

УДК 004.032.26

DOI: [10.26102/2310-6018/2025.49.2.042](https://doi.org/10.26102/2310-6018/2025.49.2.042)

Применение моделей машинного обучения семейства YOLO для задачи анализа чайного сырья по фотографии

В.А. Попов¹✉, А.В. Зубков^{1,2}

¹Волгоградский государственный технический университет, Волгоград,
Российская Федерация

²Волгоградский государственный медицинский университет, Волгоград,
Российская Федерация

Резюме. В статье проведен концепт решения задачи анализа чайного сырья по фотографии с использованием моделей семейства YOLO, а также их сравнительный анализ различных версий двух моделей YOLOv8: Nano и Small. В статье выделены метрики для сравнительного анализа. Было проведено экспериментальное сравнение характеристик обеих моделей на реальных примерах чайного сырья. Для целей исследования был собран тренировочный набор данных, который содержал в себе изображения сырья чая следующих классов ферментации: зеленый чай, красный чай, белый чай, желтый чай, улун, шу пуэр, шен пуэр. Для увеличения количества тренировочных образцов были применены методы аугментации, такие как поворот изображения, усиление резкости, искажение перспективы и размытие. Из результатов эксперимента следует, что выбор между двумя представленными моделями обуславливается решаемой задачей и количеством вычислительных ресурсов. YOLOv8s (Small) превосходит YOLOv8n (Nano) в точности, но YOLOv8n расходует меньше времени для предоставления результатов и может использоваться в условиях ограниченного количества вычислительных ресурсов, что делает ее особенно предпочтительной при обработке больших объемов данных.

Ключевые слова: анализ изображений, машинное обучение, компьютерное зрение, чайное сырье, сверточные нейронные сети.

Для цитирования: Попов В.А., Зубков А.В. Применение моделей машинного обучения семейства YOLO для задачи анализа чайного сырья по фотографии. *Моделирование, оптимизация и информационные технологии*. 2025;13(2). URL: <https://moitvvt.ru/ru/journal/pdf?id=1938> DOI: 10.26102/2310-6018/2025.49.2.042

Application of YOLO family machine learning models for the task of analyzing tea raw materials by photograph

V.A. Popov¹✉, A.V. Zubkov^{1,2}

¹Volgograd State Technical University, Volgograd, the Russian Federation

²Volgograd State Medical University, Volgograd, the Russian Federation

Abstract. This article presents a concept for analyzing tea raw materials using the YOLO family of models, as well as comparative analysis of two versions of YOLOv8: Nano and Small. The study highlights metrics used to compare these models' performance. An experimental comparison was conducted on real examples of tea raw material images. For this purpose, a training dataset was collected containing images of tea samples classified by fermentation type: green tea, red tea, white tea, yellow tea, oolong, shou puerh, and sheng puerh. To increase the number of training samples, augmentation methods were applied such as image rotation, sharpening, perspective distortion, and blurring. Based on the experiment results, it is concluded that choosing between the two presented models depends on the task at hand and available computational resources. YOLOv8s (Small) outperforms YOLOv8n (Nano) in terms of accuracy but consumes more time to provide results. On the other hand, YOLOv8n processes

data faster and can be effectively utilized under limited computing power conditions, making it particularly suitable for handling large volumes of data.

Keywords: image analysis, machine learning, computer vision, tea raw material, convolutional neural networks.

For citation: Popov V.A., Zubkov A.V. Application of YOLO family machine learning models for the task of analyzing tea raw materials by photograph. *Modeling, Optimization and Information Technology*. 2025;13(2). (In Russ.). URL: <https://moitvivr.ru/ru/journal/pdf?id=1938> DOI: 10.26102/2310-6018/2025.49.2.042

Введение

Китайская культура чаепития известна своим глубоким пониманием и ценностью уникальных качеств каждого сорта чая [1]. Важность правильного выбора сорта обусловлена влиянием внешних признаков чайного сырья на восприятие конечного продукта потребителем. Настоящая статья посвящена изучению возможностей снижения уровня сложности входа в чайную культуру путем внедрения цифровых помощников, основанных на методах машинного обучения.

Чайное сырье представлено свежими листьями чайного дерева *Camellia sinensis*, собранными в различных регионах Китая и прошедшими разнообразные процессы обработки [2]. Внешний вид чайного сырья играет ключевую роль в определении качества и разновидности чая [3]. Например, зеленый чай сохраняет естественный оттенок свежей зелени, демонстрируя минимальное изменение структуры листа благодаря отсутствию значительной ферментации [4]. Улун, напротив, демонстрирует переходный спектр цветов, обусловленных степенью окисления, варьируя от светло-зеленых до темно-коричневых оттенков. Для черного чая характерны темные тона, достигаемые полным процессом ферментации. Особенности формы и структуры листа играют важную роль в идентификации сорта [5, 6]. Так, белые чаи представлены мягкими почками, покрытыми тонким слоем белого ворса, создавая ощущение бархатистости. Зеленые чаи зачастую обладают гладкой поверхностью, тогда как черный чай приобретает ломкую и сыпучую текстуру после полного ферментирования. Пуэры демонстрируют сложную текстуру, включающую пористую поверхность и неоднородность материала, формирующегося в процессе старения.

Эти внешние признаки служат основой для оценки качества и классификации чайного сырья специалистами отрасли и опытными потребителями. Однако традиционная идентификация сортов требует значительных усилий и специализированного опыта, ограничивая доступность глубокого понимания свойств чая широкой аудиторией [7].

Развитие технологий компьютерного зрения предлагает новый подход к решению данной проблемы [8, 9]. Методы машинного обучения позволяют автоматически классифицировать образцы чайного сырья, используя характерные визуальные признаки, такие как цвет, форма и структура листьев [10]. Одним из перспективных подходов являются алгоритмы семейства YOLO (You Only Look Once). Эта группа моделей способна эффективно анализировать изображения и определять ключевые характеристики объектов, предлагая решение, которое сочетает высокую точность и скорость обработки.

Цель настоящего исследования заключается в разработке системы, основанной на моделях YOLO, позволяющей автоматизировать распознавание сортов китайского чая по изображениям. Данная система должна обеспечить быстрое и точное определение сортов чая, позволяя пользователям получать информацию о продукте без необходимости глубоких специализированных знаний. Таким образом, внедрение такой

технологии способствует упрощению восприятия чайной культуры и повышению доступности высококачественных продуктов среди широких слоев населения.

Архитектура YOLO (You Only Look Once) представлена разными модификациями, что позволяет использовать ее в различных условиях для решения задач различной степени сложности [11]. Регулярные обновления для моделей этого семейства улучшают их работоспособность, расширяя спектр применения, предоставляя возможности использования на различных устройствах: от мобильных устройств до высокопроизводительных серверов [12].

Реализация в виде фреймворка уже давно используется в научных исследованиях, связанных с машинным обучением. Так, например, модель YOLOv8 была использована для обнаружения дорожных знаков в реальном времени [13]. Обученная модель предлагала новый метод контроля за безопасностью дорожного движения.

В данном исследовании будут рассматриваться две модификации модели YOLOv8 – Nano (YOLOv8n) и Small (YOLOv8s), которые отличаются размером и вычислительными требованиями. Основной целью исследования является анализ их производительности на основе таких метрик, как mAP (mean Average Precision), потери обучения и время, необходимое для обучения [14].

Модификация YOLOv8 Nano характеризуется компактностью и низкой сложностью, что делает её привлекательной для внедрения в устройства с ограниченным объемом памяти и мощностью процессора. Модель YOLOv8 Small отличается чуть большим числом параметров и, соответственно, способна достигать большей точности, однако требует больше вычислительной мощности.

Дополнительной целью исследования является всесторонний сравнительный анализ обеих версий модели по ряду ключевых критериев, определяющих их практическое применение:

- mAP (Mean Average Precision) – средняя точность выявления объектов, характеризующая общую способность модели точно определять местоположение и класс обнаруженных объектов.
- Потери обучения (Training Loss) – показатель скорости и эффективности процесса оптимизации весов модели.
- Время обучения (Training Time) – длительность процесса обучения, которая измеряется в секундах или минутах, позволяющая оценить затраты ресурсов на обучение каждой модели.

Сбор данных

Для целей настоящего исследования был специально сформирован уникальный датасет, состоящий из фотографий шести популярных сортов чая – красного, желтого, белого, улуна, шу пуэра и шен пуэра. Он был разделен на тренировочный набор, тестовый набор и валидационный набор. Тренировочный набор предназначен для обучения модели, тестовый набор отвечает за контроль работоспособности модели путем проведения испытаний на нем, а валидационный набор задействован в процессе кросс-валидации, при котором датасет многократно делят на обучающие и тестовые подмножества.

Изначально полученный объем изображений оказался недостаточен для эффективного обучения глубокой нейронной сети, необходимой для решения поставленной задачи классификации сортов чая. Понимая важность репрезентативности данных и следуя рекомендациям предыдущих исследований, авторы решили применить методы аугментации изображений.

Методология аугментации основывалась на технике, представленной в работе [15], где были предложены различные подходы к изменению исходных изображений. Этот подход позволил увеличить число уникальных примеров в датасете, расширив возможности обучения модели и снизив риск возникновения эффекта переобучения.

Кроме количественного роста числа тренировочных образов, аугментация также позволила решить проблему повышения качества и надежности итогового результата. Моделируя разнообразные сценарии съемки (например, поворот и смещение объектов), увеличение размера выборки помогло минимизировать влияние фона и внешних факторов на конечный результат классификации. Помимо того, важным аспектом стало повышение устойчивости модели к небольшим изменениям ориентации чайных листьев и чашечек. Таким образом, улучшилась возможность идентифицировать сорта чая даже при наличии незначительных отличий.

Основные виды преобразований, использованные в рамках текущего исследования, включают:

– Поворот (Rotation) – исходные изображения подвергались ротационным преобразованиям, осуществляемым вокруг центральной оси на углы 30° , 45° , 60° , 90° , 120° и 180° . Данная операция создает дополнительное многообразие углов обзора, позволяя модели лучше понимать объект независимо от его пространственного положения.

– Усиление резкости (Sharpening) – целью данного этапа являлось акцентирование мелких деталей изображений. Алгоритм усиливает границы между объектами и фоновыми элементами, помогая машине точнее выделять характерные признаки каждого класса.

– Искажение перспективы (Perspective Warping) – эта процедура симулирует эффекты искажений, вызванные изменением угла съемки или положением камеры. Перспективные деформации создают виртуальную иллюзию глубины пространства, формируя реалистичные объекты в разных ракурсах и снижая зависимость модели от конкретной точки наблюдения.

– Размытие (Blurring) – эффект уменьшения резкости служит средством введения небольших ошибок восприятия, характерных для реальных фотоаппаратов или смартфонов. Применение данного типа преобразования имитирует ситуацию плохого освещения или быстрого движения объектов, обеспечивая дополнительную устойчивость системы перед подобными условиями эксплуатации.

Проведя аугментацию изображений вдобавок к изначальному набору изображений чая, было необходимо подсчитать количество новых образцов и высчитать процент увеличения по сравнению с показателями до аугментации. Аугментация увеличила количество изображений: тренировочного набора с 840 до 7560, валидационного набора с 280 до 2520, тестового набора с 280 до 2520. Суммарно датасет увеличился с 1400 до 12600 изображений.

На Рисунке 1 представлены примеры аугментированных изображений.



Рисунок 1 – Аугментация данных
Figure 1 – Data augmentation

Условия эксперимента

Эксперимент по обучению моделей YOLOv8 проводился в строго контролируемых лабораторных условиях, исключавших любые посторонние факторы, которые могли бы повлиять на результаты. Использовались 3 компьютера, оснащенные следующими комплектующими:

- Компьютер 1 – процессор AMD Ryzen 5 3600 3.60 GHz, оперативная память 32 ГБ DDR4, видеокарта Geforce GTX 1660 Super.
- Компьютер 2 – процессор AMD Ryzen 5 2600 3.40 GHz, оперативная память 16 ГБ DDR4, видеокарта Geforce GTX 1650 Super.
- Компьютер 3 – процессор AMD Ryzen 5 5600X 3.70 GHz, оперативная память 32 ГБ DDR4, видеокарта Geforce RTX 3060 Ti.

Операционная система Windows 10 Home x64 была установлена на каждый компьютер исключительно ради чистоты эксперимента. Были минимизированы стандартные программы и приложения, оставив лишь базовый набор драйверов видеокарт и необходимых библиотек Python.

Перед началом экспериментов операционная система была очищена от сторонних приложений и фоновых служб, кроме основных компонентов операционной среды. Установлен минимальный пакет программного обеспечения для поддержки YOLOv8, включавший TensorFlow, PyTorch, CUDA Toolkit версии 11.4 и cuDNN v8.2. Модули или библиотеки, помимо тех, что необходимы для запуска и обучения YOLOv8, установлены не было. Вся вычислительная нагрузка распределялась исключительно на графические процессоры. Зафиксировано отсутствие любых посторонних фоновых процессов и сервисов, способных снизить производительность.

Метрики

Для проведения эксперимента использовались следующие метрики:

- mAP50 – среднее значение средней точности (Average Precision) при пороге пересечения границ областей Intersection over Union (IoU), равном 50 %. Эта метрика

отражает долю правильно распознанных объектов при достаточно высокой степени совпадения прямоугольников локализации с истинными рамками.

– mAP50–95 – усредненное значение средней точности при диапазоне порогов IoU от 50 % до 95 %, равномерно распределенных с шагом 5 %. Данная метрика учитывает точность определения объектов при различных уровнях наложения областей, предоставляя комплексную оценку общей работоспособности модели.

- Train Loss и Val Loss – потери на обучении и валидации соответственно.
- Время обучения – время, затраченное на обучение модели.

Результаты и обсуждение

В Таблице 1 отображены результаты эксперимента. Каждая модель сравнивалась на одинаковом количестве эпох в соответствии с метриками, представленными в предыдущем разделе. Сравнение было проведено на базе 10 эпох с последующим увеличением на 10 до 60 эпох.

Таблица 1 – Результаты сравнения моделей YOLOv8
Table 1 – YOLOv8 models results comparison

Эпохи	mAP50 (nano)	mAP50 (small)	mAP50–95 (nano)	mAP50–95 (small)	Train Loss (nano)	Train Loss (small)	Val Loss (nano)	Val Loss (small)	Время (nano), мин	Время (small), мин
10	0,983	0,982	0,914	0,914	1,68	1,62	1,61	1,58	1153	2174
20	0,991	0,986	0,925	0,939	1,55	1,50	1,49	1,46	2293	4606
30	0,993	0,991	0,934	0,943	1,48	1,43	1,44	1,41	3408	6439
40	0,994	0,993	0,942	0,945	1,42	1,39	1,40	1,38	4530	8548
50	0,994	0,994	0,944	0,947	1,38	1,34	1,37	1,33	5664	10438
60	0,994	0,995	0,946	0,948	1,35	1,31	1,35	1,30	6751	12806

Начиная с 20 эпох модель Small демонстрирует преимущество по показателям mAP50 и mAP50–95. На 60 эпохах mAP50 для Small составляет 0,995, что на 0,001 выше, чем у Nano (0,994). Похожая ситуация наблюдается для mAP50–95, где Small также выигрывает с результатом 0,948 против 0,946 для Nano.

Модель Small показывает значения потерь ниже как на обучении, так и на валидации на протяжении всего обучения. Это подтверждается результатами на 60 эпохах, где для Small Train Loss равен 1,31, а для Nano – 1,35. Валидационные потери также ниже у Small: 1,30 против 1,35 у Nano.

Модель Nano требует значительно меньше времени на обучение по сравнению с моделью Small. На 60 эпохах время обучения для Nano составляет 6751 минут, что почти в два раза меньше, чем для Small – 12806 минут.

Лучшими показателями обладает версия Small, обученная на 60 эпохах. Графики, отображающие ее показатели, представлены на Рисунке 2.

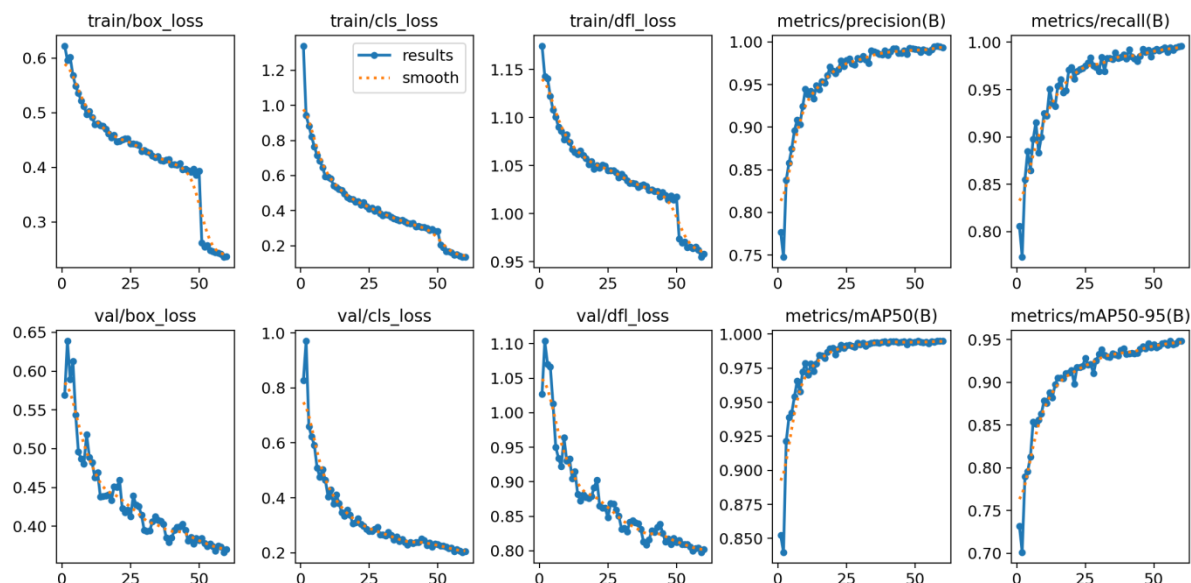


Рисунок 2 – Показатели лучшей модели
Figure 2 – Performance metrics of the best model

Заключение

Анализ показал, что выбор между YOLOv8n и YOLOv8s зависит от целей и задачи. Модель YOLOv8s (Small) превосходит YOLOv8n (Nano) по точности, однако требует больше вычислительных ресурсов и времени на обучение. В случае ограничений по времени или вычислительным мощностям, модель YOLOv8n (Nano) будет предпочтительней, предлагая более быстрые результаты при приемлемом уровне точности. Если задача требует максимальной точности, то согласно результатам исследования необходимо выбирать YOLOv8s (Small) на 50–60 эпохах обучения.

Если же небольшое снижение точности не критично, и важны компактность, быстрая скорость обучения (разница со Small практически два раза) и возможность работать в условиях ограниченных ресурсов, то рекомендуется использовать YOLOv8n (Nano), обученную на 40–50 эпохах.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Preedy V.R. *Tea in Health and Disease Prevention*. Academic Press; 2012. 1612 p.
2. Zhen Yo.-S. *Tea: Bioactivity and Therapeutic Potential*. London: CRC Press; 2002. 280 p. <https://doi.org/10.1201/b12659>
3. Ning J., Sun J., Li Sh., Sheng M., Zhang Zh. Classification of Five Chinese Tea Categories with Different Fermentation Degrees Using Visible and Near-Infrared Hyperspectral Imaging. *International Journal of Food Properties*. 2017;20(2):1515–1522. <https://doi.org/10.1080/10942912.2016.1233115>
4. Bakhshipour A., Zareiforoush H., Bagheri I. Application of Decision Trees and Fuzzy Inference System for Quality Classification and Modeling of Black and Green Tea Based on Visual Features. *Journal of Food Measurement and Characterization*. 2020;14(3):1402–1416. <https://doi.org/10.1007/s11694-020-00390-8>
5. Mukhopadhyay S., Paul M., Pal R., De D. Tea Leaf Disease Detection Using Multi-Objective Image Segmentation. *Multimedia Tools and Applications*. 2021;80(1):753–771. <https://doi.org/10.1007/s11042-020-09567-1>

6. Chen Yi. Identification of Tea Leaf Based on Histogram Equalization, Gray-Level Co-Occurrence Matrix and Support Vector Machine Algorithm. In: *Multimedia Technology and Enhanced Learning: Second EAI International Conference, ICMTEL 2020: Proceedings: Part I, 10–11 April 2020, Leicester, UK*. Cham: Springer; 2020. P. 3–16. https://doi.org/10.1007/978-3-030-51100-5_1
7. Takahashi K., Sugimoto I. Remarks on Tea Leaves Aroma Recognition Using Deep Neural Network. In: *Engineering Applications of Neural Networks: 18th International Conference, EANN 2017: Proceedings, 25–27 August 2017, Athens, Greece*. Cham: Springer; 2017. P. 160–167. https://doi.org/10.1007/978-3-319-65172-9_14
8. Lakshmanan V., Görner M., Gillard R. *Practical Machine Learning for Computer Vision*. Sebastopol: O'Reilly Media, Inc.; 2021. 482 p.
9. Burns D.A., Ciurczak E.W. *Handbook of Near-Infrared Analysis*. Boca Raton: CRC Press; 2007. 834 p. <https://doi.org/10.1201/9781420007374>
10. Shanmugamani R. *Deep Learning for Computer Vision: Expert Techniques to Train Advanced Neural Networks Using TensorFlow and Keras*. Packt Publishing Ltd; 2018. 310 p.
11. Терещук М.В., Зубков А.В., Орлова Ю.А., Молчанов Д.Р., Литвиненко В.А., Черкашин Д.Р. Классификация движений антропоморфного объекта на видеоизображении с использованием методов машинного обучения. *Вестник Дагестанского государственного технического университета. Технические науки*. 2024;51(2):154–163. <https://doi.org/10.21822/2073-6185-2024-51-2-154-163>
Tereshchuk M.V., Zubkov A.V., Orlova Yu.A., Molchanov D.R., Litvinenko V.A., Cherkashin D.R. Development of Models for Classifying the Movements of an Anthropomorphic Body From a Video Stream. *Herald of Dagestan State Technical University. Technical Sciences*. 2024;51(2):154–163. (In Russ.). <https://doi.org/10.21822/2073-6185-2024-51-2-154-163>
12. Ульев А.Д., Донская А.Р., Зубков А.В. Автоматизированное распознавание и контроль взаимодействия людей по видеоизображению. *Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение*. 2023;13(2):45–64. <https://doi.org/10.21869/2223-1536-2023-13-2-45-64>
Ulyev A.D., Donskaia A.R., Zubkov A.V. Automated Recognition and Control of Human Interaction by Video Image. *Proceedings of the Southwest State University. Series: IT Management, Computer Science, Computer Engineering. Medical Equipment Engineering*. 2023;13(2):45–64. (In Russ.). <https://doi.org/10.21869/2223-1536-2023-13-2-45-64>
13. НИКИТИН Д.В., ТАРАНЕНКО И.С., КАТАЕВ А.В. Детектирование дорожных знаков на основе нейросетевой модели YOLO. *Инженерный вестник Дона*. 2023;(7). URL: <http://www.ivdon.ru/ru/magazine/archive/n7y2023/8531>
Nikitin D.V., Taranenko I.S., Kataev A.V. Road Sign Detection Based on the YOLO Neural Network Model. *Engineering Journal of Don*. 2023;(7). (In Russ.). URL: <http://www.ivdon.ru/en/magazine/archive/n7y2023/8531>
14. Orlova Yu., Gorobtsov A., Sychev O., Rozaliev V., Zubkov A., Donskaia A. Method for Determining the Dominant Type of Human Breathing Using Motion Capture and Machine Learning. *Algorithms*. 2023;16(5). <https://doi.org/10.3390/a16050249>
15. Майорова Е.С., Зарипова Р.С. Разработка алгоритма переноса стиля изображения с использованием предобученной нейросети. *Инженерный вестник Дона*. 2024;(2). URL: <http://www.ivdon.ru/ru/magazine/archive/n2y2024/8997>

Majorova E.S., Zaripova R.S. Development of Image Style Transfer Algorithm Using Pre-Trained Neural Network. *Engineering Journal of Don*. 2024;(2). (In Russ.). URL: <http://www.ivdon.ru/en/magazine/archive/n2y2024/8997>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Попов Владислав Алексеевич, студент, Волгоградский государственный технический университет, Волгоград, Российская Федерация. *e-mail*: vladislav332598@yandex.ru

Vladislav A. Popov, Student, Volgograd State Technical University, Volgograd, the Russian Federation.

Зубков Александр Владимирович, кандидат технических наук, доцент, Волгоградский государственный технический университет, Волгоградский государственный медицинский университет, Волгоград, Российская Федерация. *e-mail*: zubkov.alexander.v@gmail.com
ORCID: [0000-0003-0425-5695](https://orcid.org/0000-0003-0425-5695)

Alexander V. Zubkov, Candidate of Engineering Sciences, Docent, Volgograd State Technical University, Volgograd State Medical University, Volgograd, the Russian Federation.

Статья поступила в редакцию 30.04.2025; одобрена после рецензирования 01.06.2025; принята к публикации 09.06.2025.

The article was submitted 30.04.2025; approved after reviewing 01.06.2025; accepted for publication 09.06.2025.