

Распознавание текста на основе скелетного представления линий с шириной и сверточных сетей

Петров Виктор Иванович

Московский физико-технический институт

Научный руководитель: Рейер Иван Александрович, к.т.н., ФИЦ ИУ РАН

- Современные методы распознавания текста основаны на нейросетевых моделях, но:
 - Требуют больших объемов разметки
 - Чувствительны к шумам и искажениям
 - Ограничены в интерпретируемости результатов
- Графовые методы могут улучшить качество распознавания за счет структурного представления текста

Постановка задачи

- Разработать метод распознавания текста, основанный на графовом представлении, который:
- Использует топологические и семантические связи символов и слов
- Повышает устойчивость к искажениям
- Снижает зависимость от обучающих данных

Цель работы

Иследуются

Различные варианты обработки графового представления символов с помощью нейросетевых методов.

Требуется

Предложить архитектуру модели, дающую лучшее качество по сравнению с современными методами при том же количестве параметров

- Проведен анализ существующих подходов
- Реализована предобработка изображений текста для выделения графовых структур
- Реализован алгоритм построения графа на основе связности символов
- Обучена базовая модель, с которой будет сравниваться качества дальнейших моделей

- Датасет MNIST предобработан : проведена бинаризация (с использование Гауссова ядра)
- Библиотека для скелетизации адаптирована под современное программное обеспечение
- Проведено исследование гиперпараметров скелетизации с визуализацией
- Построен базовый пайплайн обработки данных для дальнейшей загрузки в графовую нейросеть

- Разработка алгоритмов распознавания текста на графе
- Обучение и тестирование модели
- Сравнение с существующими методами (CNN)
- Написание статьи