitative phase imaging and machine learning. *J Biomed Opt.* – 2017 – № 22(3). – 36015 p. – DOI: 10.1117/1.JBO.22.3.036015.
13. Morozov A. O., Bazarkin A. K., Vovdenko S. V., Taratkin M. S., Balashova M. S., Enikeev D. V. Application of artificial intelligence in molecular and genetic diagnosis of prostate cancer // *Urology Herald.* – 2024. – Vol. 12, No. 1. – Pp. 117–130. – DOI: 10.21886/2308-6424-2024-12-1-117-130.
14. Babicheva, M. V. Automation of fake image recognition using neural networks / M. V. Babicheva, I. A. Tretyakov // *Problems of Artificial Intelligence.* – 2025. – No. 1(36). – Pp. 94–105. – DOI: 10.24412/2413-7383-94-105. – EDN GNSJQU.
15. Jocher G., Qiu J. Ultralytics YOLO11 [Electronic resource] / G. Jocher, J. Qiu. – Version 11.0.0. – Ultralytics, 2024. – Available at : <https://docs.ultralytics.com/models/yolo11/>
16. Application of artificial intelligence in pathological anatomy / D. S. Shvorob, T. A. Vasyaeva, E. A. Khryukin, A. V. Papakina // *Siberian Journal of Clinical and Experimental Medicine.* – 2025. – Vol. 40, No. 2. – Pp. 211–217. – DOI: 10.29001/2073-8552-2025-40-2-211-217. – EDN VCDXNM.
17. OpenCV Documentation. Open Source Computer Vision Library. Electronic resource. – Available at : [www/ URL: https://docs.opencv.org/](http://www.opencv.org/).
18. Large-Image Documentation. Kitware, Inc. <https://doi.org/10.5281/zenodo.14624225> [Electronic resource]. – Available at : [www/ URL: https://girder.github.io/large_image/](http://www.girder.github.io/large_image/)
19. Zuev, V. M. Orthogonalization of neural network weight coefficients // *Problems of Artificial Intelligence.* – 2025. – No. 1(36). – Pp. 135–140. – DOI: 10.24412/2413-7383-135-140. – EDN CRADNZ.
20. Bulten W., Kartasalo K., Chen P.C., Ström P., Pinckaers H., Nagpal K. et al. Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge // *Nat Med.* – 2022. – № 28(1). – P. 154–63. DOI: 10.1038/s41591-021-01620-2.
21. Nikitenko, K. A. Interpretability of neurosemantic models in applied domains / K. A. Nikitenko, A. V. Zvyagintseva // *Problems of Artificial Intelligence.* 2025. No. 2(37). Pp. 79–90. DOI: 10.24412/2413-7383-2025-2-37-79-90. – EDN VFRDWH.
22. Kasparov, Yu. A. Training computer vision models for the detection and segmentation of stem cells / Yu. A. Kasparov, V. M. Zuev // *Problems of Artificial Intelligence.* – 2025. – No. 3(38). – Pp. 183–189. – DOI: 10.24412/2413-7383-2025-3-38-183-189. – EDN KQRFEV.
23. Application of LeNet-5 type neural network ensembles for pneumonia diagnosis on chest X-ray images / D. N. Tumakov, A. R. Berdigulov, D. I. Khairullina, I. I. Bariev // *Problems of Artificial Intelligence.* – 2025. – No. 3(38). – Pp. 34–50. – DOI: 10.24412/2413-7383-2025-3-38-34-50. – EDN URVNAV.

RESUME

E. A. Khryukin, T. V. Martynenko, T. A. Vasyaeva

Development of a modification of the bag-of-words method for aggregating classification results of large-scale images

The analysis of whole slide histological images is an important task in medical diagnostics. A key characteristic of such images is their extremely large size, which makes it impossible to directly apply classification models to the entire image. Therefore, whole slide images are typically divided into smaller fragments that are processed individually, after which the results must be aggregated to obtain a final diagnostic decision for the entire image.

The study was conducted using a dataset of whole slide histological images with diagnostic labels established by a pathologist. At the first stage, the images were divided into fragments, and primary classification was performed using the YOLO11s-cls model. For each fragment, the most probable classes and their corresponding probabilities were recorded. To aggregate the classification results, a modification of the Bag-of-Words (BoW) method, called Weighted Bag-of-Words (WBoW), is proposed. Unlike the classical BoW approach, the proposed method takes into account the probabilistic estimates of the classes and introduces a coefficient that reduces the contribution of fragments corresponding to normal tissue in the presence of pathological features. The resulting feature representation was used as input for classical machine learning algorithms. A Support Vector Machine with a linear kernel was employed as the final classification model.

Experimental results demonstrate that the proposed method provides higher classification performance compared to both the classical Bag-of-Words approach and a naive aggregation strategy. The best results were obtained using the WBoW representation with an SVM classifier: the average classification accuracy and F1-score both reached 0.999.

These results confirm the effectiveness of the proposed modification of the BoW method for aggregating fragment-level classification results of large-scale images.

РЕЗЮМЕ

Е. А. Хрюкин, Т. В. Мартыненко, Т. А. Васяева

Разработка модификации метода мешка слов для агрегации результатов классификации изображений большого размера

Анализ полнослайдовых гистологических изображений является актуальной задачей медицинской диагностики. Существенной особенностью таких изображений является их большой размер, что не позволяет напрямую использовать модели классификации для обработки всего изображения целиком. В связи с этим изображения разбиваются на отдельные фрагменты, для которых выполняется первичная классификация, после чего возникает задача агрегирования полученных результатов для формирования итогового диагностического решения.

В качестве исходных данных использовался набор полнослайдовых изображений гистологических препаратов с известным диагностическим заключением, установленным специалистом-патологоанатомом. Первичная классификация фрагментов выполнялась с использованием модели YOLO11s-clс, для каждого фрагмента сохранялись наиболее вероятные классы и соответствующие вероятности. Для агрегирования результатов предложена модификация метода мешка слов – Weighted Bag-of-Words (WBoW). В отличие от классического метода Bag-of-Words (BoW), предложенный метод учитывает вероятностные оценки классов и вводит коэффициент, позволяющий уменьшать вклад фрагментов нормальной ткани при наличии признаков патологии. Сформированное признаковое представление использовалось в качестве входных данных для алгоритмов классического машинного обучения. В качестве алгоритма итоговой классификации применялся метод опорных векторов с линейным ядром.

Результаты экспериментальных исследований показали, что предложенный метод обеспечивает более высокое качество классификации по сравнению с классическим методом BoW и наивным алгоритмом агрегации результатов. Наилучшие показатели были получены при использовании WBoW в сочетании с SVM: средняя точность классификации и значения F1-меры составили 0.999.

Полученные результаты подтверждают эффективность предложенной модификации метода мешка слов для задачи агрегирования результатов классификации фрагментов полнослайдовых изображений большого размера.

Хрюкин Евгений Александрович – аспирант кафедры автоматизированных систем управления ФГБОУ ВО «Донецкий национальный технический университет», 283001, г. Донецк, ул. Артема, д. 58, телефон +7(949) 421-77-84, wadealer@yandex.ru, *Область научных интересов:* математическое моделирование, цифровая обработка изображений, нейронные сети, искусственный интеллект. ORCID 0009-0001-2070-2371

Мартыненко Татьяна Владимировна – доцент, кандидат технических наук, доцент кафедры автоматизированных систем управления ФГБОУ ВО «Донецкий национальный технический университет», 283001, г. Донецк, ул. Артёма, 58, телефон +7(949) 413-63-76, tatyana.v.martynenko@yandex.ru. *Область научных интересов:* искусственный интеллект, методы оптимизации, глубокое обучение, цифровая обработка изображений. ORCID 0000-0002-1483-8483

Васяева Татьяна Александровна – доцент, кандидат технических наук, декан факультета информационных систем и технологий ФГБОУ ВО «Донецкий национальный технический университет», 283001, г. Донецк, ул. Артёма, 58, телефон +7(949) 334-91-76. fist@donntu.ru. *Область научных интересов:* анализ и прогнозирование временных рядов, эволюционные вычисления, глубокое обучение для компьютерного зрения, обработка естественного языка, проектирование мультимодальных AI-ассистентов. ORCID 0000-0001-9362-2279

Khriukin Evgeniy Alexandrovich – Postgraduate Student of Automated Control Systems Department of Donetsk National Technical University, 283001, Donetsk, Artema St. 58, phone +7(949) 421-77-84, wadealer@yandex.ru, Research interests: mathematical modeling, digital image processing, neural networks, artificial intelligence. ORCID 0009-0001-2070-2371

Martynenko Tatyana Vladimirovna – Ph.D. in Engineering, Associate Professor, Associate Professor of the Department of Automated Control Systems, Donetsk National Technical University, 283001, Donetsk, Artema St. 58, phone +7(949) 413-63-76, tatyana.v.martynenko@yandex.ru. Research interests: artificial intelligence, optimization methods, deep learning, digital image processing. ORCID: 0000-0002-1483-8483

Vasyaeva Tatyana Alexandrovna – Ph.D. in Engineering, Associate Professor, Dean of the Faculty of IST of Donetsk National Technical University Donetsk National Technical University, 283001, Donetsk, Artema St. 58, phone +7(949) 334-91-76. fist@donntu.ru. Research interests: time series analysis and forecasting, evolutionary computation, deep learning for computer vision, natural language processing, design of multimodal AI assistants. ORCID 0000-0001-6578-0050

Статья поступила в редакцию 18.03.2026.