МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ «МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ (ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)»

ФАКУЛЬТЕТ ИННОВАЦИЙ И ВЫСОКИХ ТЕХНОЛОГИЙ

БАЗОВАЯ КАФЕДРА «КОРПОРАТИВНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ»

Автореферат  
выпускной бакалаврской квалификационной работы  
Ткаченко Дмитрия Андреевича  
на тему:  
«Разработка механизма автоматического перевода текста на фотографии для мобильных устройств»

*Специальность 010302 –  
Прикладные математика и информатика*

Научный руководитель:

Родюков А.В.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

подпись и дата

Студент 4 курса, 597 гр.:

Ткаченко Дмитрий Андреевич

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

подпись и дата

Москва, 2019

Ткаченко Дмитрий Андреевич

Аннотация к автореферату на тему

**«Разработка механизма автоматического перевода текста на фотографии для мобильных устройств»**

©Кафедра корпоративных информационных систем ФИВТ МФТИ

1. **Тема работы**

«Разработка механизма автоматического перевода текста на фотографии для мобильных устройств».

1. **Обоснование выбора темы**
2. **Объект, предмет и метод исследования**

**Объектом исследования** являются механизмы по распознаванию текста на фотографии, библиотеки для перевода и механизмы проектирования и реализации мобильного приложения.

**Предметом исследования** является процесс создания приложения по автоматическому переводу на платформе iOS.

**Методы исследования** — анализ предметной области, разработка дизайна и архитектуры приложения, анализ и использование реализованных библиотек перевода, реализация модели распознавания текста, ее портирование в формат, пригодный для использования на мобильном устройстве и написание программного кода.

1. **Актуальность выбранной темы**

С развитием технологий человечество вошло в новую эпоху - эру мобильных устройств. Их количество исчисляется миллиардами, люди проводят часы жизни за ними и уже вряд ли могут представить себя без них. В данной ситуации крайне важно использовать мобильные устройства как помощников в быту или в решении технических задач. Сейчас вычислительные мощности Вашего смартфона или планшета уже превосходят мощности стационарных компьютеров 2-3 летней давности. Таким образом, с их помощью можно решать и ресурсоемкие задачи, которые раньше были под силу только мощным системам. Примером такой задачи является распознавание текста на фотографии. Но только распознать текст мало - его можно и нужно использовать для дальнейшей обработки и получения информации. Но, если текст на незнакомом для Вас языке, то он нуждается в переводе. Такая ситуация часто встречается в путешествиях или на конференциях, которых проводится все больше благодаря глобализации.. Следовательно, данная тема является актуальной.

1. **Постановка исследовательской задачи**
2. **Формулировка проблемы**

На данный момент достаточно хорошо и точно решена задача распознавания текста на фотографии, но самые точные алгоритмы предполагают реализацию и обучение нейронной сети, наиболее удобным и распространенным инструментом для разработки которых является язык Python. Запустить код языка Python на мобильном устройстве не представляется возможным. Приложения для платформы iOS написаны на языках Swift и Obj-C, для которых нет библиотек по работе с нейронными сетями. Таким образом необходим иной механизм использования нейронной сети на мобильном устройстве – конвертация обученной модели в формат, пригодный для использования на мобильном устройстве. По сравнению с разработкой, предполагающей отправку фотографии на сервер и использующей нейронную сеть, а затем отправляющую ответ, будет получен куда более быстрый отклик, так как отпадут издержки на передачу информации.

1. **Цели и задачи исследования**

В рамках данной выпускной бакалаврской квалификационной работы предлагается реализовать в программном коде нейронную сеть для распознавания текста с фотографии и мобильное приложение на платформе iOS, а затем интегрировать полученную нейронную сеть в это приложение, получив решение проблемы, сформулированной в п.3.1.

В работе решаются следующие 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.
* 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).
* Тестирование полученных механизмов на корректность и производительность.
* Формулировка общих выводов, идей по улучшению.

1. **Принятые при достижении цели ограничения**

* Ограничения на выбор библиотеки по работе с нейронными сетями, накладываемые ограничением на источник механизма конвертации в формат, пригодный для работы на платформе iOS.

1. **Подход к решению задач и достижению целей**

Изучаются теоретические основы алгоритмов распознавания текста на фотографии; изучаются основы архитектурной разработки мобильного приложения; изучаются инструменты и механизмы для создания макета мобильного приложения; рассматриваются различные варианты реализации нейронных сетей для распознавания текста с фотографии; изучаются особенности языков программирования Swift и Obj-C, позволяющие наиболее эффективно и правильно с точки зрения архитектуры и ООП реализовать мобильное приложение; изучается механизм использования программного кода, написанного на языке С++ в проекте, написанном на языке Swift с целью использования библиотеки, написанной на языке C++; строится архитектурная модель мобильного приложения; вносятся изменения в архитектуру нейронной сети, которые позволяют снять некоторые ограничения, наложенные механизмом портирования данной модели; созданная и обученная нейронная сеть портируется с помощью библиотеки CoreML в формат, который может быть использован в платформе iOS; проводится тестирование на корректность работы и сравнение с другими решениями.

1. **Структура и содержание выпускной квалификационной работы**
2. **Структура и объем выпускной квалификационной работы с кратким изложением их основного содержания**

Работа состоит из введения, девяти глав, заключения и списка использованной литературы.

В первой главе производится анализ задач, которые должно решать приложение и разработка его интерфейса. Рассматриваются инструменты для создания макетов мобильных приложений и сам процесс создания.

Во второй главе рассматриваются существующие архитектуры и архитектурные решения для разработки приложения на мобильной платформе, приводится описание архитектуры создаваемого приложения.

В третьей главе речь идет о написании программного кода приложения, использованных инструментах и подходах, спецификах языка Swift, использованных при разработке.

В четвертой главе освещаются сторонние библиотеки, использованные при разработке приложения, в частности библиотека для перевода и распознавания текста.

В пятой главе приводится описание архитектуры нейронной сети для распознавания текста с фотографии, ее изменение для возможности портирования и описание результатов обучения.

В шестой главе описывается работа с библиотекой CoreML для портирования нейронной сети.

В седьмой главе описывается механизм использования C++ программного кода в проекте, написанном на языке Swift с целью предобработки фотографий.

В девятой главе описываются окончательные результаты разработки и выводы.

1. **Результаты работы**

* Изучены теоретические материалы по алгоритмам распознавания текста с фотографии.
* Изучены различные варианты реализации нейронных сетей для распознавания текста с фотографии.
* Изучены инструменты для разработки интерфейсов.
* Разработана архитектура для мобильного приложения. Реализовано само приложение.
* Изучены и реализованы механизмы использования C++ программного кода в проекте, написанном на языке программирования Swift.
* Изучен и применен механизм портирования обученной с помощью библиотеки Keras нейронной сети в формат, пригодный для использования на платформе iOS.
* Изучены и использованы сторонние библиотеки по распознаванию текста для сравнительного анализа.
* Проведено тестирование на корректность.
* Сделаны соответствующие выводы.

1. **Список литературы**