УДК 004.932
DOI

АВТОМАТИЧЕСКОЕ РАСПОЗНАВАНИЕ ЭЛЕКТРОННЫХ ОБРАЩЕНИЙ ГРАЖДАН В ВИДЕ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ СОВРЕМЕННЫХ ТЕХНОЛОГИЙ

Догадина Е.П., Долгов В.И.

ФГОБУ ВО «Финансовый университет при Правительстве Российской Федерации», Москва, e-mail: epdogadina@fa.ru

Автоматизация обработки обращений граждан, поступающих в электронном виде с приложением изображений, становится все более значимой для эффективной работы государственных и муниципальных организаций. В представленной работе реализован подход, основанный на современных методах компьютерного зрения и машинного обучения, для автоматического выделения и распознавания текстовой информации на сканах и фотографиях документов различного качества. Целью данной работы является разработка системы, которая будет автоматически анализировать и распознавать электронные обращения граждан в виде изображений в различных условиях, используя гибридные технологии – сочетание нейросетевых моделей для детекции объектов в режиме реального времени с инструментами для распознавания текста. В ходе исследования были изучены и сопоставлены архитектуры сверточных нейронных сетей для выделения текстовых областей и алгоритмы распознавания текста, что позволило проанализировать их эффективность при различных условиях, включая размытость, наклон текста и низкое качество изображений. Проведен сравнительный анализ точности и скорости работы моделей, а также выявлены основные преимущества и ограничения существующих решений. Результаты экспериментов показали, что интеграция современных методов обработки изображений и распознавания текста позволяет повысить точность и снизить количество ошибок при автоматической регистрации данных в электронных обращениях по сравнению с традиционными подходами. Применение разработанного подхода на основе нейросетевых моделей с учетом технологии оптического распознавания текста может быть рекомендовано для внедрения в автоматизированные системы документооборота, что будет способствовать ускорению рассмотрения заявлений граждан и улучшению качества предоставляемых государственных услуг.

Ключевые слова: нейросетевые модели, оптическое распознавание символов, детекция текстовых областей, текстовое изображение, электронные обращения, автоматизация обработки обращений, компьютерное зрение, машинное обучение

Статья подготовлена по результатам исследований, выполненных за счет бюджетных средств по государственному заданию Финуниверситета.

AUTOMATIC RECOGNITION OF ELECTRONIC CITIZEN APPEALS PRESENTED AS IMAGES USING MODERN TECHNOLOGIES

Dogadina E.P., Dolgov V.I.

Financial University under the Government of the Russian Federation, Moscow, e-mail: epdogadina@fa.ru

The automation of processing citizen appeals submitted electronically with attached images is becoming increasingly important for the effective operation of governmental and municipal organizations. This paper presents an approach based on modern computer vision methods and machine learning for the automatic detection and recognition of textual information in scans and photographs of documents of varying quality. The goal of this work is to develop a system that will automatically analyze and recognize electronic appeals from citizens in the form of images in various conditions using hybrid technologies – a combination of neural network models for real-time object detection with text recognition tools. The study examines and compares the architectures of convolutional neural networks for text area detection and algorithms for text recognition, enabling an assessment of their effectiveness under various conditions, including image blur, text skew, and low image quality. A comparative analysis of model accuracy and processing speed was carried out, and the main advantages and limitations of existing solutions were identified. Experimental results have shown that the integration of advanced image processing methods and text recognition algorithms increases accuracy and reduces the number of errors in the automatic registration of data in electronic appeals compared to traditional approaches. The application of the developed approach based on neural network models, taking into account optical character recognition technology, can be recommended for implementation in automated document management systems, which will help to speed up the processing of citizens' applications and improve the quality of public services provided.

Keywords: neural network models, optical character recognition, detection of text areas, textual images, electronic citizen appeals, automation of appeal processing, computer vision, machine learning

The article was prepared based on the results of research carried out at the expense of budgetary funds under a state assignment to the Financial University.

Введение

В условиях стремительного развития цифровых технологий и внедрения принципов «умного управления» все большее количество гражданских обращений поступает в электронном виде. Электронные обращения становятся важным инструментом взаимодействия между населением и органами государственной и муниципальной власти. Однако наряду с текстовыми сообщениями все чаще используются изображения, содержащие как рукописный, так и печатный текст – фотографии заявлений, сканы документов, скриншоты, графические схемы и другие материалы.

Обработка таких обращений вручную требует значительных временных и трудовых ресурсов, повышает риск ошибок и задержек при регистрации и направлении запроса в соответствующие структурные подразделения. Это обосновывает необходимость автоматизации процессов распознавания, классификации и маршрутизации обращений, представленных в виде изображений [1].

Современные технологии компьютерного зрения, методы оптического распознавания текста (optical character recognition, OCR), а также методы машинного и глубокого обучения предоставляют широкие возможности для решения этих задач. Применение нейросетевых моделей позволяет эффективно извлекать текстовую информацию из изображений различного качества, классифицировать обращения по тематическим категориям, выявлять ключевые фрагменты текста и автоматически направлять заявки в соответствующие отделы для оперативного рассмотрения.

Целью настоящей работы является исследование современных методов классификации и автоматического распознавания текстовой информации, содержащейся в электронных обращениях граждан в виде изображений, с использованием технологий компьютерного зрения и машинного обучения. Особое внимание уделяется оценке эффективности различных подходов для повышения точности и оперативности обработки обращений в условиях цифровой трансформации государственных услуг.

Анализ современных подходов

В последние годы направление, связанное с обработкой изображений и распознаванием текста из сканов и фотографий документов, развивается особенно динамично [2]. Благодаря развитию методов компьютерного зрения и машинного обучения появилась возможность извлекать данные из неструктурированных источников и автоматизировать процессы, ранее полностью зависящие от ручного труда.

В ряде российских исследований показано, что применение алгоритмов глубокого обучения позволяет ускорить обработку документов. Так, созданный модуль интеллектуального распознавания используется для автоматизации приема документов абитуриентов; его внедрение привело к снижению числа ошибок при извлечении данных из паспортов и дипломов и уменьшению времени проверки [3]. В другой работе представлена система для анализа текстов контрактов и прогнозирования их исполнения. В этом случае интеграция методов машинного обучения повысила точность оценки рисков и сократила сроки принятия решений [4]. Все это подтверждает, что интеллектуальные технологии эффективно решают задачи электронного документооборота.

Классические OCR-системы, в частности Tesseract, обычно применяются при распознавании сканов с хорошо читаемым печатным текстом. Однако известно, что при работе со сложными изображениями уровень точности таких систем уменьшается [5]. В новых разработках, к примеру EasyOCR, объединяются алгоритмы для выделения текстовых областей и нейросетевые методы распознавания. Такой подход дает заметное преимущество при работе с изображениями, содержащими сложную структуру [6]. Модели, обученные на значительных объемах данных, показывают стабильные результаты даже при низком качестве исходных изображений [7].

Детекция текстовых областей считается обязательным этапом для современных систем распознавания. Алгоритмы, построенные на базе глубоких сверточных сетей, в том числе YOLO, уверенно выделяют текстовые фрагменты даже при наличии поворота, перекрытий или сложного фона [8]. Среди сегментационных методов CRAFT позволяет определять отдельные символы и связи между ними, что помогает корректно восстанавливать структуру текста [9]. В архитектурах, основанных на трансформерах, например TrOCR, используются предобученные представления и учитывается языковой контекст, что приводит к повышению точности распознавания [10].

Этап предобработки изображений оказывает значительное влияние на итоговую точность OCR. Здесь используются приемы очистки фона, устранения шумов, коррекции контраста и геометрии. При применении методов увеличения разрешения, например EDSR, удается повысить читаемость размытых фрагментов текста [11]. Для повышения устойчивости моделей при анализе сложных изображений находят применение процедуры, снижающие не-

определенность сети, а также анализ в пониженных разрешениях [12].

Распознавание рукописного текста попрежнему относится к наиболее трудоемким направлениям. Современные решения, в которых комбинируются сверточные и рекуррентные нейросети с механизмами внимания, приблизились по точности к человеческим показателям, что подтверждается результатами специализированных исследований [13]. Существенный прирост качества достигается за счет обучения моделей на синтетических данных. Генерация искусственных изображений с разнообразным фоном расширяет обучающий корпус и улучшает итоговые результаты распознавания [14].

В ряде зарубежных работ акцент делается на создании комплексных систем, объединяющих распознавание текста и интеллектуальную обработку данных. Так, система LMV-RPA реализует голосование между несколькими алгоритмами распознавания и языковыми моделями, что обеспечивает высокий уровень точности и уменьшает время анализа [15]. В решении PreP-OCR совмещается восстановление качества изображения с последующей языковой коррекцией, что ведет к заметному снижению количества ошибок [16]. Бенчмарк OCRBench используется для оценки эффективности мультимодальных моделей при работе с различными типами текстов и выявляет существующие трудности при анализе документов со сложной структурой [17]. Модель OCR-free Document Understanding Transformer сочетает в себе обработку изображения и текста, что обеспечивает хорошие результаты в различных сценариях применения [18]. В обзоре [19] представлены современные подходы к пониманию структуры документов на базе трансформеров, а также выделена значимость этих решений для повышения качества обработки.

Сравнительный анализ подходов подтверждает, что современные OCR-системы постепенно переходят от самостоятельных алгоритмов к интегрированным решениям. В таких системах используются нейросетевые и мультимодальные методы, реализуется предобработка изображений и учитывается контекст. Эти технологии внедряются в государственные информационные системы, что способствует повышению точности извлечения информации, ускоряет обработку обращений и улучшает коммуникацию между гражданами и органами власти.

Материалы и методы исследования

В работе представлены ключевые этапы и методы распознавания документов. Первым шагом является подготовка данных, которая включает сбор информации, формирование размеченных выборок (определение и обозначение важных элементов документов) и начальная обработка изображений. Для выполнения задачи детекции текстовых блоков на изображениях были использованы сверточные нейронные сети (CNN).

Среди моделей нейронных сетей сверточные нейронные сети занимают особое место благодаря использованию сверточных слоев. Эти слои представляют собой набор матриц, результат работы предыдущего слоя, умноженных на фильтры, веса которых определяются в процессе обучения модели. Количество матриц и фильтров зависит от решаемой задачи, и при увеличении количества фильтров повышается предсказательная способность модели, но требуется большая вычислительная мощность [3, 4]. На рис. 1 представлен алгоритм работы сверточной нейросетевой архитектуры.

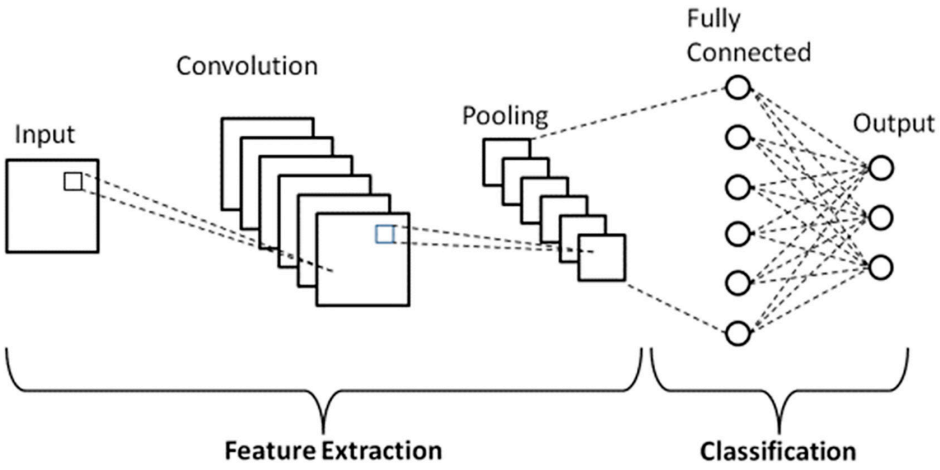

Рис. 1. Алгоритм работы сверточной нейросети (CNN)
Источник: составлено авторами на основе [1, 4, 5]

Таблица 1

Сравнительный анализ YOLOv11 с другими моделями

Характеристика	YOLOv11	YOLOv8	YOLOv9	YOLOv10	Faster R-CNN
Архитектура	Transformer-based backbon	CNN-based	CNN-based	CNN-based	Region Proposal Network (RPN)
Скорость работы, мс	13,5	23	21	19,3	17
Точность детекции	Высокая	Средняя	Средняя	Средняя	Высокая
Поддержка задач	Обнаружение, сегментация, классификация	Обнаружение, сегментация, классификация	Обнаружение, сегментация, классификация	Обнаружение, сегментация, классификация	Обнаружение
Вычислительные затраты	Средние	Средние	Средние	Средние	Высокие
Уровень ложных детекций	Низкий	Средний	Средний	Средний	Низкий

Источник: составлено авторами на основе [5, 20, 21].

В работе использовалась гибридная CNN YOLOv11, которая работает по принципу разделения изображения на сетку и предсказания ограничивающих рамок для найденных объектов. В сравнении с ранними версиями, эта модель использует улучшенные блоки извлечения признаков и оптимизированные механизмы обработки данных, повышающие эффективность в реальном времени. Благодаря высокой точности и быстродействию YOLOv11 активно применяется в видеонаблюдении, биометрических системах и автоматическом анализе изображений. В табл. 1 представлено сравнение YOLOv11 с другими CNN-моделями для детекции изображений.

Архитектура YOLOv11 была улучшена за счет использования трансформаторного бэкбона, что положительно сказалось на производительности и точности. Faster R-CNN обеспечивает высокую точность, но проигрывает в скорости по сравнению с YOLOv11. YOLOv8, YOLOv9 и YOLOv10 имеют более высокое время инференса, чем YOLOv11, и уступают ему по точности.

Для электронной подачи документов гражданами особенно важно, чтобы система OCR могла точно распознавать текст с различных типов изображений, будь то отсканированные копии, фотографии паспортов или других удостоверяющих документов. Рассмотрим наиболее подходящие OCR-методы: Neural OCR, EasyOCR и Tesseract OCR.

Современные нейросетевые решения на основе глубокого обучения (Neural OCR) показывают наилучшие результаты при обработке документов, особенно если используются GPU для ускорения вычислений. Такие системы способны эффективно обрабатывать размытые или низкокачественные изображения, а также наклоненный или нестандартный текст.

Однако, если требуется open-source решение с хорошим балансом между точностью и скоростью, EasyOCR считается оптимальным выбором. Tesseract OCR может подойти для простых задач, например, при работе с четкими отсканированными документами в формате PDF или PNG. Однако он требует тщательной предварительной обработки изображений и не всегда обеспечивает высокую точность на мобильных фото или документах с фоном [4, 5].

Таблица 2

Ключевые факторы при работе с OCR-методами

Фактор	Рекомендация
Точность	Neural OCR > EasyOCR > Tesseract
Поддержка языков	Neural OCR, EasyOCR
Открытость и стоимость	Tesseract, EasyOCR

Источник: составлено авторами на основе [5, 10, 12].

Ключевые факторы при работе с OCR-методами представлены в табл. 2.

Для электронной подачи документов гражданами лучше всего использовать Neural OCR, особенно если есть возможность задействовать GPU и работать с облачными сервисами. Если нужен open-source вариант, можно выбрать EasyOCR.

Результаты исследования и их обсуждение

В ходе работы были рассмотрены и реализованы современные алгоритмы распознавания изображений для автоматической идентификации и классификации различных видов электронных обращений граждан. Использование гибридных нейросетевых моделей, таких как YOLOv11, позволило эффективно выделять текстовые области на изображениях разного качества, что существенно повысило точность последующего распознавания текста.

Сравнительный анализ показал, что архитектура YOLOv11, усовершенствованная за счет внедрения трансформаторного бэкбона, обеспечивает высокую производительность и точность по сравнению с более ранними версиями YOLO и моделью Faster R-CNN, особенно по таким параметрам, как скорость обработки и количество ложных срабатываний. Благодаря использованию оптимизированных механизмов обработки данных, YOLOv11 демонстрирует лучшие результаты при работе в режиме реального времени и позволяет надежно применять систему для автоматического анализа обращений.

В работе используется видеокарта NVIDIA GTX 1070 6 Gb RAM. На рис. 2 представлен график метрик точности по нескольким порогам IoU от 0,5 до 0,95 (mAP50-95 – mean Average Precision) и времени обработки кадра (FPS – Frames Per Second) для исследуемых моделей YOLOv11.

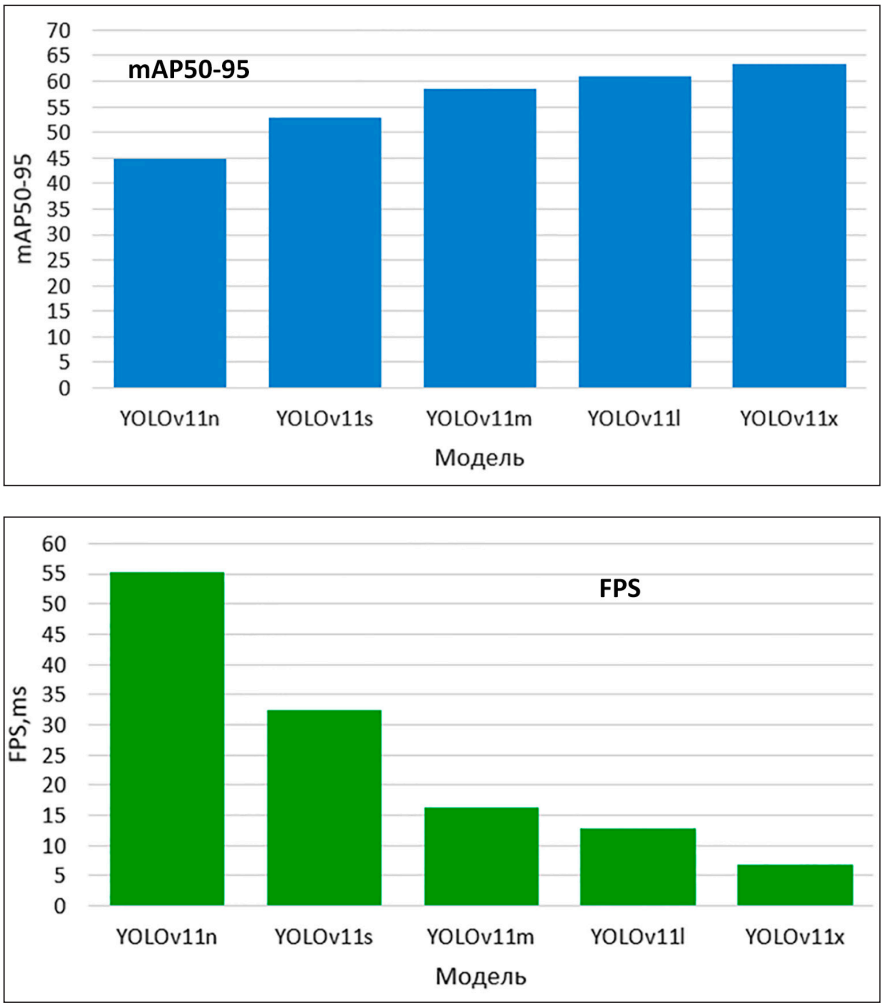

Рис. 2. Сравнение метрик mAP50-95 и FPS для моделей YOLOv11
Источник: составлено авторами на основе полученных данных в ходе исследования

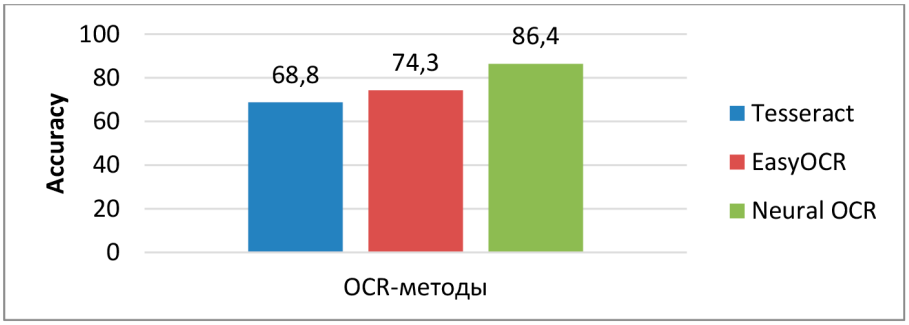

Рис. 3. Задача распознавания текста
Источник: составлено авторами на основе полученных данных в ходе исследования

Метрика mAP50-95 позволяет видеть, как изменяется качество детекции при разных версиях модели. Для каждого порога IoU определяется площадь под кривой точности (Precision-Recall curve). Это значения AP (Average Precision), которые показывают, насколько хорошо модель находит объекты на уровне заданного порога. После расчета AP для каждого порога (например, 0,50; 0,55; 0,60; ...; 0,95) берется среднее значение этих AP. Это дает итоговую метрику mAP50-95. FPS для модели YOLOv11 является важной метрикой, которая определяет, как быстро модель может обрабатывать поток изображения, в том числе в реальном времени.

По данным рис. 2 можно сделать вывод, что лучшее влияние на детекцию объектов обеспечивает модель YOLOv11x. В связи с этим дальнейшие эксперименты проводились для этой модели.

Для распознавания текстовой информации были исследованы и сопоставлены различные методы оптического распознавания символов, в том числе Neural OCR, EasyOCR и Tesseract OCR. Было выявлено, что нейросетевые решения, использующие глубокое обучение, дают наилучшие результаты при обработке документов с размытыми, наклоненными или нестандартными текстами, а также на изображениях с низким качеством. EasyOCR был отмечен как эффективный открытый инструмент для работы с большим количеством языков и типовых задач, а Tesseract OCR рекомендуется использовать для хорошо отсканированных документов, несмотря на некоторые ограничения при обработке фотографий с фоном.

На рис. 3 представлена диаграмма задачи распознавания текста с учетом метрики Accuracy (доли верно распознанных символов).

Внедрение данных алгоритмов и инструментов способствует не только повышению точности и скорости обработки электронных обращений, но и оптимизации работы государственных и муниципальных структур. Использование автоматизированных систем обработки документов позволяет снизить нагрузку на сотрудников, ускорить реагирование на запросы граждан, повысить прозрачность и качество предоставляемых услуг. Вместе с тем важным остается вопрос информационной безопасности и защиты персональных данных при массовой цифровизации документооборота.

Таким образом, полученные результаты подтверждают целесообразность применения современных методов компьютерного зрения и машинного обучения для обработки электронных обращений граждан в виде изображений. Интеграция таких решений в государственные системы способна значительно повысить эффективность и качество предоставления услуг населению.

Заключение

Современные алгоритмы распознавания изображений и методы оптического распознавания текста позволяют эффективно автоматизировать обработку электронных обращений граждан, представленных в виде изображений.

В данной статье была подробно рассмотрена проблема распознавания документов в различных условиях с применением гибридных систем, сочетающих возможности модели YOLOv11 для детекции объектов и эффективные инструменты распознавания текста, такие как Tesseract, Neural OCR и EasyOCR. Результаты проведенного исследования подтверждают, что гибридный подход обеспечивает значительное улучшение скорости обработки документов, позволяя справляться с различными электронными обращениями, поступающими от граждан. Комбинируя возможности YOLOv11 для локализации текстовых обла-

стей с передовыми алгоритмами распознавания текста, в работе достигнута высокая степень достоверности в идентификации информации, содержащейся в документах. Такой подход не только оптимизирует процесс обработки, но и открывает новые возможности для автоматизации работы с документами в различных сферах.

Кроме того, результаты исследований показали, что использование этой модели в практике может значительно сократить временные затраты на ручную обработку данных и минимизировать вероятность ошибок, что является критически важным для обеспечения качества информации. Перспективы дальнейших исследований связаны с расширением функциональных возможностей систем распознавания, повышением их устойчивости к ошибкам и адаптацией к специфическим условиям практического применения.

Авторы заявляют об отсутствии явных или потенциальных конфликтов интересов.

Список литературы

1. Ivanyuk V. Neural Network Model for the Multiple Factor Analysis of Economic Efficiency of an Enterprise // Lecture Notes in Computer Science. 2021. Vol. 12855 LNAI. P. 278–289.; URL: https://link.springer.com/chapter/10.1007/978-3-030-87897-9_26 (дата обращения: 06.08.2025). DOI: 10.1007/978-3-030-87897-9_26. EDN: HFSOBY.

2. Тишина Л.В. Разработка модуля интеллектуального распознавания документов средствами машинного зрения // Интерэкспо Гео-Сибирь. 2022. Т. 7. № 2. С. 136–140. EDN: KSKGTZ.

3. Корчагин С.А., Догадина Е.П., Мелентьев В.В., Никитин П.В., Сердечный Д.В. Система поддержки принятия решений по выдаче банковских гарантий на основе прогнозирования исполнения контрактов с использованием методов машинного обучения и технологий парсинга // Современные наукоемкие технологии. 2023. № 7. С. 41–47. URL: https://top-technologies.ru/ru/article/view?id=39692 (дата обращения: 04.07.2025). DOI: 10.17513/snt.39692.

4. Lozhkin A.G., Mayorov K.N., Bozek P. Convolutional neural networks training for autonomous robotics // Management Systems in Production Engineering. 2021. Vol. 29, Is. 1. P. 75–79. DOI: 10.2478/mspe-2021-0010.

5. Andriyanov N. Methods for preventing visual attacks in convolutional neural networks based on data discard and dimensionality reduction // Applied Sciences. 2021. Vol. 11, Is. 11. Article 5235. URL: https://www.mdpi.com/2076-3417/11/11/5235 (дата обращения: 06.07.2025). DOI: 10.3390/app11115235.

6. Baek Y., Lee B., Han D., Yun S., Lee H. Character region awareness for text detection // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 9365–9374. DOI: 10.1109/CVPR.2019.00959.

7. Shi B., Bai X., Yao C. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2017. Vol. 39, Is. 11. P. 2298–2304. DOI: 10.1109/TPAMI.2016.2646371.

8. Xiao L., Zhou P., Xu K., Zhao X. Multi-directional scene text detection based on improved YOLOv3 // Sensors. 2021. Vol. 21, Is. 14. Art. 4870. DOI: 10.3390/s21144870.

9. Li M., Lv T., Chen J., Cui L., Lu Y., Florencio D., Zhang C., Li Z., Wei F. TrOCR: Transformer-based optical character recognition with pre-trained models // Proceedings of the AAAI Conference on Artificial Intelligence. 2023. Vol. 37, No. 11. P. 13094–13102. DOI: 10.1609/aaai.v37i11.26538.

10. Alahmadi M.D., Alshangiti M. Optimizing OCR performance for programming videos: the role of image super-resolution and large language models // Mathematics. 2024. Vol. 12, Is. 7. Art. 1036. DOI: 10.3390/math12071036.

11. Maliński K., Okarma K. Analysis of image preprocessing and binarization methods for OCR-based detection and classification of electronic integrated circuit labeling // Electronics. 2023. Vol. 12. Art. 2449. DOI: 10.3390/electronics12112449.

12. Мухамадиева К.Б. Обнаружение и распознавание неординарного текста из видеокадров // Вестник Донецкого национального университета. Серия Г: Технические науки. 2021. № 1. С. 45–57. EDN: WCKKZA.

13. AlKendi W., Gechter F., Heyberger L., Guyeux C. Advancements and Challenges in Handwritten Text Recognition: A Comprehensive Survey // Journal of Imaging. 2024. Vol. 10, Is. 1. Art. 18. URL: https://www.mdpi.com/2313-433X/10/1/18 (дата обращения: 04.07.2025). DOI: 10.3390/jimaging10010018.

14. Gupta A., Vedaldi A., Zisserman A. Synthetic Data for Text Localisation in Natural Images // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 2315–2324. URL: https://ieeexplore.ieee.org/document/7780623 (дата обращения: 04.07.2025). DOI: 10.1109/CVPR.2016.254.

15. Abdellatif O., Ayman A., Hamdi A. LMV-RPA: Large Model Voting-Based Robotic Process Automation // In: Abraham A., Alimi A.M., Haqiq A., Sehaba K., Gandhi N. (eds) Advances on Intelligent Computing and Data Science II. ICACIn 2024. Lecture Notes on Data Engineering and Communications Technologies, vol 255. Springer, Cham. 2025. P. 134–144. [Электронный ресурс]. URL: https://link.springer.com/chapter/10.1007/978-3-031-91354-9_11 (дата обращения: 04.07.2025). DOI: 10.1007/978-3-031-91354-9_11.

16. Guan S., Lin M., Xu C., Liu X., Zhao J., Fan J., Xu Q., Greene D. PreP-OCR: A Complete Pipeline for Document Image Restoration and Enhanced OCR Accuracy // Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (ACL). 2025. P. 15413–15425. [Электронный ресурс]. URL: https://aclanthology.org/2025.acl-long.749 (дата обращения: 04.07.2025). DOI: 10.18653/v1/2025.acl-long.749.

17. Liu Y., Li Z., Huang M., Yang B., Yu W., Li C., Yin X., Liu C., Jin L., Bai X. OCRBench: On the Hidden Mystery of OCR in Large Multimodal Models // Science China Information Sciences. 2024. Vol. 67. Art. 220102. [Электронный ресурс]. URL: https://link.springer.com/article/10.1007/s11432-024-4235-6 (дата обращения: 04.07.2025). DOI: 10.1007/s11432-024-4235-6.

18. Kim G., Hong T., Yim M., Nam J., Park J., Yim J., Hwang W., Yun S., Han D., Park S. OCR-free Document Understanding Transformer. In: Avidan S. et al. (eds) Computer Vision – ECCV 2022. Lecture Notes in Computer Science, vol 13688. Springer, Cham. 2022. P. 498–517. [Электронный ресурс]. URL: https://link.springer.com/chapter/10.1007/978-3-031-19815-1_29 (дата обращения: 07.07.2025). DOI: 10.1007/978-3-031-19815-1_29.

19. Abdallah A., Eberharter D., Pfister Z., Jatowt A. A survey of recent approaches to form understanding in scanned documents // Artificial Intelligence Review. 2024. Vol. 57. Art. 342. DOI: 10.1007/s10462-024-11000-0.

20. Никитин Д.В., Тараненко И.С., Катаев А.В. Детектирование дорожных знаков на основе нейросетевой модели YOLO // Инженерный вестник Дона. 2023. № 7 (103). С. 91–99. EDN: MPZLRZ.

21. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // IEEE Trans Pattern Anal Mach Intell. 2017. Vol. 39, Is. 6. P. 1137–1149. URL: https://ieeexplore.ieee.org/document/7485869 (дата обращения: 06.07.2025). DOI: 10.1109/TPAMI.2016.2577031.