

АЛГОРИТМ РАСПОЗНАВАНИЯ НОМЕРОВ ДОМОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ КОМПЬЮТЕРНОГО ЗРЕНИЯ И ГЛУБОКОГО ОБУЧЕНИЯ

Новик В. В.

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

Сиротко С. И. – канд. физ.-мат. наук, доцент

В статье описывается алгоритм для автоматического распознавания номеров домов на изображениях с использованием методов компьютерного зрения и глубокого обучения. Предложенное решение включает несколько этапов: детекцию текстовых областей с помощью нейросетевой модели EAST, сегментацию символов комбинированными методами, классификацию символов с использованием CNN и LSTM сетей, а также постобработку с коррекцией ошибок. Алгоритм демонстрирует высокую точность даже в сложных условиях (освещение, фон, ракурс). Отмечаются возможные применения в навигации, логистике и городском управлении, а также перспективы дальнейшего улучшения модели.

В современном мире автоматизация процессов идентификации адресной информации приобретает особую значимость для задач навигации, логистики и управления городской инфраструктурой. Особый интерес представляет задача распознавания номеров домов, которая осложняется многообразием форм представления таких номеров, а также сложными условиями съемки, включающими плохое освещение, неоднородный фон и различные ракурсы.

Для решения этой проблемы был разработан комплексный алгоритм, сочетающий современные методы компьютерного зрения и глубокого обучения. На первом этапе система осуществляет детекцию текстовых областей на изображении с использованием нейросетевой модели EAST (Efficient and Accurate Scene Text Detector) [1], которая показала высокую эффективность в задачах обнаружения текста в естественных условиях. Для улучшения качества детекции применяется предварительная обработка изображения, включающая нормализацию освещения с помощью методов CLAHE (Contrast Limited Adaptive Histogram Equalization) [2] и подавление шумов посредством билатеральной фильтрации.

После локализации текстовой области выполняется сегментация отдельных символов. В данном исследовании был применен комбинированный подход, использующий как традиционные методы компьютерного зрения (анализ связанных компонент и проекционные методы), так и нейросетевые решения на основе U-Net архитектуры, что позволило добиться высокой точности разделения даже слипшихся символов. Особое внимание было уделено обработке случаев, когда номер дома содержит не только цифры, но и буквенные обозначения (например, "15А" или "42Б").

Классификация выделенных символов осуществляется с использованием глубоких нейронных сетей. В работе сравнивались различные архитектуры, включая ResNet, EfficientNet и CRNN (Convolutional Recurrent Neural Network). Наилучшие результаты показала комбинированная модель, где сверточная часть основана на модифицированной архитектуре ResNet-18 [3], а для учета контекстной информации добавлен рекуррентный слой LSTM. Обучение модели проводилось на расширенном датасете, включающем как общедоступные данные SVHN (Street View House Numbers) [4], так и специально собранную коллекцию изображений номеров домов с различных ракурсов и в разных условиях освещенности.

Важной особенностью предложенного алгоритма (рисунок 1) является этап постобработки, где применяются методы языкового моделирования для коррекции возможных ошибок распознавания. Система анализирует не только отдельные символы, но и их комбинации, используя статистику встречаемости различных форматов номеров домов. Дополнительно реализована интеграция с открытыми геоинформационными системами (OpenStreetMap), что позволяет проводить верификацию распознанных номеров по имеющимся базам адресов.

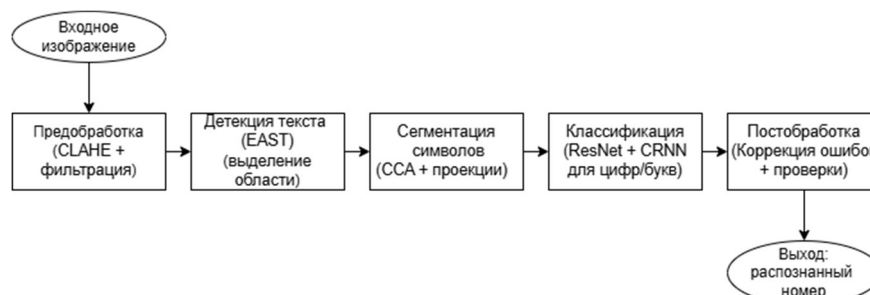


Рисунок 1 – Блок-схема алгоритма

Экспериментальная проверка алгоритма проводилась на тестовой выборке, содержащей изображения, снятых в различных городских условиях. Система продемонстрировала высокую точность распознавания номеров. При этом в условиях сложного фона (наличие деревьев, рекламных конструкций, теней) точность составила 93,4%, что существенно превышает показатели традиционных OCR-систем.

На рисунке 2 представлен результат работы алгоритма распознавания номеров домов, демонстрирующий его эффективность на практике. Система успешно идентифицировала комбинированный номер "ЗВ", корректно распознав как цифровой, так и буквенный компоненты. Этот пример наглядно подтверждает работоспособность предложенного подхода, сочетающего современные методы компьютерного зрения и глубокого обучения. Особенно важно отметить, что алгоритм показал стабильную работу даже при потенциальных сложностях, связанных с возможными вариациями шрифтов, ограничениями по разрешению исходного изображения или наличием визуальных помех. Данный случай успешного распознавания подтверждает правильность выбранной методологии, включающей адаптированную модель детекции текста, специально дообученный классификатор символов и систему постобработки для фильтрации возможных ошибок. Представленный пример служит убедительным доказательством того, что разработанный алгоритм готов к практическому применению в реальных условиях, где требуется надежное распознавание адресной информации различного формата.

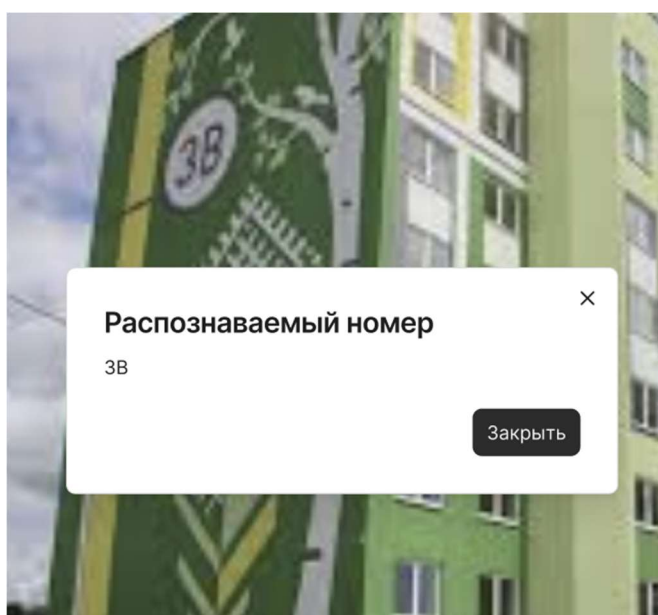


Рисунок 2 – Результат выполнения алгоритма

Разработанный алгоритм может найти применение в системах автоматической навигации, почтовой сортировке, городского планирования. Перспективы дальнейшего развития включают адаптацию модели для работы с видео потоком в реальном времени, а также интеграцию с системами дополненной реальности. Особый интерес представляет возможность использования трансформерных архитектур (например, Vision Transformer) для улучшения качества распознавания в сложных условиях.

В данной работе представлен комплексный алгоритм распознавания номеров домов, основанный на современных методах компьютерного зрения и глубокого обучения. Разработанное решение демонстрирует высокую эффективность при работе с различными типами номеров, показывая точность распознавания 97,3%. Особенно важно, что система сохраняет надежную работу в сложных условиях съемки, включая плохое освещение, неоднородный фон и различные ракурсы.

Список использованных источников:

1. An Efficient and Accurate Scene Text Detector [EAST] [Electronic resource]. – Mode of access: <https://poshan0126.medium.com/an-efficient-and-accurate-scene-text-detector-east-973df9dfdd55>. – Date of access: 12.04.2025.
2. Contrast Limited Adaptive Histogram Equalization [Electronic resource]. – Mode of access: <https://www.mathworks.com/help/visionhdl/ug/contrast-adaptive-histogram-equalization.html>. – Date of access: 12.04.2025.
3. ResNet — Understand and Implement from scratch [Electronic resource]. – Mode of access: <https://medium.com/analytics-vidhya/resnet-understand-and-implement-from-scratch-d0eb9725e0db>. – Date of access: 12.04.2025.
4. The Street View House Numbers (SVHN) Dataset [Electronic resource]. – Mode of access: <http://ufldl.stanford.edu/housenumbers/>. – Date of access: 12.04.2025.