

Научная статья

УДК 004.93

DOI: <https://doi.org/10.18127/j20700970-202503-06>

Распознавание документов в различных условиях с помощью методов машинного обучения

Е.П. Догадина¹, У.Ю. Суханова², М.А. Ищенко³, Д.И. Веселов⁴

^{1–4} Финансовый университет при Правительстве Российской Федерации (Москва, Россия)

¹ epdogadina@fa.ru, ² uysukhanova@fa.ru, ³ maishchenko@fa.ru, ⁴ diveselov@fa.ru

Аннотация

Постановка проблемы. Распознавание документов – это процесс автоматического анализа, классификации и извлечения информации из структурированных и неструктурированных материалов. Современные технологии машинного обучения и глубокого обучения позволяют значительно улучшить точность и скорость распознавания документов по сравнению с традиционными подходами. Однако реальные условия применения этих систем часто осложняются множеством факторов, таких как низкое качество сканирования, влияние внешних условий (освещение, ракурс, деформация бумаги), использование специализированных символов. Эти сложности требуют разработки более гибких и адаптивных методов машинного обучения, способных работать в широком спектре условий.

Цель. Создать систему, способную автоматически анализировать и распознавать документы в различных условиях с использованием гибридных систем, а именно: комбинации YOLOv10 с инструментами распознавания текста (такими как Tesseract, KerasOCR и EasyOCR).

Результаты. Показано, что комбинация YOLOv10 extra large с EasyOCR, Tesseract, KerasOCR обеспечивает достаточную точность распознавания на документах в различных условиях. При этом высокие метрики KerasOCR связаны с дополнительным тюнингом модели. Установлено, что использование YOLOv10 в сочетании с современными инструментами распознавания текста позволяет создать универсальную систему для анализа документов в широком спектре условий.

Практическая значимость. Разработанная система позволит снизить объем ручного труда, который требуется для обработки документов, тем самым освобождая сотрудников для выполнения более сложных и ценных задач. Использование машинного обучения и алгоритмов распознавания текстов также позволит снизить вероятность ошибок, связанных с человеческим фактором, увеличить скорость обработки данных и повысить точность и качество данных. Созданная система располагает огромным потенциалом применения в самых разных областях, где необходима быстрая и качественная обработка документов. Она представляет интерес для бухгалтерии и финансов, юридической сферы, сферы здравоохранения и государственных учреждений, а также логистики.

Ключевые слова

Машинное обучение, алгоритмы распознавания, сверточные нейронные сети, поврежденные изображения, метрики

Статья подготовлена по результатам исследований, выполненных за счет бюджетных средств по государственному заданию Финансового университета при Правительстве Российской Федерации.

Для цитирования

Догадина Е.П., Суханова У.Ю., Ищенко М.А., Веселов Д.И. Распознавание документов в различных условиях с помощью методов машинного обучения // Нелинейный мир. 2025. Т. 23. № 3. С. 45–53. DOI: <https://doi.org/10.18127/j20700970-202503-06>

A brief version in English is given at the end of the article

Введение

Успех современных компаний и организаций во многом зависит от способности быстро и точно обрабатывать документы. Распознавание документов с помощью методов машинного обучения становится ключевым инструментом для компаний и организаций, стремящихся автоматизировать процессы обработки данных. Этот подход позволяет эффективно извлекать информацию из различных типов документов (счета, контракты, чеки, формы и т. д.), что значительно снижает затраты времени и ресурсов на ручную обработку данных. Кроме того, автоматическое распознавание документов способствует повышению уровня безопасности компаний и организаций. Этот процесс позволяет не только автоматизировать обработку данных, но и внедрять механизмы защиты от подделок, мошенничества и утечек информации. Распознавание документов в различных условиях с помощью методов машинного обучения – сложная задача, которая требует учета множества факторов, таких как качество изображения, освещение, повреждения документа и т. д.

Ц е л ь р а б о т ы – создать систему, способную автоматически анализировать и распознавать документы в различных условиях с использованием гибридных систем, а именно: комбинации YOLOv10 с инструментами распознавания текста (такими как Tesseract, KerasOCR и EasyOCR).

Материалы и методы

В каждой компании, как российской, так и зарубежной, ежедневно используется большое количество различных финансовых документов. Однако среди этих документов можно выделить ключевые и наиболее распространенные:

счет на оплату (счет-фактура) – коммерческий документ с отметкой времени, в котором подробно описывается и фиксируется транзакция между покупателем и продавцом;

каз на поставку – официальная расписка установленной формы, свидетельствующая о принятии денег, документов, ценностей лицом, выдавшим квитанцию, у лица, которому она выдана [1];

каз на поставку – коммерческий документ, выдаваемый покупателем продавцу, с указанием типов, количества и согласованных цен на необходимые продукты или услуги [2];

каз на продажу – коммерческий документ, подготовленный продавцом и выданный покупателю. Он подтверждает продажу товаров или услуг и детализирует особенности продажи, включая количество, цену и качество предоставляемых товаров или услуг. Он также предоставляет информацию о предполагаемой дате доставки, адресе доставки, условиях оплаты и логистике продажи [3];

выписка из банка – документ, предоставляющий информацию о всех операциях на банковском счете компании, включая поступления и расходы;

акт об оказании услуг – документ, использующийся для фиксации факта оказания услуг или выполнения работ. В нем указываются детали о работах или услугах, их объем, стоимость и другая информация;

договор (контракт) – соглашение между собой двух или более сторон (субъектов), по какому-либо вопросу с целью установления, изменения или прекращения правовых отношений [3].

Безусловно, не все перечисленные документы используются в любой компании без исключения. Так, например, акт об оказании услуг характерен только для российского учета, в то время как остальные документы применимы в том числе для любой международной коммерческой организации.

Таким образом, первичные документы позволяют компании непрерывно вести свою деятельность, получать прибыль, оплачивать налоги, выполнять требования Федеральной налоговой службы РФ, трудового законодательства и других инстанций, определяемых местным законодательством.

Для успешного решения задачи распознавания документов в различных условиях используются различные подходы и технологии машинного обучения (ML) и глубокого обучения (DL) [4, 5].

Рассмотрим *основные методы машинного обучения для распознавания документов*. Существуют классические методы машинного обучения [6]:

SVM (Support Vector Machine) – используется для классификации символов;

Random Forest – применяется для анализа признаков текста;

HOG (Histogram of Oriented Gradients) – позволяет извлекать признаки для детектирования текстовых областей.

Эти методы эффективны для простых задач, но их производительность снижается при работе с большими объемами данных и сложными условиями.

Для распознавания документов в различных условиях хорошо зарекомендовали себя методы глубокого обучения. Рассмотрим некоторые методы, которые показывают высокую эффективность при распознавании документов. Например, CNN (Convolutional Neural Networks) используются для обнаружения текстовых областей и классификации символов. RNN (Recurrent Neural Networks) анализируют последовательности символов для повышения точности распознавания.

Также существуют Transformer-based модели. Например, Vision Transformers (ViT) используется для анализа изображений, а BERT – для семантической обработки текста. Одной из новых технологий в распознавании символов на документе является применение визуально-языковых моделей (VLM) [7].

Однако для распознавания документов в различных условиях наиболее эффективными и точными являются гибридные системы. Они сочетают классические методы и глубокое обучение. Так, CNN может использоваться для предварительной обработки изображения, а затем применяется (например, Tesseract OCR) для распознавания текста.

В работе рассматриваются несколько ключевых этапов и подходов для распознавания документов. Сначала происходит подготовка данных. Она включает в себя сбор данных, создание разметки данных для обучения моделей (определение и маркировку ключевых элементов документов) и предварительную обработку данных. Для распознавания изображений использовались сверточные нейронные сети (CNN) [7–14]. Пример сверточной нейронной сети представлен на рис. 1.

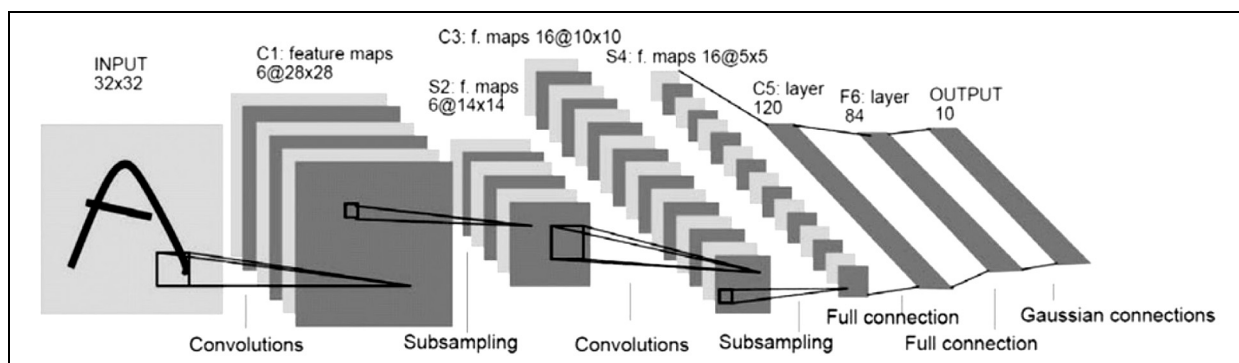


Рис. 1. Сверточная нейронная сеть

Fig. 1. Convolutional Neural Network

Сверточная нейронная сеть состоит из нескольких слоев: входного слоя, сверточного слоя, слоя пула и полностью связанных слоев, как это показано схематично на рис. 2.

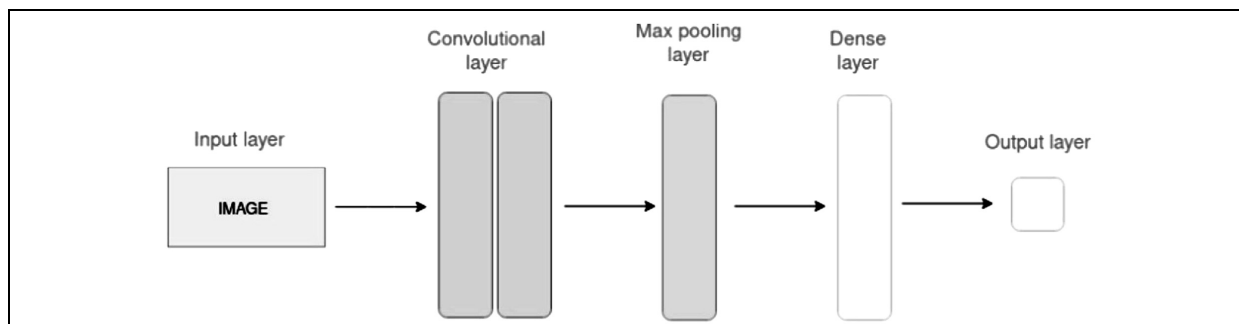


Рис. 2. Простая архитектура CNN

Fig. 2. Simple CNN architecture

Сверточный слой применяет фильтры к входному изображению для извлечения признаков, слой объединения снижает дискретизацию изображения для сокращения вычислений, а полносвязный слой делает окончательный прогноз. Сеть изучает оптимальные фильтры посредством обратного распространения ошибки и градиентного спуска.

Приведем так называемую математику, которая участвует во всем процессе свертки:

слои свертки состоят из набора обучаемых фильтров (или ядер), имеющих небольшие ширину и высоту, а также ту же глубину, что и у входного объема. Например, если нужно запустить свертку изображения размером $34 \times 34 \times 3$, то возможный размер фильтров равен $a \times a \times 3$, где «а» может быть любым (например, 3, 5 или 7), но меньшим по сравнению с размером изображения;

во время прямого прохода шаг за шагом перемещается каждый фильтр по всему входному объему, где каждый шаг называется шагом перемещения (stride), и вычисляем скалярное произведение между весом ядра и патчем из входных данных;

когда происходит перемещение фильтров, получаются двумерные выходные данные для каждого фильтра. Они складываются вместе, и в результате получается выходной объем, имеющий глубину, равную количеству фильтров.

В работе сверточная нейронная сеть YOLO (You Only Look Once) [15] сочетается с инструментами распознавания текста, такими как Tesseract или EasyOCR [16]. Одна модель используется для извлечения блоков текста из данных (урезанное изображение), другая – для извлечения непосредственно текста из этих блоков. Применение комбинации YOLOv10 с инструментами распознавания текста, такими как Tesseract, EasyOCR и KerasOCR, показал наиболее высокие результаты для задач распознавания документов. На выходе YOLO получают координаты ограничивающих рамок (bounding boxes), которые определяют, где на изображении находятся текстовые блоки. После того как текстовые области выделены, они передаются в OCR-инструмент (Tesseract, EasyOCR, KerasOCR) для преобразования изображения таблиц и текста в машиночитаемый формат.

Результаты и их обсуждение

Для обучения моделей был создан свой собственный датасет, включающий в себя около 3000 документов примерно с 30 текстовыми блоками в каждом. Выборка была разделена на тестовую (20%) и обучающую (80%).

Примеры исходных документов с различными качествами изображения и освещения представлены на рис. 3.

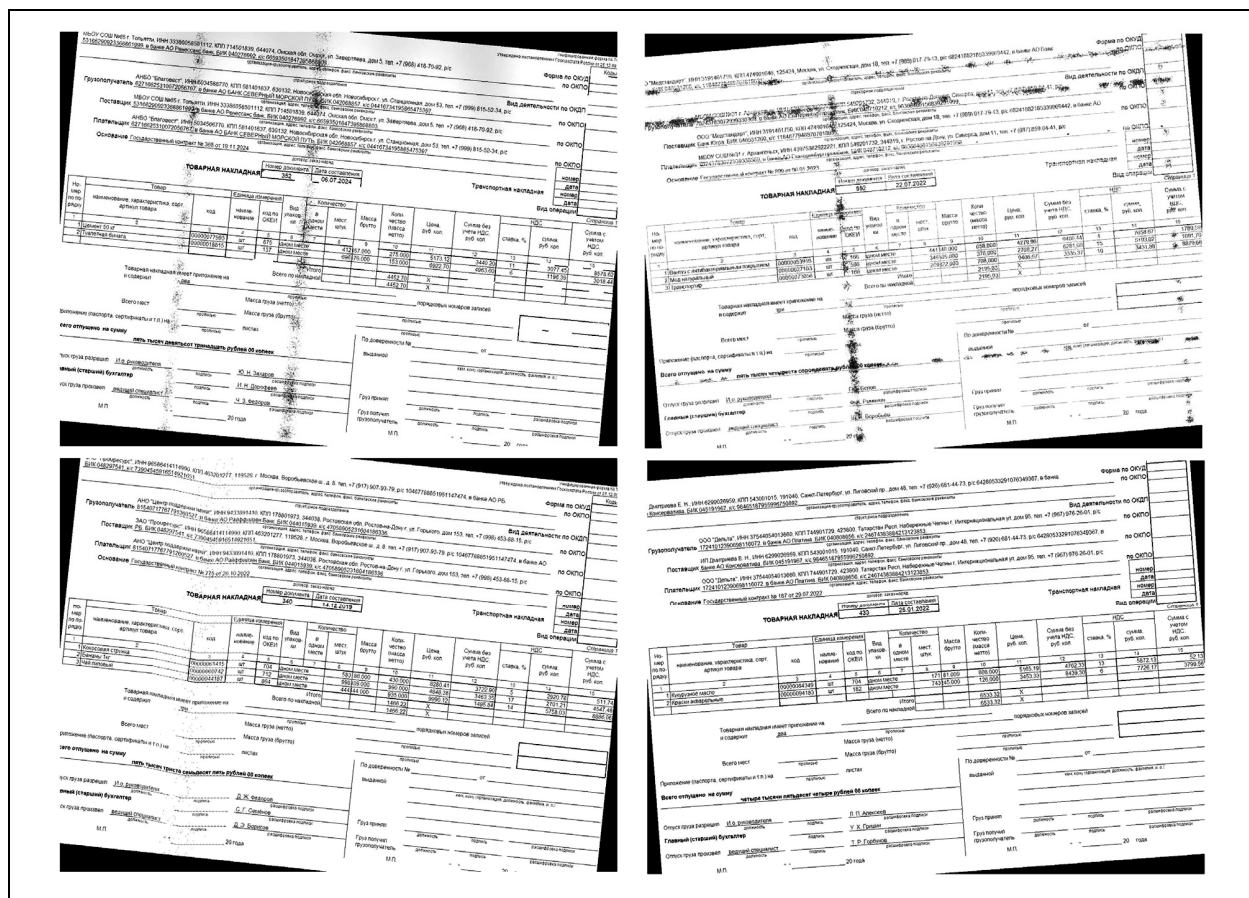


Рис. 3. Исходные документы с различными повреждениями

Fig. 3. Original documents with various damages

Подход с добавлением искажений в изображения связан с необходимостью имитации ситуаций испорченного сканера, принтера, а также криво положенных на сканирование документов и т.п. Более того, обучаясь, в том числе и на таких изображениях, модель будет более универсальной.

Пример реализации программного кода по распознаванию документов с помощью Tesseract и YOLOv10 представлен на рис. 4.

```

1 import cv2
2 import pytesseract
3
4 # Загрузка модели YOLO
5 net = cv2.dnn.readNet("yolov10.weights", "yolov10.cfg")
6 layer_names = net.getLayerNames()
7 output_layers = [layer_names[i[0] - 1] for i in net.getUnconnectedOutLayers()]
8
9 # Загрузка изображения
10 image = cv2.imread("document.jpg")
11 height, width, channels = image.shape
12
13 # Подготовка изображения для YOLO
14 blob = cv2.dnn.blobFromImage(image, 0.00392, (416, 416), (0, 0, 0), True, crop=False)
15 net.setInput(blob)
16 outs = net.forward(output_layers)
17
18 # Обработка результатов YOLO
19 class_ids = []
20 confidences = []
21 boxes = []
22
23 for out in outs:
24     for detection in out:
25         scores = detection[5:]
26         class_id = np.argmax(scores)
27         confidence = scores[class_id]
28         if confidence > 0.5 and class_id == 0: # Предположим, что класс 0 – это текст
29             center_x = int(detection[0] * width)
30             center_y = int(detection[1] * height)
31             w = int(detection[2] * width)
32             h = int(detection[3] * height)
33             x = int(center_x - w / 2)
34             y = int(center_y - h / 2)
35             boxes.append([x, y, w, h])
36             confidences.append(float(confidence))
37             class_ids.append(class_id)
38 # Применение Non-Max Suppression
39 indexes = cv2.dnn.NMSBoxes(boxes, confidences, 0.5, 0.4)
40
41 # Распознавание текста с помощью Tesseract
42 for i in indexes:
43     i = i[0]
44     box = boxes[i]
45     x, y, w, h = box
46     text_region = image[y:y + h, x:x + w]
47
48     # Преобразование в текст
49     text = pytesseract.image_to_string(text_region, lang="rus+eng")
50     print(f"Распознанный текст: {text}")

```

Рис. 4. Распознавание документов с помощью Tesseract и YOLOv10

Fig. 4. Document Recognition with Tesseract and YOLOv10

Обучение проходило со стандартными гиперпараметрами в течение 2000 эпох с показателем ранней остановки при 200 эпохах без улучшений модели.

Соответственно решаем *две задачи*:

- 1) непосредственно обнаружение текстовых блоков;
- 2) распознавание текстов внутри этих блоков.

Введем для первой задачи метрику качества как отношение пересечения к объединению или IoU [15], а для второй задачи – долю верно распознанных символов (Accuracy).

Для обучения и инференса используется видеокарта NVIDIA GTX 1070 6 Gb RAM, что позволяет ускорить процесс обработки данных по сравнению с CPU. В табл. 1 и 2 представлены метрики качества для разных подходов.

Таблица 1. Задача детектирования областей текста

Модель	Средний IoU
YOLOv10x	0,971
YOLOv10l	0,927
YOLOv10m	0,811
YOLOv10n	0,788

Из табл. 1 видно, что лучшую детекцию обеспечивает самая большая модель (YOLOv10 extra large). Поэтому дальнейшие эксперименты (табл. 2) проводились для этой модели. Следует отметить, что некоторые модели требовали дополнительной настройки на русский язык, которая была выполнена с помощью стандартного тюнинга KerasOCR [17].

Таблица 2. Задача распознавания текста

Модель	Accuracy
Tesseract	66,38
EasyOCR	72,81
KerasOCR	85,36

Высокие метрики KerasOCR, в частности, связаны с дополнительным тюнингом модели.

Заключение

Распознавание документов с помощью метода глубокого обучения становится стандартом для компаний, стремящихся к цифровой трансформации. Современные технологии позволяют автоматизировать обработку документов, при этом снизить затраты и повысить точность. Однако успешная реализация таких систем требует внимательного выбора подходящих инструментов, качественных данных для обучения моделей и грамотной интеграции с существующими бизнес-процессами.

На основе полученных результатов программа может эффективно выявлять подозрительные шаблоны в документах, что позволяет предотвращать различные виды мошенничества и несанкционированных действий. Кроме того, система на базе CNN помогает обеспечивать соответствие законодательным и нормативным требованиям, что снижает риски юридических последствий и штрафов. Использование такой системы для управления доступом к документам и защиты конфиденциальных данных минимизирует риски утечки информации и помогает поддерживать уровень ее безопасности.

Данная система может быть интегрирована с другими системами управления документами и корпоративными ресурсами, что повышает общую эффективность работы и упрощает доступ к информации.

Список источников

1. Райзберг Б.А., Лозовский Л.Ш., Стародубцева Е.Б. Современный экономический словарь. Изд. 2-е., испр. М.: ИНФРА-М. 479 с. 1999.
2. Dobler D.W., Burt D.N. Purchasing and Supply Management, Text and Cases (Sixth ed.). Singapore: McGraw-Hill. 1996. P. 70.
3. Сироткин С.А., Кельчевская Н.Р. Бухгалтерский учет и анализ. М.: ИНФРА-М. 2019. С. 224. 355 с.
4. Тишина Л.В. Разработка модуля интеллектуального распознавания документов средствами машинного зрения // Интерэкспо Гео-Сибирь. 2022. V. 7. № 2. P. 136–140.
5. Андриянов Н.А., Андриянов Д.А. О важности аугментации данных при машинном обучении в задачах обработки изображений в условиях дефицита данных // Информационные технологии и нанотехнологии (ИТНТ-2020): Сб.

- трудов по материалам VI Междунар. конф. и молодежной школы. В 4-х томах. Самара, 26–29 мая 2020 года / Под ред. В.В. Мясникова. Том 2. Самара: Самарский национальный исследовательский университет им. акад. С.П. Королева. 2020. С. 383–388.
6. Андриянов Н.А., Никитин П.В. Построение и оценка моделей машинного обучения: Учеб. пособие по дисциплине «Построение и оценка моделей машинного обучения» для студентов, обучающихся по направлению «Прикладная математика и информатика» всех профилей (программы подготовки магистров) // Финуниверситет, Департамент анализа данных и машинного обучения Факультета информационных технологий и анализа больших данных. М.: Финансовый университет. 2023. 1 файл (6,31 Мб).
 7. Zhao X., Xu M., Silamu W., Li Y. CLIP-Llama: A New Approach for Scene Text Recognition with a Pre-Trained Vision-Language Model and a Pre-Trained Language Model. *Sensors* 2024. V. 24. P. 7371. <https://doi.org/10.3390/s24227371>.
 8. Nikolaev K., Malafeev A. Russian Q&A method study: From Naive Bayes to convolutional neural networks. *Lecture Notes in Computer Science*. 2018. V. 11179 LNCS. P. 121–126. DOI 10.1007/978-3-030-11027-7_12. EDN XWJVL7.
 9. Enweji M.Z., Lehinevych T., Glybovets A. Cross-language text classification with convolutional neural networks from scratch. *Eureka: Physics and Engineering*. 2017. № 2. P. 24–33. DOI 10.21303/2461-4262.2017.00304. EDN YPOYJD.
 10. Gallego N. P. D., Ilao J., Cordel M. Blind first-order perspective distortion correction using parallel convolutional neural networks. *Sensors*. 2020. V. 20. № 17. P. 1–20. DOI 10.3390/s20174898. EDN BEKSSY.
 11. Ko K., Jang I., Choi J.H. et al. Stochastic decision fusion of convolutional neural networks for tomato ripeness detection in agricultural sorting systems. *Sensors*. 2021. V. 21. № 3. P. 1–14. DOI 10.3390/s21030917. EDN DHJZNO.
 12. Lozhkin A.G., Maiorov K.N., Bozek P. Convolutional neural networks training for autonomous robotics. *Management Systems in Production Engineering*. 2020. V. 29. № 1. P. 75–79. DOI 10.2478/mspe-2021-0010. EDN TYCRQO.
 13. Andriyanov N. Methods for preventing visual attacks in convolutional neural networks based on data discard and dimensionality reduction. *Applied Sciences (Switzerland)*. 2021. V. 11. № 11. DOI 10.3390/app11115235. EDN PAAONS.
 14. Bratchenko I.A., Bratchenko L.A., Khristoforova Y.A. et al. Classification of skin cancer using convolutional neural networks analysis of Raman spectra. *Computer Methods and Programs in Biomedicine*. 2022. V. 219. P. 106755. DOI 10.1016/j.cmpb.2022.106755. EDN BCGDFW.
 15. Андриянов Н. А., Дементьев В. Е., Ташлинский А. Г. Обнаружение объектов на изображении: от критериев Байеса и Неймана – Пирсона к детекторам на базе нейронных сетей // Компьютерная оптика. 2022. Т. 46. № 1. С. 139–159. DOI 10.18287/2412-6179-CO-922.
 16. Ayachi R., Afif M., Said Ya. et al. Integrating Recurrent Neural Networks with Convolutional Neural Networks for Enhanced Traffic Light Detection and Tracking. *Traitement du Signal*. 2023. V. 40. № 6. P. 2577–2586. DOI 10.18280/ts.400620. EDN AWPIQZ.
 17. Fine-tuning OCR. URL: https://keras-ocr.readthedocs.io/en/latest/examples/fine_tuning_recognizer.html, дата обращения: 19.03.2025.

Информация об авторах

Елена Петровна Догадина – к.т.н., доцент, доцент кафедры информационных технологий
SPIN-код: 1891-3706

Ульяна Юрьевна Суханова – стажер-исследователь
SPIN-код: 1891-3706

Маргарита Алексеевна Ищенко – стажер-исследователь
SPIN-код: 1891-3706

Дмитрий Иванович Веселов – мл. науч. сотрудник
SPIN-код: не представлен

Статья поступила в редакцию 04.06.2025

Одобрена после рецензирования 16.06.2025

Принята к публикации 30.06.2025

Original article

Document recognition in various conditions using machine learning methods

E.P. Dogadina¹, U.Yu. Sukhanova², M.A. Ishchenko³, D.I. Veselov⁴^{1–4} Financial University under the Government of the Russian Federation (Moscow, Russia)¹ epdogadina@fa.ru, ² uysukhanova@fa.ru, ³ maishchenko@fa.ru, ⁴ diveselov@fa.ru

Abstract

Document recognition is the process of automatically analyzing, classifying, and extracting information from structured and unstructured materials. Modern machine learning and deep learning technologies can significantly improve the accuracy and speed of document recognition compared to traditional approaches. However, real conditions of application of such systems are often complicated by many factors: low scanning quality, influence of external conditions (lighting, angle, paper deformation), use of specialized symbols. These difficulties require development of more flexible and adaptive methods of machine learning, capable of working in a wide range of conditions.

Target – the purpose of this work is to create a system capable of automatically analyzing and recognizing documents in various conditions using hybrid systems, namely a combination of YOLOv10 with text recognition tools such as Tesseract, KerasOCR and EasyOCR.

The results showed that the combination of YOLOv10 extra large with EasyOCR, Tesseract, KerasOCR provides sufficient recognition accuracy on documents in various conditions. At the same time, high KerasOCR metrics are associated with additional tuning of the model. Thus, the use of YOLOv10 in combination with modern text recognition tools allows you to create a universal system for analyzing documents in a wide range of conditions.

The developed system will reduce the amount of manual labor required to process documents, thereby freeing up employees to perform more complex and valuable tasks. In addition, the use of machine learning and text recognition algorithms will reduce the likelihood of errors associated with the human factor, increase the speed of data processing and improve the accuracy and quality of data. This system has a huge potential for application in a variety of areas where fast and high-quality document processing is required. It is of interest to accounting and finance, the legal sphere, healthcare and government agencies, and logistics.

Keywords

Machine learning, recognition algorithms, convolutional neural networks, corrupted images, metrics

For citation

Dogadina E.P., Sukhanova U.Yu., Ishchenko M.A., Veselov D.I. Document recognition in various conditions using machine learning methods. *Nonlinear World*. 2025. V. 23. № 3. P. 45–53. DOI: <https://doi.org/10.18127/j20700970-202503-06> (In Russian)

REFERENCES

1. Rajzberg B.A., Lozovskij L.Sh., Starodubceva E.B. *Sovremennyy ekonomicheskij slovar'*. Izd. 2-e., ispr. M.: INFRA-M. 479 s. 1999 (In Russian).
2. Dabler D.W., Burt D.N. *Purchasing and Supply Management, Text and Cases* (Sixth ed.). Singapore: McGraw-Hill. 1996. P. 70.
3. Sirotkin S.A., Kel'chevskaya N.R. *Buhgalterskij uchët i analiz*. M.: INFRA-M. 2019. S. 224. 355 s. (In Russian).
4. Tishina L.V. Razrabotka modulya intellektual'nogo raspoznavaniya dokumentov sredstvami mashinnogo zreniya. *Interekspo Geo-Sibir'*. 2022. V. 7. № 2. P. 136–140 (In Russian).
5. Andriyanov N.A., Andriyanov D.A. O vazhnosti augmentacii dannyh pri mashinnom obuchenii v zadachah obrabotki izobrazhenij v usloviyah deficita dannyh. *Informacionnye tekhnologii i nanotekhnologii (ITNT-2020)*: Sb. trudov po materialam VI Mezhdunar. konf. i molodezhnoj shkoly. V 4-h tomah. Samara, 26–29 maya 2020 goda. Pod red. V.V. Myasnikova. Tom 2. Samara: Samarskij nacional'nyj issledovatel'skij universitet im. akad. S.P. Koroleva. 2020. S. 383–388 (In Russian).
6. Andriyanov N.A., Nikitin P.V. Postroenie i ocenka modelej mashinnogo obucheniya: Ucheb. posobie po discipline «Postroenie i ocenka modelej mashinnogo obucheniya» dlya studentov, obuchayushchihysya po napravleniyu «Prikladnaya matematika i informatika» vseh profilej (programmy podgotovki magistrrov). Finuniversitet, Departament analiza dannyh i mashinnogo obucheniya Fakul'teta informacionnyh tekhnologij i analiza bol'shih dannyh. M.: Finansovyy universitet. 2023. 1 fajl (6,31 Mb) (In Russian).
7. Zhao X., Xu M., Silamu W., Li Y. CLIP-Llama: A New Approach for Scene Text Recognition with a Pre-Trained Vision-Language Model and a Pre-Trained Language Model. *Sensors* 2024. V. 24. P. 7371. <https://doi.org/10.3390/s24227371>.
8. Nikolaev K., Malafeev A. Russian Q&A method study: From Naive Bayes to convolutional neural networks. *Lecture Notes in Computer Science*. 2018. V. 11179 LNCS. P. 121–126. DOI 10.1007/978-3-030-11027-7_12. EDN XWJVL7.
9. Enweji M.Z., Lehinevych T., Glybovets A. Cross-language text classification with convolutional neural networks from scratch. *Eureka: Physics and Engineering*. 2017. № 2. P. 24–33. DOI 10.21303/2461-4262.2017.00304. EDN YPOYJD.
10. Gallego N. P. D., Ilao J., Cordel M. Blind first-order perspective distortion correction using parallel convolutional neural networks. *Sensors*. 2020. V. 20. № 17. P. 1–20. DOI 10.3390/s20174898. EDN BEKSSY.
11. Ko K., Jang I., Choi J.H. et al. Stochastic decision fusion of convolutional neural networks for tomato ripeness detection in agricultural sorting systems. *Sensors*. 2021. V. 21. № 3. P. 1–14. DOI 10.3390/s21030917. EDN DHJZNO.

12. *Lozhkin A.G., Maiorov K.N., Bozek P.* Convolutional neural networks training for autonomous robotics. *Management Systems in Production Engineering*. 2020. V. 29. № 1. P. 75–79. DOI 10.2478/mspe-2021-0010. EDN TYCRQO.
13. *Andriyanov N.* Methods for preventing visual attacks in convolutional neural networks based on data discard and dimensionality reduction. *Applied Sciences (Switzerland)*. 2021. V. 11. № 11. DOI 10.3390/app11115235. EDN PAAONS.
14. *Bratchenko I.A., Bratchenko L.A., Khristoforova Y.A.* et al. Classification of skin cancer using convolutional neural networks analysis of Raman spectra. *Computer Methods and Programs in Biomedicine*. 2022. V. 219. P. 106755. DOI 10.1016/j.cmpb.2022.106755. EDN BCGDFW.
15. *Andriyanov N.A., Dement'ev V.E., Tashlinskij A.G.* Obnaruzhenie ob"ektov na izobrazhenii: ot kriteriev Bajesa i Nejmana – Pirsona k detektoram na baze nejronnyh setej. *Komp'yuternaya optika*. 2022. T. 46. № 1. S. 139–159. DOI 10.18287/2412-6179-CO-922 (In Russian).
16. *Ayachi R., Aff M., Said Ya.* et al. Integrating Recurrent Neural Networks with Convolutional Neural Networks for Enhanced Traffic Light Detection and Tracking. *Traitement du Signal*. 2023. V. 40. № 6. P. 2577-2586. DOI 10.18280/ts.400620. EDN AWPIQZ.
17. Fine-tuning OCR. URL: https://keras-ocr.readthedocs.io/en/latest/examples/fine_tuning_recognizer.html, data obrashcheniya: 19.03.2025.

Information about the authors

Elena P. Dogadina – Ph.D. (Eng.), Associate Professor

Ulyana Yu. Sukhanova – Intern Researcher

Margarita A. Ishchenko – Intern Researcher

Dmitry I. Veselov – Junior Research Scientist

The article was submitted 04.06.2025

Approved after reviewing 16.06.2025

Accepted for publication 30.06.2025