

## СИСТЕМА РАСПОЗНАВАНИЯ ТЕКСТА НА ГРАФИЧЕСКИХ ИЗОБРАЖЕНИЯХ

*Воронцова П.В.*

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

*Деменковец Д.В. – старший преподаватель*

В данной работе рассматриваются система распознавания текста на графических изображениях. Описываются такие алгоритмы, как предобработка данных и обучение нейронной сети. Представлены примеры работы медианной и монохромной фильтраций, которые используются для предварительной обработки изображения к обучению нейронной сети.

Системы распознавания текста на графических изображениях – это технология, которая позволяет компьютеру распознавать и переводить на понятный для него язык текст, который находится на изображении. Система может работать с различными языками и шрифтами, что делает ее универсальной и применимой в различных областях [1].

При создании системы распознавания текста на графических изображениях была выполнена следующая последовательность действий.

Выполняется сбор данных, который включает в себя накопление графических изображений с текстом, чтобы использовать их для обучения системы. Это достаточно трудоемкий процесс, так как необходимо собрать большой объем изображений, на которых присутствует текст. Для набора данных используются различные источники. В основном источниками являются открытые базы данных и различные интернет-ресурсы. Также при сборе данных важно учитывать разнообразие текстов на изображениях для эффективной работы модели по распознаванию различных видов текста. Некоторые изображения могут быть защищены авторским правом или содержать конфиденциальную информацию. Поэтому необходимо убедиться в том, что использование этих изображений не нарушает законодательство.

После этого выполняется предобработка данных. Этот этап включает в себя предварительную обработку изображений для исключения шумов, улучшения контрастности и приведения их к единому формату. Для обработки используются 2 фильтра: медианный и монохромный. Медианный фильтр применяется для минимизации шума и сглаживания острых краев букв (засечек), искажающих его размер и препятствующий дальнейшему распознаванию. Суть этого фильтра заключается в замене каждого пикселя на изображении на медианное значение яркости его окрестности. Пример работы медианной фильтрации приведен на рисунке 1.

Далее изображение обрабатывается с помощью монохромного фильтра, который выполняет четкую бинаризацию изображения, при этом фиксируя границы символа. При монохромной фильтрации цветные пиксели изображения преобразуются в черно-белые (монохромные) пиксели. Этот фильтр позволяет сделать изображение более четким и тем самым упростить процесс распознавания символов текста. Пример работы монохромной фильтрации приведен на рисунке 2. При этом монохромные изображения занимают значительно меньше места в памяти компьютера, ускоряя обработку данных [2].

Полный алгоритм предобработки массива графических изображения приведен на рисунке 3.

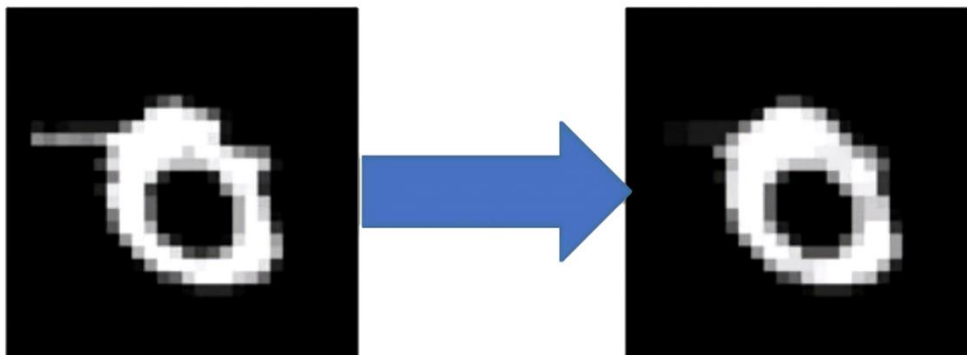


Рисунок 1 – Медианная фильтрация

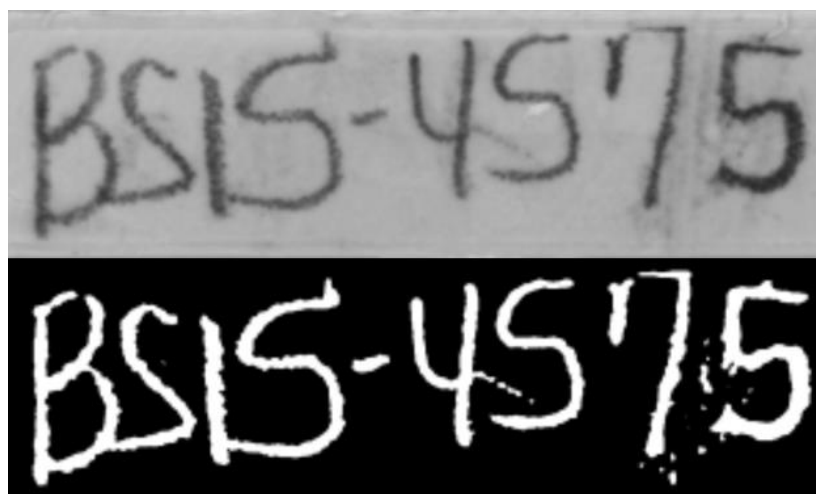


Рисунок 2 – Монохромная фильтрация



Рисунок 3 – Предобработка массива изображений

После предобработки массива изображений следует процесс обучения модели. Для этого используются различные алгоритмы машинного обучения, такие как сверточные нейронные сети или методы опорных векторов [3]. Алгоритм обучения нейронной сети приведен на рисунке 4.



Рисунок 4 – Обучение нейронной сети

Завершающими этапами являются тестирование и оптимизация. Модель тестируется на новых изображениях. Выполняется расчет параметров и оптимизация модели для достижения наилучшей точности распознавания текста в изображении.

В результате была разработана система, позволяющая извлекать текстовую информацию из изображений. Возможен вариант использования данной системы во многих областях. К примеру, распознавание номеров автомобилей, оптическое распознавание символов на чеках и документах, анализ рукописного текста. Это позволит ускорить и автоматизировать процесс обработки текстовых данных с изображений.

**Список использованных источников:**

1. "OCR and Its Applications in Industry." *International Journal of Computer Science and Mobile Computing*. [Электронный ресурс]. – Режим доступа: <https://www.ijcsmc.com/docs/papers/March2020/V9I3202002.pdf> – Дата доступа: 15.03.2023.
2. "Цифровая обработка сигналов" [Электронный ресурс]. – Режим доступа: [https://dl.booksee.org/genesis/43000/0d06ddd492b9413c9dd6999773695a56/as/\[Glinchenko\\_A.S.\]\\_Cifrovaya\\_obrabotka\\_signalov\(BookSee.org\).pdf](https://dl.booksee.org/genesis/43000/0d06ddd492b9413c9dd6999773695a56/as/[Glinchenko_A.S.]_Cifrovaya_obrabotka_signalov(BookSee.org).pdf) – Дата доступа: 10.03.2023.
3. "Практический курс по нейронным сетям." [Электронный ресурс]. – Режим доступа: <https://books.ifmo.ru/file/pdf/2394.pdf> – Дата доступа: 12.03.2023.