

*Е.В. Козлова, В.В. Новый*¹

*Витебский государственный университет им. П.М. Машиерова,
г. Витебск, Беларусь*

¹SPIN-код: 8336-8713

ОБНАРУЖЕНИЕ ПЕЧАТНЫХ СИМВОЛОВ НА ЦИФРОВЫХ ИЗОБРАЖЕНИЯХ МЕТОДАМИ ГЛУБОКОГО ОБУЧЕНИЯ

Современные технологии обработки и распознавания изображений активно применяются в самых различных областях, от автоматизации документооборота до систем безопасности. Одной из ключевых задач в данной области является распознавание текстовой информации, представленной в виде печатных цифровых символов. Этот процесс часто требует высокоточных методов, способных обеспечивать высокую эффективность и точность при различных условиях, таких как недостаточное качество изображения, отличающиеся шрифты, трудноразличимые по схожести символы и различные шумовые помехи.

Цель работы заключается в разработке приложения для оптического распознавания символов с использованием методов глубокого обучения для распознавания на входном цифровом изображении печатных цифровых символов десятичной системы счисления и распознавания текстов на изображении.

В ходе работы применялись такие методы исследования, как изучение и сравнительный анализ. В качестве материала в исследовании использовались научные статьи, официальные Интернет-ресурсы, документация по применяемым технологиям. Для обучения и тестирования нейронной сети были использованы два набора входных данных, содержащих печатные цифровые символы на белом фоне и цветные символы на цветном фоне. В качестве методов распознавания применяются свёрточные нейронные сети (CNN), оптическое распознавание символов.

Оптическое распознавание символов (англ. optical character recognition (OCR)) – это процесс классификации оптических об-

разов, содержащихся в цифровых изображениях на соответствующие буквенно-цифровые и другие символы [1, с. 9]. Таким образом, оптическое распознавание символов – удобная и гибкая архитектура по преобразованию символов из изображений или сканированных документов в машинно-читаемый текст.

Оптическое распознавание символов включает в себя несколько последовательных этапов: предварительная обработка изображения, сегментация, распознавание символов на изображении и вывод текстового результата.

Цветные изображения часто содержат избыточную информацию, которая не нужна для распознавания текста. Поэтому изображения преобразуются в оттенки серого. Это упрощает дальнейшую обработку и уменьшает объем данных, с которыми работает модель. Изображения могут содержать различные виды шума, такие как пятна, линии или артефакты, которые могут мешать распознаванию текста. Для удаления шума используются фильтры, такие как медианный фильтр или гауссово размытие. Эти методы помогают сгладить изображение и убрать мелкие дефекты [2].

Сегментация текста – это процесс разделения изображения на отдельные текстовые блоки, строки и символы. Этот шаг важен для точного распознавания, так как модель должна обрабатывать текст по частям. Сегментация может быть выполнена с помощью различных методов.

Определенные текстовые блоки сортируются перед этапом распознавания для исключения потери смысловой нагрузки распознаваемого текста. Для этапа распознавания была разработана и обучена свёрточная нейронная сеть. Определяемые блоки символов выделяются на изображении. Результатом оптического распознавания символов является изображение с выделенными печатными символами и выводом текста к изображению.

Таким образом мы получаем самостоятельный программный модуль, который может применяться для распознавания номерных знаков, сигналов светофора, при сканировании документов, автоматизации анализа данных на цифровом изображении при автоматической обработке и многих других видах деятельности.

Заключение. В результате выполнения работы было создано приложение, использующее методы глубокого обучения

для выделения и распознавания заданного набора печатных символов на цифровых изображениях.

Литература

1. *Arindam Chaudhuri, Krupa Mandaviya, Pratixa Badelia, Soumya K Ghosh. Optical Character Recognition Systems for Different Languages with Soft Computing. Springer, 2017.*
2. *Береснев Д.В. Исследования методов распознавания текстовых документов с использованием компьютерного зрения // Искусственный интеллект в промышленных, коммерческих, медицинских и финансовых приложениях: сборник статей научно-технического семинара студентов кафедры «Инженерной кибернетики». Москва, 30–31 мая 2024 года. М.: Национальный исследовательский технологический университет «МИСИС», 2024. Вып. 2. С. 23–29.*