

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

- 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

Исследуются

Различные варианты обработки графового представления символов с помощью нейросетевых методов.

Требуется

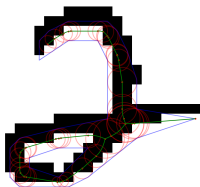
Предложить архитектуру модели, дающую лучшее качество по сравнению с современными методами при том же количестве параметров

- Проведен анализ существующих подходов
- Датасет MNIST предобработан : проведена бинаризация (с использованием 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
- Построены первые графовые нейросети, обученные на предобработанном датасете

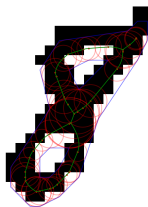
Визуализация полученных скелетных представлений

Скелеты #16 и #17

#16



#17



Промежуточный результат обучения графовой нейросети

```
Total valid graphs: 59999
Graph 0: Nodes=28, Edges=27, Label=5
Graph 1: Nodes=36, Edges=36, Label=0
Graph 2: Nodes=23, Edges=22, Label=4
Number of classes: 10
params: 67722
Epoch: 001, Train Loss: 1.8057, Train Acc: 0.4088, Test Acc: 0.5466
Epoch: 002, Train Loss: 1.4741, Train Acc: 0.5844, Test Acc: 0.5701
Epoch: 003, Train Loss: 1.4230, Train Acc: 0.6098, Test Acc: 0.5542
Epoch: 004, Train Loss: 1.4002, Train Acc: 0.6183, Test Acc: 0.5757
Epoch: 005, Train Loss: 1.3706, Train Acc: 0.6292, Test Acc: 0.5781
Epoch: 006, Train Loss: 1.3818, Train Acc: 0.6263, Test Acc: 0.5861
Epoch: 007, Train Loss: 1.3664, Train Acc: 0.6319, Test Acc: 0.5460
Epoch: 008, Train Loss: 1.3614, Train Acc: 0.6313, Test Acc: 0.6072
Epoch: 009, Train Loss: 1.3538, Train Acc: 0.6368, Test Acc: 0.5270
Epoch: 010, Train Loss: 1.3619, Train Acc: 0.6341, Test Acc: 0.5457
Best Test Accuracy: 0.6072
```

Процесс обучения

- Подбор оптимальной архитектуры для графовой нейросети
- Обучение и тестирование различных архитектур
- Сравнение с существующими методами (CNN)
- Написание статьи