

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**Построение и обучение свёрточной нейронной сети для задачи
определения пола и возраста человека по фотографии**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 441 группы
направления 02.03.03 Математическое обеспечение и администрирование
информационных систем
факультета компьютерных наук и информационных технологий
Фёдорова Андрея Алексеевича

Научный руководитель:

ассистент кафедры ИиП

А.А. Казачкова

подпись, дата

Зав. Кафедрой ИиП:

доцент

М.В. Огнева

подпись, дата

Саратов 2020

ВВЕДЕНИЕ

Актуальность темы. Нейронные сети, или точнее, искусственные нейронные сети, представляют собой технологию, имеющую истоки во множестве дисциплин: математике, компьютерных науках, статистике, нейрофизиологии, физике и технике. Они сохраняют актуальность и повсеместно используются в наши дни, находя своё применение в таких разнородных областях, как моделирование, анализ временных рядов, распознавание образов, обработка сигналов и управление благодаря одной важной особенности – умении обучаться на основе данных при участии учителя или без его вмешательства.

Проводится множество исследований по искусственным нейронным сетям, и связано это с тем, что способ обработки информации человеческим мозгом в корне отличается от алгоритмов, используемых обычными цифровыми компьютерами. Можно сказать, что мозг – это весьма непростой, нелинейный, параллельный компьютер (система обработки информации). У мозга есть возможность формировать свои структурные компоненты, которые называются *нейронами*, таким образом, чтобы они могли исполнять конкретные задачи (например, распознавание образов, обработку сигналов органов чувств, моторные функции) на порядок быстрее, чем с этими задачами могут справиться самые быстродействующие современные компьютеры. Примером такой задачи обработки информации может быть обычное зрение. В функции зрительной системы входит создание представления окружающего мира в виде, обеспечивающего возможность взаимодействия с этим миром. Мозг поочередно исполняет ряд задач распознавания (например, распознавание ранее знакомого лица в незнакомом окружении). На выполнение такого ряда задач мозг тратит 100-200 миллисекунд, в то время как на выполнение подобных задач даже меньшей сложности компьютер может потратить несколько дней.

В данной работе в распознавании образов в качестве образа выступает изображение человека. Говоря об актуальности использования нейронных

сетей, стоит отметить, что в настоящее время существует несколько популярных сервисов, использующих нейронные сети, для нахождения на изображении лица человека, определении возраста, пола, национальности человека.

Одним из таких сервисов является how-old.net – сервис от Microsoft, который определяет пол и возраст по фотографии человека. Демонстрация сервиса прошла в рамках конференции Build 2015. После этого how-old.net за несколько часов стал очень популярен в сети Интернет, набрав более 40 тысяч пользователей, загрузивших свыше 250 тысяч фотографий. В основе сервиса лежит система машинного обучения и распознавания лиц, а также поисковые инструменты Bing. Для использования сервиса надо загрузить фотографию на сайт, и система почти сразу выдаст предполагаемый пол и возраст изображённых на снимке людей.

FaceApp – мобильное приложение для iOS и Android, разработанное российской компанией Wireless Lab, использующей технологию нейронной сети для генерации очень реалистичных преобразований лиц на фотографиях. Приложение может видоизменять лицо, сделать его улыбающимся или более агрессивным, омолодить или состарить и даже изменить пол человека на фотографии. FaceApp был запущен на iOS в январе 2017 года и на Android в феврале 2017 года.

В сентябре 2019 года стало популярным мобильное приложение Gradient. Оно определяет пол человека на фотографии и на кого он похож из знаменитых людей. Также оно в процентном соотношении определяет черты какой национальности более присущи человеку на фотографии. Правда неизвестно насколько широко приложение использует именно нейронные сети и глубокое обучение, так как для одной и той же фотографии могут быть выданы совершенно разные результаты. Популярность приложения скорее объясняется интересной идеей, а не сложностью реализации.

Получившееся в результате данной работы приложение схоже с вышеприведёнными сервисами. В нём, как и во многих из них, будут использованы свёрточные нейронные сети для обработки изображения.

Цель бакалаврской работы – изучение искусственных нейронных сетей, их возможностей и методов обучения, создание, обучение моделей свёрточной нейронной сети для определения пола и возрастной категории человека по фотографии, а также использование обученных моделей с помощью пользовательского интерфейса.

Поставленная цель определила **следующие задачи**:

1. Изучение общего принципа работы нейронной сети и в частности принципа работы свёрточной нейронной сети
2. Отбор, формирование и распределение данных для обучения нейронной сети
3. Формирование архитектур для двух моделей свёрточной нейронной сети – для распознавания пола и возрастной категории человека по фотографии
4. Обучение моделей нейронной сети на отобранных данных для распознавания по фотографии человека его пола и возрастной категории
5. Визуализация процесса обучения моделей
6. Применение полученных моделей нейронных сетей
7. Создание пользовательского интерфейса для использования моделей

Методологические основы построения и обучения свёрточной нейронной сети для задачи определения пола и возраста человека по фотографии представлены в работах С. Хайкина, Ф. Уоссермена, Я. Гудфеллоу, И. Бенджио, А. Курвилль, Т. Рашида, С. Николенко, А. Кадурина, Е. Архангельской

Практическая значимость бакалаврской работы. На основе свёрточных нейронных сетей было создано приложение, распознающее пол и возрастную категорию человека по фотографии и позволяющее показать статистику по распознанным людям на нескольких фотографиях. Такой тип распознавания людей может использоваться для различного рода

статистического анализа. Например, для сбора информации целевой аудитории для рекламодателей, так как на некоторых рекламных экранах, мониторах на улицах города или в общественном транспорте устанавливаются камеры, которые фиксируют лица людей, смотревших на экран с рекламой какое-то время. Автоматически распознав пол и возрастную категорию этих людей и подсчитав статистику, рекламодатели могут понять какие именно люди больше всего заинтересованы в их товарах или услугах.

Структура и объём работы. Бакалаврская работа состоит из введения, 3-ёх разделов, заключения, списка использованных источников и 8 приложений. Общий объём работы – 92 страницы, из них 65 страниц – основное содержание, включая 23 рисунка, список использованных источников информации – 36 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Нейронные сети» посвящен общим понятиям, связанным с нейронными сетями, здесь описан общий принцип работы классической нейронной сети. Также в этом разделе приведена классификация нейронных сетей по различным признакам. Здесь же подробно описан принцип работы свёрточной нейронной сети, используемой для реализации распознавания лиц людей на изображениях в бакалаврской работе.

Примечание [LAK1]: связанным

Также в этом разделе описаны функции активации, которые нужны для того, чтобы активировать нейроны, проходящие некоторое пороговое значение, отсеивая тем самым «шумы». Приведены уравнения, графики, описания, преимущества и недостатки таких функций активации как ступенчатая, линейная, сигмоидальная, гиперболический тангенс, ReLu (rectified linear unit).

Далее идёт речь об обучении нейронной сети как в глобальном смысле, так и про метод градиентного спуска и метод обратного распространения ошибки. Суть метода градиентного спуска заключается в нахождении локального минимума и максимума функции потерь для минимизации этой потери. Суть метода обратного распространения ошибки в том, чтобы после того как был получен результат на выходном слое сети, вычислялась ошибка и на её основании выполнялась бы обратная передача информации к входному слою с целью корректировки весов нейронов.

Далее описаны сами функции потерь, которые используются для расчёта ошибки между реальным и полученным результатом. Приведены формулы и описания следующих функций потерь: средняя квадратичная ошибка, бинарная кросс-энтропия и категориальная кросс-энтропия.

В качестве вывода описаны достоинства нейронных сетей, среди которых нелинейность, отображение входной информации в выходную, адаптивность, очевидность ответа, контекстная информация, отказоустойчивость, масштабируемость и др.

Примечание [LAK2]: «отображение входной информации в выходную» звучит не очень круто, все так делают! Может имелась в виду «нелинейность отображения»?

Второй раздел «Распознавание образов» посвящен распознаванию образов в глобальном смысле и различным инструментам, библиотекам, фреймворкам, которые можно использовать для распознавания образов.

В глобальном смысле описывается принцип машинного распознавания образов, также отмечается, что распознавание образов с помощью нейронных сетей схоже с тем, распознаёт образы человеческого мозга.

Примечание [LAK3]: тут, кажется, пропущено слово "как"

Из инструментов и средств для распознавания образов приведены OpenCV – это библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом; PyTorch – современная библиотека для глубокого обучения, разработанный для сервисов Facebook; Tensorflow — довольно молодой фреймворк для глубокого машинного обучения, разрабатываемый в Google Brain; Keras – реализация TensorFlow, высокоуровневый API для построения и обучения моделей, включающий поддержку для TensorFlow-специфичной функциональности.

Примечание [LAK4]: библиотека разработанный

В бакалаврской работе для разработки моделей нейронной сети используется именно Keras. Так как OpenCV – это библиотека именно компьютерного зрения, и для глубокого обучения и создания собственной модели нейронной сети Keras подходит больше. Фреймворку PyTorch недостаёт интерфейсов для мониторинга и визуализации. И Keras, как надстройка над TensorFlow, имеет более «дружественный» интерфейс, чем PyTorch и тот же TensorFlow. Также приведён график популярности самых распространённых инструментов для работы с нейронными сетями, на котором видно, что Keras является самым востребованным.

Далее приведён часто используемый пример типичной архитектуры свёрточной нейронной сети, состоящей из слоёв, каждый из которых подробно описан далее.

Третий раздел «Применение искусственных нейронных сетей для распознавания образов» посвящён реализации двух архитектур моделей свёрточной нейронной сети для определения пола и возрастной категории.

Примечание [LAK5]: тут надо очередной раз повторить "людей на изображениях", а то мало ли

Модели обучались на изображениях, которые по скрипту сохранялись с сайта, на **котором сетей каждый раз** при обновлении страницы генерируется изображение несуществующего человека.

Примечание [LAK6]: наверное "сетью генерируется" или "сеть генерирует"

После разделения входных данных на обучающую, проверочную и тестовую выборки были построены архитектуры моделей. Для получения лучшего результата (точности модели) при построении архитектуры первой модели для определения пола подбирались **нужных** слои и более подходящие **гиперпараметры в этих слоях**. При применении этой же архитектуры к модели для определения возрастной категории точность модели получилась значительно хуже, поэтому было уменьшено количество некоторых слоёв, изменены гиперпараметры и, конечно же, функция активации.

Примечание [LAK7]: нужныЕ

Примечание [LAK8]: лучше просто "гиперпараметры", не все же они относятся к отдельным слоям, есть и общие для модели в целом

Далее был визуализирован процесс обучения обеих моделей с помощью графиков, на которых видно, как в процессе обучения увеличивалась точность и уменьшались потери моделей.

С помощью таких расширения языка Python как PySide 2 и PyQt был сконструирован пользовательский интерфейс приложения, которое демонстрирует работу обеих моделей. Для этого в логику интерфейса были добавлены возможности для загрузки ранее сохранённых моделей и для предсказания пола и возрастной категории человека по фотографии. Также в приложение была добавлена возможность отобразить статистику по распознанным людям на нескольких фотографиях, показывающую процентное соотношение между полами и возрастными категориями.

ЗАКЛЮЧЕНИЕ

В данной работе были изучены типы и принцип работы искусственных нейронных сетей, а также принцип распознавания образов на изображении с помощью свёрточной нейронной сети. Были изучены различные виды функций активации нейрона и функций потерь нейронной сети, а также способ обучения нейронной сети с помощью градиентного спуска и метода обратного распространения. Были рассмотрены основные актуальные библиотеки и фреймворки, позволяющие создавать и обучать нейронные сети для решения задачи распознавания образов.

Примечание [LAK9]: обратного распространения чего?

Модели нейронной сети были обучены на заранее отобранных и распределённых данных для распознавания по фотографии человека его пола и возрастной категории, процесс обучения был визуализирован с помощью графиков потерь и точности. Для использования обученных моделей разработано графическое приложение. Также обученные модели можно определённым образом использовать и для статистического анализа, результаты которого в дальнейшем можно применять в различных целях, например, в таргетированной онлайн-рекламе (реклама, направленная на определённую целевую аудиторию), которая отображалась бы для людей, принадлежащих одной категории, определяемой по полу или возрасту. Одновременно с этим можно дообучать модели на уже новых полученных данных, как это было сказано в работе.

Примечание [LAK10]: размеченных

Основные источники информации:

- Хайкин, Саймон. Нейронные сети: полный курс, 2 издание / С. Хайкин. Пер. с англ. — М.: Издательский дом «Вильямс», 2006. — 1104 с.: ил. — Парал. тит. англ
- Уоссермен, Филипп. Нейрокомпьютерная техника : Теория и практика / Ф. Уоссермен. Перевод с англ. Ю. А. Зуева, В. А. Точенова; Под ред. А. И. Галушкина. - М. : Мир, 1992. - 184 с
- Гудфеллоу Я., Бенджио И., Курвилль А. Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.: цв. ил.
- Рашид, Тарик. Создаём нейронную сеть / Т. Рашид. Пер. с англ. — СПб. : ООО “Альфа-книга”, 2017. — 272 с.: ил. — Парал. тит. англ.
- С. Николенко, А. Кадури, Е. Архангельская. Глубокое обучение. — СПб.: Питер, 2018. — 480 с.: ил. — (Серия «Библиотека программиста»)