

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

УДК 004.3.049.77

ДАЦИК
Антон Андреевич

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ

АВТОРЕФЕРАТ

диссертации на соискание степени
магистра информатики и вычислительной техники

по специальности 1-40 81 04 – Обработка больших объемов информации

Минск 2020

Работа выполнена на кафедре информатики учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель: **АНИСИМОВ Владимир Яковлевич**,
кандидат физико-математических наук, доцент
кафедры информатики «Белорусский
государственный университет информатики и
радиоэлектроники»

Рецензент: **БУЛОВА Александр Дмитриевич**,
кандидат технических наук, доцент кафедры
экономической информатики «Белорусский
государственный экономический университет»

Защита диссертации состоится «24» июня 2020 г. года в 10⁰⁰ часов на заседании Государственной экзаменационной комиссии по защите магистерских диссертаций в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220013, Минск, ул. Гикало, 9, копр. 4, ауд. 111, тел. 293-85-91, e-mail: inform@bsuir.by

С диссертацией можно ознакомиться в библиотеке учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

ВВЕДЕНИЕ

В настоящее время всё более широкое распространение получают биометрические системы идентификации человека. Традиционные системы идентификации требуют знания пароля, наличия ключа, идентификационной карточки либо иного идентифицирующего документа, который можно забыть, потерять или подделать. В отличие от них биометрические системы основываются на уникальных биологических характеристиках человека, которые трудно подделать и которые однозначно определяют конкретного человека. К таким характеристикам относятся отпечатки пальцев, форма ладони, узор радужной оболочки, изображение сетчатки глаза. Лицо, голос и запах каждого человека также индивидуальны.

Распознавание человека по изображению лица выделяется среди биометрических систем тем, что, во-первых, не требует специального дорогостоящего оборудования. Для большинства приложений достаточно персонального компьютера и обычной видеокамеры. Во-вторых, отсутствует физический контакт человека с устройствами. Не надо ни к чему прикасаться или специально останавливаться и ждать срабатывания системы. В большинстве случаев достаточно просто пройти мимо или задержаться перед камерой на несколько секунд.

К недостаткам распознавания человека по изображению лица следует отнести то, что сама по себе такая система не обеспечивает 100%-ной надёжности идентификации. Там, где требуется высокая надёжность, применяют комбинирование нескольких биометрических методов.

На данный момент проблеме распознавания человека по изображению лица посвящено множество работ, однако в целом она ещё далека от разрешения. Основные трудности состоят в том, чтобы распознать человека по изображению лица независимо от изменения ракурса и условий освещённости при съёмке, а также при различных изменениях, связанных с возрастом, причёской и т.д.

Распознавание изображений пересекается с распознаванием образов. Такие задачи не имеют точного аналитического решения. При этом требуется выделение ключевых признаков, характеризующих зрительный образ, определение относительной важности признаков путём выбора их весовых коэффициентов и учёт взаимосвязей между признаками.

Изначально эти задачи выполнялись человеком-экспертом вручную, путём экспериментов, что занимало много времени и не гарантировало качества. В новых методах выделение ключевых признаков осуществляется путём автоматического анализа обучающей выборки, но тем не менее большая часть информации о признаках задаётся вручную. Для автоматического применения таких анализаторов выборка должна быть достаточно большой и охватывать все возможные ситуации (например, изменения ракурса, внешности, условий освещённости и т.п.).

Нейросетевые методы предлагают иной подход к решению задачи распознавания образов. Архитектура и функционирование нейронных сетей (НС) имеют биологические прообразы. Веса в нейронной сети не вычисляются путём решения аналитических уравнений, а подстраиваются различными локальными методами (например, разновидностями градиентного спуска) при обучении. Обучаются нейронные сети на наборе обучающих примеров. В процессе обучения НС происходит автоматическое извлечение ключевых признаков, определение их важности и построение взаимосвязей между ними. Обученная НС может успешно применять опыт, полученный в процессе обучения, на неизвестные образы за счёт хороших обобщающих способностей.

Таким образом, применение нейронных сетей для задачи распознавания человека по изображению лица является перспективным направлением, на что и направлено основное внимание данной работы.

В данной работе описаны методы распознавания человека по изображению лица и возможности использования их совместно с нейросетевыми методами. Дан обзор нейросетевых методов распознавания изображений и их применения к распознаванию человека по изображению лица. Описаны различные способы представления изображения и учёта его свойств.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

Магистерская диссертация относится к актуальной области современных информационных технологий – использованию нейронных сетей для решения задачи распознавания лиц. Актуальность работы обусловлена тем, что в настоящее время всё более широкое распространение получают биометрические системы идентификации человека и именно использование нейронных сетей для решения данной задачи является наиболее перспективным направлением, на что и направлено основное внимание данной работы.

Результаты работы носят как теоретический (комплекс рекомендаций, методик и т.п.), так и практический характер: предложенная композиция из нейронных сетей может применяться в системах биометрической идентификации, системах детекции с идентификацией личности, системах анализа видеоконтента за счет надежности, а также высокой скорости идентификации.

Степень разработанности проблемы

Исследование задачи распознавания человека по изображению лица с использованием нейронных сетей осуществлялось на основе работ российских и белорусских ученых: А. Глазунова, Д. Самалю, Г. Вороновского, А. Скурихина, В. Базаревского, а также зарубежных авторов: Ш. Панканти, Э. Джейна, А. Ранга, П. Пернера, М. Зейлера и др.

Одним из недостатков исследований, представленных в современной технической литературе, является рассмотрение отдельных частей задачи распознавания лиц, т.е. поиска лица на изображении и идентификации человека.

Предложенная работа направлена на устранение этого недостатка, в работе целостно рассмотрена задача распознавания лица на изображении, а также проведены исследования различных архитектур нейронных сетей для достижения необходимой точности и производительности при реализации системы распознавания лиц.

Цель и задачи исследования

Целью диссертационной работы является изучение алгоритмов обучения и архитектур нейронных сетей для распознавания лиц на изображениях, а также разработка программного обеспечения для решения задач обнаружения лиц и дальнейшего распознавания в режиме реального времени.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Исследовать современные архитектуры нейронных сетей;
2. На основании произведенного анализа выбрать необходимые технологии и инструменты для реализации поставленной задачи;
3. Спроектировать и разработать программный продукт для распознавания лиц на изображениях и видео;
4. Провести экспериментальные исследования разработанного программного продукта.

Область исследования

Содержание диссертации соответствует образовательному стандарту высшего образования второй ступени (магистратуры) ОСВО 1-39 81 01-2012 специальности 1-40 81 04 «Обработка больших объемов информации».

Теоретическая и методологическая основа исследования

В основу диссертации легли работы белорусских и зарубежных ученых в области использования глубоких нейронных сетей для решения задач компьютерного зрения.

Информационная база исследования сформирована на основе литературы, открытой информации, технических нормативно-правовых актов, сведений из электронных ресурсов, а также материалов научных конференций и семинаров.

Научная новизна

Научная новизна и значимость полученных результатов работы заключается в исследовании архитектур глубоких нейронных сетей для

решения задачи распознавания лиц, в предложенной композиции нейронных сетей для решения данной задачи, а также разработке мобильного приложения, использующего композицию из BlazeFace и FaceNet сетей.

Теоретическая значимость работы заключается в детальном анализе архитектур и подходов к использованию нейронных сетей для распознавания лиц.

Практическая значимость диссертации состоит в проведенных исследованиях различных архитектур, также предложенная композиция из нейронных сетей может применяться в системах биометрической идентификации, системах детекции с идентификацией личности, системах анализа видеоконтента за счет надежности, а также высокой скорости идентификации.

Основные положения, выносимые на защиту

1. Систематизация основных классов решаемых задач в распознавании человека по изображению лица, преимущества нейросетевых методов распознавания лица человека на изображении.

2. Архитектуры сверточных нейронных сетей для решения задачи распознавания лица. Анализ их точности и производительности работы.

3. Обзор TensorFlow Lite для конвертации моделей сверточных нейронных сетей в формат, который может быть использован на мобильных устройствах, а также разработка экспериментального мобильного приложения, которое использует композицию из BlazeFace и FaceNet.

Апробация диссертации и информация об использовании ее результатов

Основные положения диссертационной работы использовались в разработке системы идентификации работников в ИУП «Адмитад».

Публикации

Результаты исследований, вошедшие в диссертацию, были опубликованы в статье в рецензируемом издании: Дацик А.А., Стасюк А.В. Алгоритмы обнаружения объектов. Образование и наука в России и за рубежом выпуск N15(2019) Vol.63.

Структура и объем работы

Диссертация состоит из введения, общей характеристики работы, четырех глав с краткими выводами по каждой главе, заключения, библиографического списка и приложений.

В первой главе приведен обзор основных классов решаемых задач в распознавании человека по изображению лица.

Во второй главе рассмотрены нейросетевые методы распознавания человека на изображении, а также рассмотрены различные типы искусственных нейронных сетей.

В третьей главе рассмотрены архитектуры сверточных нейронных сетей для решения задачи распознавания лица, также проведены эксперименты для сравнения точности и производительности разных архитектур.

В четвертой главе приведен обзор средств для использования сверточных сетей на мобильных устройствах. А также приведены результаты разработки экспериментального мобильного приложения, работающего с помощью BlazeFace и FaceNet сетей.

В приложении представлены публикации автора и результаты исследований производительности и точности сетей.

Общий объем работы составляет 53 страницы, из которых основного текста – 40 страниц, 20 рисунков, 7 таблиц, список использованных источников из 22 наименований на 2 страницах и 3 приложения на 8 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** рассмотрено современное состояние биометрических систем идентификации, приведены преимущества распознавания человека по изображению лица перед традиционными способами идентификации. Рассмотрены основные плюсы нейросетевых методов для решения задач распознавания образов.

В общей характеристике работы показана актуальность проводимых исследований, степень разработанности проблемы, сформулированы цель и задачи диссертации, обозначена область исследований, научная (теоретическая и практическая) значимость исследований, а также апробация работы.

В первой главе приведен обзор задач распознавания человека на изображениях лица.

Из анализа следует, что задачи распознавания человека на изображении лица делятся на три больших класса: поиск в больших базах данных, контроль доступа и контроль фотографий в документах. Они различаются как по требованиям, предоставляемым к системам распознавания, так и по способам решения, и поэтому представляют собой отдельные классы.

Также можно сделать вывод, что различны и требования, предъявляемые к ошибкам первого и второго рода для таких классов. Ошибкой первого рода (type I error, misdetection) называется ситуация, когда объект заданного класса не распознаётся (пропускается) системой. Ошибка второго рода (type II error, false alarm) происходит, когда объект заданного класса принимается за объект другого класса.

В данной главе отмечается различие понятий верификации и распознавания (идентификации). В задаче верификации неизвестный объект заявляет, что он принадлежит к некоторому известному системе классов. Система подтверждает или опровергает это заявление. В системах верификации ошибкой первого рода является ситуация, когда объект, принадлежащий к известным системе классам, принимается за объект, относящийся к неизвестным системе классам, и в доступе ему отказывают. Ошибка второго рода совершается, когда объект неизвестного класса принимается за объект, относящийся к известным системе классам, и ему разрешается доступ. При распознавании требуется отнести объект к одному из n известных классов или выдать заключение о том, что этот объект не относится к известным классам.

Во второй главе рассмотрены нейросетевые методы распознавания лица человека.

Приведены основные задачи, решаемые при помощи нейронных сетей:

- разбиение пространства признаков на области, соответствующие классам (классификация, распознавание, кластеризация);
- извлечение ключевых характеристик, сжатие и реконструкция образов;
- аппроксимация функции многих переменных с любой заданной точностью;
- прогнозирование временных рядов;
- решение оптимизационно-комбинаторных задач;
- распознавание с учётом топологии пространства.

Большинство из этих задач прямо или косвенно связаны с распознаванием изображений.

Также во второй главе рассмотрена архитектура многослойной нейронной сети (рисунок 1).

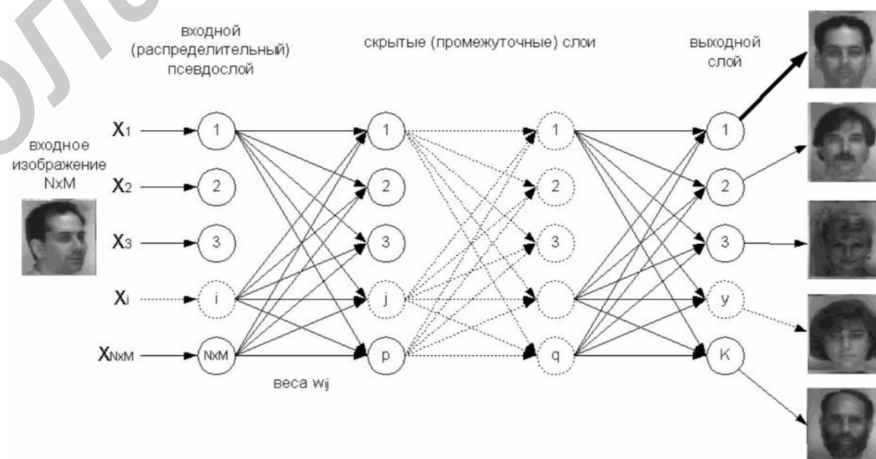


Рисунок 1 – Архитектура многослойной нейронной сети и её применение для распознавания изображений. Нейрон с максимальной активностью (здесь первый) указывает принадлежность к распознанному классу

Рассмотрены следующие проблемы многослойных нейронных сетей:

- проблема локального минимума;
- выбор архитектуры сети (количество нейронов, слоев, характер связей);
- выбор шага (скорости) обучения.

Проведен анализ генетического алгоритма как одного из наиболее перспективных методов, применяемых на этапе обучения нейронной сети.

Рассмотрены типы нейронных сетей позволяют учесть топологию пространства изображения. Ведь изображение – это не просто n -мерный вектор, составленный из яркостей пикселей. Изображение имеет топологию, определяемую через двумерное локальное соседство пикселей.

Принципы работы таких сетей основываются на разбиении изображения на маленькие участки и иерархическом сопоставлении как взаимного их расположения, так и содержания.

Такие сети являются наиболее перспективными для распознавания изображений.

В третьей главе была рассмотрена задача распознавания лиц с более практической точки зрения. Как уже было упомянуто ранее распознавание лиц представляет собой фактически последовательность нескольких связанных проблем:

1. Рассмотреть изображение и найти на нём все лица.
2. Преобразование найденных лиц в тензор признаков.
3. Поиск меры схожести между лицами в некой базе данных и найденными на фотографии или видео.

Компьютеры неспособны к такого рода высокому уровню обобщения, поэтому приходится учить их каждому шагу в этом процессе отдельно.

Необходимо построить конвейер, с помощью которого можно будет находить решение на каждом шаге процесса распознавания лица отдельно и передавать результат текущего шага на следующий. Другими словами, соединить несколько алгоритмов обучения машины в одну цепь.

Таким образом, вначале были рассмотрены современные архитектуры сверточных нейронных сетей для поиска лица на изображении, затем рассмотрена FaceNet, как кандидат на роль преобразователя изображения лица в тензор признаков.

Проведены исследования, чтобы понять насколько эти сети быстро работают на CPU 2,5 GHz Intel Core i7 и с какой точностью определяют положения лица на изображении. Также были проведены аналогичные эксперименты на GPU Tesla P100-PCI-E-16GB.

В данной главе были рассмотрены следующие архитектуры сверточных нейронных сетей: Yolo v3, MobileNetV2, MTCCN и BlazeFace, как наиболее новая и оптимизированная для работы на мобильных устройствах.

Были приведены отличия типичной SSD модели от SSD (рисунок 2).

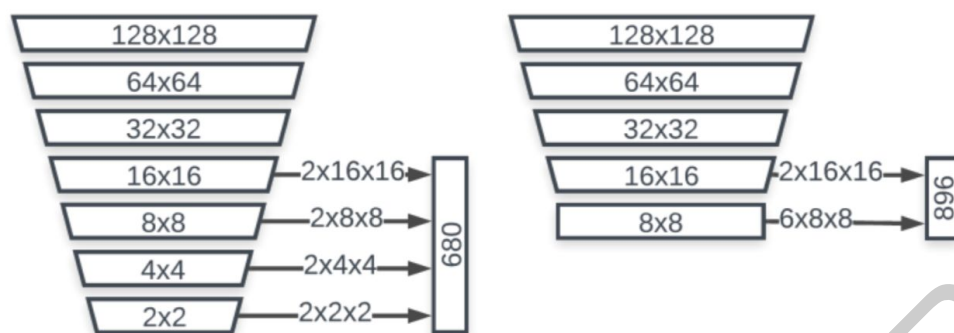


Рисунок 2 – Поиск якорей в SSD (слева) и BlazeFace (справа)

В плане производительности BlazeFace сеть превзошла Yolo, MTCCN, Mobilenet (таблица 1).

Таблица 1 – Результаты замеров времени, затрачиваемого на поиск лица каждой сетью на одном фрейме

Нейронная сеть	Время поиска (сек)
MTCCN	0.475
Yolo	0.232
MobileNet	0.164
BlazeFace	0.012

Далее была рассмотрена архитектура сети FaceNet. Задачей данной сети является преобразование найденного изображения лица в вектор признаков для дальнейшего сравнения его с уже занесенными в базу данных лицами.

Особенностью FaceNet является ее функция потерь называемая TripletLoss. Основная задача данного метода – уместить лица в пространстве размерности R^d , в котором будет минимальное расстояние между лицами одной личности и максимальное – между лицами из разных классов (рисунок 3).

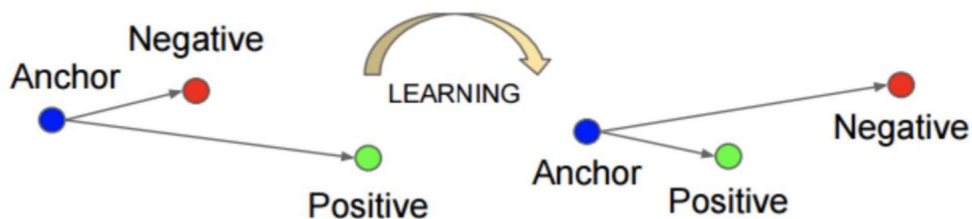


Рисунок 3 – Работа TripletLoss

В четвертой главе было разработано экспериментальное мобильное приложения под ОС Android. Интерес представляет, как эти сети будут работать на не флагманских мобильных устройствах, а также с какой скоростью приложение будет распознавать лица людей, которые добавлены в базу данных.

Также в данной главе был изучен TensorFlow Lite как инструмент, который позволяет исполнение модели на различных устройствах, в том числе мобильных телефонов и интернета вещей, что дает вам более чем 3-кратное ускорение вывода по сравнению с оригинальным TensorFlow. С помощью него можно получить машинное обучение на своем Raspberry Pi или на телефоне.

Связка из BlazeFace для определения положения лица и FaceNet для дальнейшего его распознавания показала достаточно хорошие результаты в плане производительности, а также точности.

Ниже показана производительность работы BlazeFace и FaceNet в зависимости от разрешения видео файла на мобильном устройстве с процессором Qualcomm Snapdragon 625 MSM8953 с тактовой частотой 2000 МГц, а также с графическим ускорителем с частотой 650 МГц (таблица 2).

Таблица 2 – Производительность работы BlazeFace и FaceNet

FPS (1080x1920)	FPS (720x1280)	FPS (540x960)
10.97	13.32	20.50

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Дан обзор нейросетевых методов распознавания человека на изображении. Указаны преимущества и перспективы нейросетевых методов. Проведены исследования различных алгоритмов обучения, предобработки изображений, архитектур нейронных сетей, которые могут использоваться для решения задачи распознавания лица человека на изображении. Проведен сравнительный анализ производительности и точности распознавания нескольких современных архитектур нейронных сетей. Отмечены наиболее перспективные архитектуры для данной задачи.

2. Предложена композиция из нейронных сетей, которая способна работать на различных устройствах, в том числе мобильных телефонах и интернета вещей. Данный способ обработки изображений позволяет получить достаточно большой прирост в производительности даже на относительно слабых устройствах. На флагманских моделях мобильных телефонах процедура распознавания лица человека занимает несколько миллисекунд и отлично работает в режиме реального времени.

3. Разработан экспериментальный программный продукт под ОС Android с использованием композиции из оптимизированных моделей нейронных сетей.

Рекомендации по практическому использованию результатов

1. Полученные результаты формируют теоретическую и практическую базу для разработки ПО компьютерных систем для решения задач поиска человека в большой базе данных, контроля доступа, контроля фотографии в документах.

2. Предложенная композиция из нейронных сетей может применяться в системах биометрической идентификации, системах детекции с идентификацией личности, системах анализа видеоконтента за счет надежности, а также высокой скорости идентификации.

3. Результаты работы могут использоваться при подготовке персонала для разработки и обслуживания компьютерных систем, решающих задачи детекции и(или) распознавания человека на изображении или видео.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ

Статьи в рецензируемых журналах

1. Дацик А.А., Стасюк А.В. Алгоритмы обнаружения объектов. Образование и наука в России и за рубежом выпуск N15(2019) Vol.63.

Библиотека БГУИР