

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет
информатики и радиоэлектроники»

УДК 004.383

Папроцкий
Эдуард Валерьевич

**Применение нейронных сетей глубокого обучения для решения задач
распознавания изображений**

АВТОРЕФЕРАТ

диссертации на соискание степени
магистра информатики и вычислительной техники

по специальности 1-40 81 01 – Информатика и технологии разработки про-
граммного обеспечения

Научный руководитель
Егорова Н.Г.
к.т.н., доцент

Минск 2020

ВВЕДЕНИЕ

После новаторской работы Виолы-Джонсона о структуре обнаружения объектов на изображении, многочисленные методы были предложены для обнаружения лица за прошедшее десятилетие. Ранние исследования в научной литературе были в основном направлены на извлечение различных объектов с помощью ручной работы с экспертами в области компьютерного зрения и обучение эффективных классификаторов для распознавания с помощью традиционных алгоритмов МО.

В последние годы методы глубокого обучения, особенно CNN, добился значительных успехов в различных задачах компьютерного зрения, начиная от классификации изображений до обнаружения объектов и семантической сегментации и так далее. В отличие от традиционного компьютерного зрения подходы, методы глубокого обучения избегают ручной разработки и доминируют во многих общеизвестных тестах, таких как ImageNet Large Scale Visual Recognition Challenge. Наряду с популярностью глубокого обучения компьютерному зрению, всплеск исследований все больше внимания уделяется глубокому обучению для решения задач по распознаванию лиц. В общем случае обнаружение лица можно рассматривать как особый тип задачи обнаружения объектов с помощью компьютера. Таким образом, исследователи попытались решить проблему обнаружения лиц, изучив некоторые методы обучения для общих задач обнаружения объектов. Один из очень важных и очень успешных основ для обнаружения общего объекта является метод R-CNN на основе области, который является вид расширения CNN для решения задач обнаружения объектов. Разнообразие последних достижений для лица обнаружение часто следует за этой линией исследования, расширяя R-CNN и его улучшенные варианты.

Распознавание лица можно разделить на две задачи: обнаружение лица и обнаружение ориентиров лица (также называемые ключевые точки). Общий подход для решения задач распознавания лиц состоит в том, чтобы последовательно применять моделей CNN для решения каждой из задач – такой подход носит название многозадачное распознавание лица. Однако создание систем распознавания состоящих из множества последовательно применяемых моделей CNN является сложной задачей, поскольку каждая CNN должна обучается отдельно от других и имеет дело с ошибками, которые были допущены предыдущими моделями.

Недавние исследования показывают, что многозадачные CNN, которые выдают несколько прогнозирующих результатов, могут предложить более высокую точность и лучшую скорость, чем однозадачные CNN. Современные многозадачные CNN основанные на тепловых картах ключевых точек страдают низкой точностью обнаружения лиц, а подходы, основанные на регрессии, имеют худшее обнаружение ориентиров.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка новых и улучшение существующих методов распознавания изображений в рамках решения задачи по распознаванию лиц.

Для достижения поставленной цели необходимо решить следующие основные задачи:

1. Анализ современных архитектур нейронных сетей глубокого обучения и функций потерь.
2. Анализ современных методов распознавания лиц.
3. Разработка нового метода распознавания лиц на основе архитектуры нейронной сети Mask R-CNN для распознавания лиц и выделения их ориентиров.
4. Реализовать ПО для распознавания лиц на основе предложенного метода.
5. Провести экспериментальные исследования разработанного метода распознавания лиц и сравнить с современными аналогами.

Объектом исследования являются системы по распознаванию и выделению контуров (масок) на лицах.

Предметом исследования является математическое и программное обеспечение компьютерных систем для решения задач распознавания лиц.

Основной гипотезой, положенной в основу диссертационной работы, является возможность использования компьютеров общего назначения с ОС Windows или Linux для решения задачи распознавания лиц на изображении и обнаружения ключевых контуров на лицах.

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя Г. Н. Егоровой, заключается в формулировке целей и задач исследования.

Публикации

По теме диссертации опубликовано 3 печатных работ, из них 1 статья в рецензируемом издании, 2 работы в сборниках трудов и материалов международных конференций.

Структура и объем работы

Диссертация состоит из введения, общей характеристики работы, пяти глав, заключения, списка использованных источников, списка пуб-

ликаций автора и приложений. В первой главе представлен анализ предметной области, выявлены основные существующие проблемы в рамках тематики исследования, показаны направления их решения. Вторая глава представляет из себя теоретические сведения на о. Третья глава посвящена разработке нового распознавания лица: выбора структуры нейронной сети, функции активации и потери, разработка алгоритма распознавания. В четвертой главе предложена практическая реализация ПО системы для распознавания лица и его основных контуров. В заключительной пятой главе представлены результаты экспериментальных исследований, анализ характеристик и практического применения разработанного метода.

Общий объем работы составляет 68 страницы, из которых основного текста – 40 страниц, 44 рисунков на 13 страницах, 3 таблиц на 2 страницах, список использованных источников из 70 наименований на 4 страницах и 2 приложения на 8 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** определена область и указаны основные направления исследования, показана актуальность темы диссертационной работы, дана краткая характеристика исследуемых вопросов, обозначена практическая ценность работы.

В **первой главе** описывается то, как развитие подходов для решения задачи распознавания образов на изображении. Описываются различные методы решения задачи в истории, а также способы решения задачи при помощи различных нейронных сетей. Отдельно рассматриваются научные работы, посвященные использованию многозадачных свёрточных нейронных сетей для решения поставленной задачи.

В **второй главе** посвящается теоретическим деталям решения задачи распознавания изображений. В ней описывается теоретическая база нейронных сетей, механизм их работы и обучения. Отдельно рассматривается свёрточные нейронные сети, описывается механизмы их работы и детали реализации и способы применения в решения поставленной задачи по распознаванию лиц и их контуров.

В **третьей главе** делается анализ существующих методов распознавания и выделения контуров лица, рассматривают их достоинства и недостатки. На основе анализа предлагает новый метод на основе нейронной сети Mas R-CNN и описывается общий алгоритм и проводится выбор функции потери для нейронной сети.

В **четвёртой главе** описывается реализация ПО, которое реализует предложенный метод для решения задачи распознавания лица и его контуров. В качестве языка программирования для решения задачи был выбран Python, в качестве библиотеки для создания нейронных сетей была выбрана Tensorflow, для обработки изображений библиотека Pillow. А также приводится архитектура нейронной сети для реализации предложенного метода.

В заключительной пятой главе проводится обучение разработанной нейронной сети и последующее её тестирование ПО. Для обучения происходит выбор подходящих наборов данных, которые бы отражали суть решаемой задачи. На основе анализа полученных результатов делается вывод об эффективности реализованного метода.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Произведён анализ однозадачных CNN на предмет выявления ключевых недостатков, на их основе предложены способы их решения путём создания нового метода распознавания.

2. Предложен метод распознавания лица и его контуров (ключевых точек), который позволяет избавиться от проблем методов, основанных на регрессии (низкая точность обнаружения контуров) и тепловых картах (несоответствие между дискретным пространством тепловых карт и непрерывным расположением ориентиров лица). Результаты тестирования показали эффективность метода и возможность его применения на практике.

3. Предложена архитектура нейронной сети, на базе Mask R-CNN, реализующая предложенный метод. Предложено использовать слой RoIAlign для извлечения объектов, с помощью которого достигается большая точность. В добавок для улучшения распознавания маленьких лиц на изображениях используется: концепция контекстных модулей (помогают увеличивать восприятия признаков лица и усиливают точность слоёв прогнозирования) и пирамида признаков (feature pyramids, для передачи информации в сети с более больших слоёв в более мелкие).

4. Определены оптимальные гиперпараметры для обучения нейронной сети.

5. Проведена подготовка входных наборов данных для обучения нейронной сети.

6. Проведён анализ определения оптимальной функции потерь для нейронной сети.

7. Разработано ПО на языке Python с использованием современных библиотек для МО которые помогают в реализации предложенной архитектуры нейронной сети.

Рекомендации по практическому использованию результатов

1. Полученные результаты формируют теоретическую и практическую базу для использования метода распознавания лиц и их ключевых контуров при решении задач в системах компьютерного зрения. Они могут быть использованы для модернизации и дальнейшего развития существующих систем.

2. Разработанные метод распознавания лица и архитектура нейронной сети может применяться в системах компьютерного зрения для повышения качества распознавания лиц и дальнейшей их локализации в зависимости от поставленной задачи.

3. Разработанное ПО может применяться для решения задач распознавания лиц и их последующей локализации в системах компьютерных зрения.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

1-А. Папроцкий Э.В. Оценка влияния позиционирования изображений в свёрточных нейронных сетях для решения задачи сегментации / Э.В. Папроцкий // Научное сообщество студентов XXI столетия. ТЕХНИЧЕСКИЕ НАУКИ. – Россия, Москва: 2020. – № 6(89). – с. 67–71.

2-А. Папроцкий Э.В. Исследование эффективности выбора ФП в свёрточных нейронных сетях для решения проблемы переобучения в системах распознавания лиц / Э.В. Папроцкий // LXXXVII Молодёжный научный форум. – Россия, Москва: 2020 – № 9(91). – с. 119–123.

3-А. Папроцкий Э.В., Подбор учебных программ в образовательных системах с использованием нейронных сетей глубокого обучения / Э.В. Папроцкий // Компьютерные системы и сети: материалы XI научно-методической конференции. – Минск: БГУИР, 2020. – с. 228-229.