

## РАСПОЗНАВАНИЕ ТЕКСТА НА ИЗОБРАЖЕНИЯХ ПОСРЕДСТВОМ НЕЙРОННЫХ СЕТЕЙ

*Курбанов С.С.*

*Белорусский государственный университет информатики и радиоэлектроники,  
г. Минск, Республика Беларусь*

*Научный руководитель: Боброва Н.Л. – к.т.н., доцент, доцент кафедры Информатики*

**Аннотация.** Распознавание текста с изображений является важным процессом в машинном обучении. В настоящее время наблюдается большой всплеск хранения информации, содержащейся в бумажном документе, в цифровом формате. Это помогает в сохранении информации, легком хранении, а также позволяет извлекать информацию по мере необходимости. Различные этапы распознавания текста: предварительная обработка, сегментация, извлечение признаков, классификация и постобработка. Этап предварительной обработки включает в себя ряд операций, наиболее важной из которых является преобразование цветного изображения в бинарное изображение, что позволяет отделить текст от фона. Шаг сегментации помогает в разделении символов. Извлечение признаков позволяет получить наиболее важную информацию из изображения, чтобы облегчить распознавание текста. Процесс классификации позволяет идентифицировать текст в соответствии с четко определенными правилами. Затем выполняется постобработка для уменьшения ошибок. Распознавание текста имеет первостепенное значение в нескольких приложениях. В этой статье обсуждается модуль распознавания текста, а также представлены различные приложения распознавания текста из изображений.

**Ключевые слова:** распознавание текста; предварительная обработка; классификация; постобработка

**Введение.** Распознавание текста приобрело большую известность в последние годы, поскольку оно вошло в широкую область приложений, таких как автоматическое считывание номерных знаков, вывесок. В наши дни всё и все стали цифровыми. Большая часть передачи информации в наши дни происходит с помощью изображений или отсканированных документов. Существует огромное количество информации, которая хранится и одновременно используется.

**Основная часть.** Распознавание текста – это область, которая обусловлена необходимостью сохранять и иметь доступ к информации, содержащей документы, более простым и быстрым способом. Одним из удобных способов переноса информации с бумаги или книг является их сканирование, при котором информация преобразуется в изображение, что предотвращает повторное использование отсканированной информации в виде текста. Таким образом, необходимо разработать инструменты для преобразования их в редактируемый вид. Целью данной статьи является изучение различных этапов процесса распознавания текста с целью преобразования текстовых изображений в редактируемые документы. Одним из популярных методов распознавания текста является оптическое распознавание символов (OCR). Он преобразует отсканированные изображения текста в редактируемый формат.

Распознавание текста – утомительная работа, так как она включает в себя распознавание текста разных шрифтов, стилей и с разным фоновым шумом. Кроме того, распознавание рукописного текста еще сложнее из-за различий в размере букв, ориентации и интервалах между буквами, которые варьируются от одного человека к другому. Таким образом, существует потребность в разработке автоматизированной системы распознавания текста, которая может идентифицировать текстовый компонент, присутствующий в изображении или сцене, и преобразовывать его в распознаваемый машиной формат.

Процесс распознавания текста начинается с захвата изображения требуемого документа, его предварительной обработки для получения нужной части, а затем его сегментации для извлечения присутствующего в нем текстового содержимого. В данной статье обсуждаются различные этапы задачи распознавания текста с изображений.

В этом разделе представлен краткий обзор существующих работ, проводимых в области распознавания текста.

В первой работе [1] рассматриваются изображения с цветным фоном и описывается метод предварительной обработки, улучшающий производительность механизма оптического распознавания символов (OCR) Tesseract. Здесь выполняется первая сегментация текста, чтобы отделить текст от цветного фона путем разделения исходного изображения на  $k$  изображений. Затем классификатор распознает изображение, содержащее текст. Было достигнуто улучшение примерно на 20% по сравнению с производительностью Tesseract OCR за счет использования предварительной обработки.

Работа С. Акопяна, О.В. Беляева, Т.П. Плехова и Д.Ю. Турдакова [2] основана на конвейере извлечения текста, который используется для извлечения текста из изображений различного качества, полученных из социальных сетей. Их работа в основном сосредоточена на разделении входных изображений на различные классы, а затем в зависимости от классов выполняется предварительная обработка. Далее следует распознавание текста с использованием механизма OCR. В данной работе используется набор данных, собранный из социальных сетей.

OCR используется для идентификации текстового компонента, присутствующего в изображениях. В третьей рассматриваемой работе [3] авторы предложили алгоритм извлечения текста из отсканированного документа. В этой работе для сегментации использовался алгоритм Отсуса, а для обнаружения асимметрии использовалось преобразование Хафа. Также была применена технология OCR для идентификации символов. Они провели эксперименты и проверили предложенный алгоритм на различных изображениях, взятых из разных источников. Средняя точность оказалась равной 93%.

К. Картик, К.Б. Равиндракумар, Р. Фрэнсис и С. Иланканнан [4] подробно обсудили различные этапы обнаружения текста, выделив различные используемые методы. Они также сделали упор на распознавание рукописного текста, которое является одной из сложных областей. В результате их исследования было обнаружено, что наилучшие результаты можно получить при сокращении времени вычислений, а также можно сегментировать многоязычные символы и повысить скорость распознавания символов.

Ануприя Шривастава, Амудха Дж., Дипа Гупта и Кшитий Шарма [5] в своей работе разработали систему, основанную на сверточной нейронной сети и LSTM – сети долгой краткосрочной памяти. Разработанная модель идентифицирует тексты по изображениям в горизонтальном, изогнутом или ориентированном стиле. Модель состоит из четырех компонентов. Первый компонент выполняет извлечение объектов на низком уровне. Второй компонент использует общий подход свертки для извлечения объектов высокого уровня. Нерелевантные функции игнорируются третьим компонентом. Четвертый компонент предсказывает последовательности символов.

Прадик Мадхукар Манваткар и доктор Кавита Р. Сингх [6] рассмотрели в своей статье различные методы извлечения символов из изображений. Базовая архитектура системы распознавания текста по изображениям также описана в их работе. Они также обсудили последовательность методов обработки изображений для извлечения текста из отсканированного изображения. В их статье также описаны различные области применения.

Модуль распознавания текста должен выполнять ряд задач. На вход модуля подается изображение, содержащее текст. Выходом модуля является текстовая информация в машиночитаемом виде. Модуль распознавания текста должен выполнять следующие задачи: предварительная обработка, сегментация, извлечение признаков и классификация (рисунок 1).

Отсканированный документ обычно представляет собой изображение. Первым шагом является предварительная обработка, которая заключается в преобразовании изображения в формат, пригодный для дальнейшей обработки. Текстовое изображение может содержать шум или искажено. На этом этапе изображение улучшается путем удаления шума, а затем преобразуется в двоичное. Шум, присутствующий на изображении, играет важную роль в успешном распознавании текста. Удаление шума повышает вероятность точного распознавания текста и обеспечивает более точный вывод. Различные фильтры, такие как фильтр Гаусса, средний фильтр, могут использоваться для удаления шума. Затем выполняется нормализация для обеспечения однородности, за которой следует бинаризация для преобразования серого изображения в бинарное изображение.

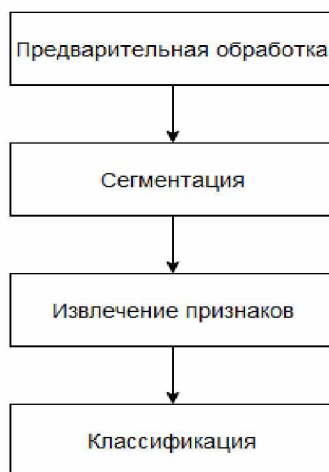


Рисунок 1 – Различные задачи по распознаванию текста

После завершения предварительной обработки отдельные символы разделяются с помощью процесса сегментации. Затем жизненно важные данные извлекаются из необработанных данных с использованием этапа извлечения признаков. Для целей сегментации можно использовать различные методы, такие как анализ основных компонентов (PCA), линейный дискриминационный анализ (LDA), анализ независимых компонентов (ICA), цепной код (CC), гистограмму и т. д. [6].

Следующим шагом является классификация, которая включает в себя распознавание каждого символа и отнесение его к нужному классу символов, таким образом преобразуя текст в машиночитаемую форму. Для этой цели можно использовать различные классификаторы, основанные на искусственной нейронной сети (ANN), метод опорных векторов (SVM). Постобработка включает в себя сохранение распознанного текста в формате, пригодном для дальнейшей обработки.

В последние годы несколько приложений обнаружения текста получили известность. Основным из них является автоматическое обнаружение текста из отсканированного документа. Недавние технологические достижения еще больше способствовали развитию методов распознавания текста. Распознавание текста позволяет автоматизировать различные области. Его можно использовать для автоматического считывания номерных знаков в пунктах взимания платы; автоматизированное считывание подписей на чековых листах; тегирование изображений и анализ данных сцены.

Распознавание текста также обеспечивает автоматизированное хранение и доступ к огромным документам в сфере здравоохранения, офисах. Это позволяет создать базу данных, в которой текст можно легко найти. Доступ к созданным базам данных и их обновление не требуют особых усилий, что экономит много бумажной работы. Автоматическое обнаружение текста с голосовой поддержкой также является благом для

людей с нарушениями зрения. Его также можно использовать для распознавания текстовой составляющей в видео.

Автоматизированная текстовая идентификация также может быть использована для повышения интеллектуальности транспортных систем. Его также можно использовать в аэропортах для проверки паспорта и извлечения информации. Обнаружение текста также позволяет автоматически вводить данные для деловых документов. Рассматриваемая процедура делает огромное количество книг доступными в Интернете, что позволяет сохранять знания и делиться ими. Распознавание текста также обеспечивает автоматизацию в промышленности, поддерживая автоматическое считывание этикеток и цифр. С дальнейшим развитием технологии распознавание текста нашло широкое применение практически во всех областях.

**Заключение.** В этой статье представлен краткий обзор различных шагов, используемых при распознавании текста с изображений. Также кратко обсуждалась работы, проводимых в этой области, также дан обзор базовой модели системы распознавания текста, описывающей поток распознавания текста с изображений.

### Список литературы

1. Маттео Бризинелло, Ратко Грби, Деян Стефанови и Роберт Пекай-Кова, *Оптическое распознавание символов на изображениях с цветным фоном*, 2018 г., 8-я Международная конференция IEEE по бытовой электронике — Берлин (ICSE-Берлин).
2. М.С. Акопян, О.В. Беляева, Т.П. Плехов и Д.Ю. Турдаков, *Распознавание текста на изображениях из социальных сетей*, 2019.
3. Неха Агравал, Арашдин Каур, *Алгоритмический подход к распознаванию текста из печатных/набранных текстовых изображений*, 2018.
4. К. Картик, К.Б. Равиндракумар, Р. Фрэнсис, С. Иланканнан, Шаги, связанные с распознаванием текста, и последние исследования в области оптического распознавания символов; *Исследование, Международный журнал новейших технологий и инженерии (IJRTE)*, ISSN: 2277-3878, том 8, выпуск 1, май 2019 г.
5. Ануприя Шривастава, Амудха Дж., Дипа Гупта, Киштий Шарма, *Модель глубокого обучения для распознавания текста в изображениях*, 10-я ICCCNT 2019, 6–8 июля 2019 г..
6. Прадик Мадхукар Манваткар, доктор Кавита Р. Сингх, *Технический обзор распознавания текста из изображений*, 2021 г..

UDC 621.3.049.77–048.24:537.2

## TEXT RECOGNITION ON IMAGES USING NEURAL NETWORKS

*Kurbanov S.S.*

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus*

*Bobrova N.L. – Cand. of Sci. (Tech)., associate professor, associate professor of the department of Computer Science*

**Annotation.** Text recognition from images is an important process in machine learning. Currently, there is a significant surge in the storage of information contained in paper documents in digital format. This aids in information preservation, easy storage, and enables information retrieval as needed. The various stages of text recognition include pre-processing, segmentation, feature extraction, classification, and post-processing. The pre-processing stage involves a series of operations, with the most crucial one being the conversion of a color image into a binary image, separating the text from the background. The segmentation step helps in separating individual characters. Feature extraction allows obtaining the most important information from the image to facilitate text recognition. The classification process identifies the text according to well-defined rules. Post-processing is then performed to reduce errors. Text recognition holds paramount importance in several applications. This article discusses the text recognition module and presents various applications of text recognition from images.

**Keywords:** text recognition, pre-processing, classification, post-processing