

Разработка механизма автоматического перевода текста на фотографии для мобильных устройств

Ткаченко Дмитрий Андреевич
ФПМИ МФТИ
Научный руководитель:
Родюков Александр Витальевич

Цели работы:

- Изучение передовых на данный момент алгоритмов распознавания текста с изображения и собственная имплементация.
- Изучение и использование механизма портирования готовой модели в формат, пригодный для использования на мобильной платформе iOS.
- Развитие навыков полного цикла разработки: дизайн и разработка интерфейса, изучение и внедрение сторонних библиотек и API, архитектурная и прикладная разработка.

Поставленные задачи:

- Изучение теоретического материала по теме распознавания текста с фотографии и мобильной разработки на платформе iOS.
- Изучение уже созданных вариантов реализации алгоритмов распознавания текста с фотографии и формального обоснования корректности их работы.
- Изучение и анализ инструментов, позволяющих создать и обучить требуемую нейронную сеть, а затем портировать ее в формат, воспринимаемый устройством на платформе iOS.

Поставленные задачи:

- Разработка архитектуры мобильного приложения с целью разделения ответственности компонент, быстродействия приложения и отсутствия ошибок во время исполнения.

- Программная реализация нейронной сети на языке программирования Python с помощью библиотеки для работы с нейронными сетями Keras (так как используется механизм конвертации из этого формата) и программная реализация мобильного приложения на платформе iOS на языках Swift, Obj-C и C++ (так как для предобработки фотографии используется библиотека OpenCV, доступная для C++, но не для Swift или Obj-C).

Результаты проделанной работы:

1. Завершена разработка интерфейса и архитектуры мобильного приложения.
2. Завершена разработка мобильного приложения, производится тестирование.
3. Изучена и реализована одна из архитектур нейронной сети для распознавания текста. Обученная модель портирована в формат, пригодный для использования на платформе iOS. Эта модель встроена в приложение.
4. Использованы сторонние библиотеки распознавание текста для сравнительного анализа.
5. Изучен и реализован механизм использования C++ программного кода в проекте, реализованном на языке Swift для предобработки фотографий с помощью библиотеки OpenCV.
6. Проводится сравнение со сторонними моделями распознавания текста.
7. На финальной стадии написание текста НИР.

Описание работы:

Содержание работы разбито на 9 частей, описание каждой из которых содержит

- 1) Анализ задач, которые должно решать приложение и разработка его интерфейса, инструменты для данной разработки.
- 2) Анализ существующих архитектур и подходов в архитектурной разработке, создание архитектуры приложения.
- 3) Написание программного кода приложения, использованные инструменты и подходы, специфики языка, использованные при разработке.
- 4) Сторонние библиотеки, использованные при разработке, в частности библиотека распознавания текста.

Описание работы:

- 5) Описание архитектуры нейронной сети, ее изменение для возможности портирование и описание процесса обучения.
- 6) Работа с библиотекой CoreML для портирования нейронной сети.
- 7) Механизм использования C++ программного кода в проекте, реализованном на языке Swift с целью предобработки фотографий.
- 8) Результаты тестирования приложения.
- 9) Общие результаты разработки и выводы.

Спасибо за внимание