

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра дискретной математики и информационных технологий

**РАЗРАБОТКА WEB-ПРИЛОЖЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ СХОДСТВА
ЛЮДЕЙ ПО ФОТОГРАФИЯМ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 421 группы
направления 09.03.01 — Информатика и вычислительная техника
факультета КНиИТ
Кузьминой Виктории Андреевны

Научный руководитель
ассистент

А. А. Трунов

Заведующий кафедрой
доцент, к. ф.-м. н.

Л. Б. Тяпаев

Саратов 2023

ВВЕДЕНИЕ

С развитием информационных технологий человечество стремится разработать устройства, способные функционировать, интерпретируя человеческую деятельность. Еще в прошлом столетии ученые грезилы идеей замены людей роботами, способными выполнять различные задачи. В 1950-х годах ученые, занимающиеся классическим искусственным интеллектом, основанным на логике и графах, искали новые сферы для его применения. В то же время появилось новое течение в компьютерной науке, сторонники которого считали, что для решения компьютерами сложных задач одной логики недостаточно. Необходимо было приблизиться к функционированию мозга и тем самым сделать системы способными программировать самих себя, опираясь на механизмы обучения мозга [1]. Вместо простого воспроизведения логической цепочки рассуждений человека были предложены попытки смоделировать непосредственно сам носитель данной логики, природный процессор, которым является мозг человека. С появлением этой идеи первопроходцы в машинном обучении стали работать над созданием искусственных нейронов — своеобразной архитектуры сети математических функций.

Сейчас развитие информационных технологий поражает своими стремительными темпами. Это позволило им охватить большинство сфер человеческой деятельности. В наши дни настоящим трендом стали нейронные сети, окружающие человека повсюду, и все, что с ними связано. Например, в индустрии развлечений нейро-сетевые технологии используются при создании анимации, компьютерной графики, а также звуковых эффектов. Нейронные сети способны выполнять сложные вычисления и моделирование различных процессов, что, бесспорно, находит свое применение в научной отрасли. Более того, компьютерное зрение стоит на страже безопасности наших данных и, более того, нашего здоровья и жизней. Примерами такому применению нейронных сетей служат разблокировка мобильного устройства при помощи лица пользователя, распознавание злокачественных новообразований или система автоматической парковки автомобилей.

Актуальность разрабатываемого приложения может быть обусловлена цифровизацией различных слоев общества, катализатором которой стал COVID-19, в том числе старшего поколения, представители которого часто пребывают в поиске своих старых знакомых, сослуживцев или же просто

друзей детства.

Целью данной выпускной квалификационной работы является разработка web-приложения для определения сходства людей по фотографиям.

Для достижения поставленной цели необходимо выполнение следующих задач:

- систематизация знаний об общих принципах функционирования нейронных сетей, их характеристиках и особенностях;
- сбор и систематизация знаний о программном обеспечении для разработки нейронной сети;
- сбор и систематизация знаний о программном обеспечении для разработки web-приложения;
- выбор наиболее оптимальных программных средств для реализации;
- программная разработка нейронной сети, позволяющей сравнивать лица людей;
- разработка web-приложения для сравнения двух фотографий, загруженных пользователем;
- интеграция кода нейронной сети в разрабатываемое web-приложение.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Выпускная квалификационная работа состоит из трех глав.

В первой главе раскрывается тема искусственного интеллекта и рассматриваются общие принципы функционирования нейронных сетей, их характеристики и особенности.

Искусственный интеллект — это научная область, которая занимается разработкой интеллектуальных компьютерных систем, то есть систем, обладающих возможностями, которые мы традиционно связываем с человеческим разумом, — понимание языка, обучение, способность рассуждать, решать проблемы и так далее [2]. Искусственный интеллект охватывает различные области научной деятельности.

Разделом искусственного интеллекта является машинное обучение. Данный раздел является очень важным, но далеко не единственным. Он представляет собой науку и искусство программирования компьютеров для того, чтобы они были способны обучаться решению различных задач на основе данных.

В 1997 году американский ученый, профессор Университета Карнеги-Меллон, основатель первой в мире кафедры машинного обучения и автор первого учебника по этому предмету, дал формальное определение алгоритмов в области машинного обучения. Звучит оно так: «Говорят, что компьютерная программа учится на опыте E в отношении некоторого класса задач T и показателя производительности P , если ее производительность по задачам T , измеренная P , улучшается с опытом E ».

В свою очередь машинное обучение включает в себя множество других методов обучения.

Все задачи, которые решаются при помощи машинного обучения, можно отнести к различным типам. К наиболее распространенным относятся следующие:

- задача регрессии;
- задача классификации;
- задача кластеризации;
- задача уменьшения размерности;
- задача выявления аномалий;
- задача распознавания объектов на изображении.

Относительно недавно начали появляться задачи распознавания на изоб-

ражениях и видео тех или иных объектов. Именно нейронные сети способны решить данную задачу. Однако это не все. Нейронные сети также используются для распознавания и синтеза речи и машинного перевода.

В основу искусственных нейронных сетей легла идея биологических нейросетей человеческого мозга. В 1888 г. испанский доктор Рамони Кайал экспериментально показал, что мозговая ткань состоит из большого числа связанных друг с другом однотипных узлов — нейронов [3].

Компьютерные нейронные сети представляют собой математические модели, основанные на реальных биологических нейросетях живых существ. Для того, чтобы перевести такую модель в программу, в языке программирования обязательно должны быть математические модули [4].

Нейронная сеть состоит из последовательности связанных между собой слоев нейронов, которые могут получать и передавать информацию.

Выделяют 3 типа слоев искусственной нейронной сети:

1. Входной слой;
2. Скрытые (вычислительные) слои;
3. Выходной слой.

Более подробно процесс функционирования искусственной нейронной сети можно описать следующим образом:

1. Поступившие на вход сигналы умножаются на веса. Сигнал входа x_n умножается на соответствующий ему вес w_n . Результатом становится произведение этих величин $x_n w_n$.
2. Следующим этапом является передача произведения $x_n w_n$ в сумматор, задачей которого является суммирование всех входных сигналов в соответствии с формулой 1. Функциональная роль сумматора заключается в агрегировании всех входных сигналов в некоторое число — взвешенную сумму, которая будет характеризовать поступивший сигнал в целом. Такую взвешенную сумму можно представить как степень общего возбуждения нейрона.
3. В последствии взвешенная сумма передается на вход функции активации. Активационными функциями для нейронов служат линейные, сигмоидные и пороговые функции [5]. Это необходимо для того, чтобы понимать, рассматривать ли внешним связям данный нейрон как возбужденный или же игнорировать его.

4. Все выходные значения будут лежать в интервале от 0 до 1. Решением нейронной сети станет выходное значение наибольшей суммы.

$$\sum_{i=1}^n x_i w_i \quad (1)$$

Теория классических нейронных сетей активно развивается еще с середины XX века, однако в последнее время стала популярна теория глубоких нейронных сетей. Идея создания глубоких нейронных сетей заключается в рассмотрении нейросети как графа вычислений с наложением слоев нейронов друг на друга [6]. Глубокой считается та нейронная сеть, у которой между входным и выходными слоями присутствует несколько скрытых слоев. Основным отличием глубоких нейронных сетей является то, что в процессе обучения они выделяют основные признаки самостоятельно, перебирая их и присваивая им различный приоритет для более точного решения задачи.

Сверточная нейронная сеть — это специальная архитектура искусственных нейронных сетей, входящая в состав технологий глубокого обучения, которая была предложена Яном Лекуном в 1988 году и нацелена на распознавание образов.

На вход такой нейронной сети подается изображение. Все слои представляются в виде матриц. Каждый нейрон последующего слоя, выступающий в роли ячейки матрицы, соединен только с нейронами определенной области предыдущего слоя. Сначала идут сверточные слои, на которых к изображению применяются матричные фильтры, так называемые «карты признаков», направленные на выявление наличия какого-либо признака. Это приводит к уменьшению размеров каждого последующего слоя.

Для более радикального уменьшения размера слоев и, как следствие, получения более компактной архитектуры используется слой подвыборки — пулинг. Он разбивает изображение на небольшие прямоугольники, из которых впоследствии оставляет ровно 1 пиксел. У пулинга существует несколько разновидностей:

- max-pooling (наиболее употребителен);
- min-pooling;
- average-pooling.

Сверточные нейронные сети позволяют решать задачи компьютерного

зрения, к которым относятся:

- Classification — классификация изображения по типу объекта, которое оно содержит;
- Semantic segmentation — определение всех пикселей объектов определённого класса или фона на изображении. Если несколько объектов одного класса перекрываются, их пиксели никак не отделяются друг от друга;
- Object detection — обнаружение всех объектов указанных классов и определение охватывающей рамки для каждого из них;
- Instance segmentation — определение пикселей, принадлежащих каждому объекту каждого класса по отдельности.

Сверточные нейронные сети могут служить не только в развлекательных целях, но и выполнять серьезную работу. Так, они находят свое применение в автомобилях, помогая распознавать пешеходов, в системах безопасности аэропортов, распознавая потенциально опасные предметы, находящиеся в багаже.

Теория распознавания образов — это раздел кибернетики, развивающий методы классификации объектов по конечному числу признаков. Задачи такого типа в повседневной жизни встречаются довольно часто. Например, находясь за рулем транспортного средства, мы обращаем внимание на дорожные знаки, светофоры, и, зная правила дорожного движения и анализируя ситуацию вокруг, мы можем принять решение о том, как именно должен пролегать наш маршрут. Во всех живых существах эволюционно заложена способность к распознаванию объектов окружающего мира. Так, какое-нибудь некрупное животное может заметить и узнать хищника и, спрятавшись, предупредить угрозу своей жизни. Однако вопрос создания искусственных систем, которые смогут распознавать образы, остается открытым.

Для нейронной сети изображение прежде всего является набором пикселей, и главной сложностью при работе нейросети становится задача распознавания цельных объектов в наборе пикселей. Для обработки изображений используются различные алгоритмы, большинство из которых нацелены на нахождение особых точек. Такие точки являются уникальными и позволяют отделять объекты друг от друга. Для нахождения данных точек используются детекторы, обеспечивающие инвариантность нахождения одних и тех же точек относительно преобразований изображения. А для того, чтобы однозначно описать эти особые точки, используются дескрипторы [7]. Дескриптор

— это идентификатор особой точки, который выделяет ее из остального множества особых точек. Он, в свою очередь, должен обеспечивать инвариантность нахождения соответствий между особыми точками относительно преобразований изображений.

Во второй главе описывается процесс разработки нейронной сети для определения сходства людей по фотографиям.

Предварительно были рассмотрены различные виды программного обеспечения для реализации задачи и выбраны наиболее оптимальные. Языком программирования для написания кода нейронной сети был выбран Python 3, а средой разработки для написания кода нейронной сети был выбран Jupyter Notebook.

За основу была взята сверточная нейронная сеть компании Microsoft ResNet, точнее ее часть, не отвечающая за классификацию. При помощи оставшейся сверточной части возможно будет извлечь из того или иного изображения необходимые признаки, в нашем случае таковыми будут являться черты человеческого лица. Векторы признаков, которые будет выдавать сверточная нейронная сеть в результате своей работы для каждой из сравниваемых фотографий, будут сравниваться между собой.

В реализации данных действий поможет использование библиотеки dlib, которая содержит предварительно обученную сверточную нейронную сеть, выдающую векторы признаков. Векторы признаков, извлекаемые из фотографий, в dlib носят название дескрипторы. Дескриптор представляет собой набор из 128 вещественных чисел. dlib преобразует изображение человеческого лица в 128-мерное векторное пространство.

Сама библиотека dlib является кроссплатформенной программной библиотекой общего назначения, написанной на языке программирования C++. Данная библиотека имеет API для Python. Именно он и будет использоваться для дальнейшей работы.

Для реализации задуманного потребуются предварительно обученные модели:

1. «shape_predictor_68_face_landmarks.dat»;
2. «dlib_face_recognition_resnet_model_v1.dat».

После того, как из обеих фотографий будут извлечены дескрипторы лица, можно будет приступить непосредственно к их сравнению. Библиотека dlib

для такого сравнения рекомендует к использованию евклидово расстояние. Пороговым значением будет являться 0,6. Если евклидово расстояние между дескрипторами лиц составляет меньше 0,6, значит, вероятно, сравниваемые фотографии принадлежат одному и тому же человеку. Если же евклидово расстояние оказалось больше порогового значения, значит можно судить об отсутствии сходства между людьми.

Проведя сравнение лиц на различных фотографиях с использованием двух других метрик, можно убедиться в корректности работы нейронной сети и опытным путем установить пороговое значение для интерпретации результатов сравнения. Для данных метрик в таблице 1 приведены диапазоны принимаемых результатов значений и подобранные пороговые значения.

Таблица 1 — Характеристики метрик

Метрика	Диапазон принимаемых значений (верхняя граница приближительна)	Пороговое значение
Метрика Чебышева	0.0 — 0.25	1.7
Метрика Манхэттен	0.0 — 9.0	5.5

В третьей главе описывается процесс разработки web-приложения для определения сходства людей по фотографиям, а также последующее внедрение кода нейронной сети в разрабатываемое приложение.

Для использования web-приложения понадобится только открыть браузер и зайти на нужный сайт. Для работы с мобильным приложением понадобится его предварительная установка через сторонний сервис. В отличие от web-приложения, мобильное приложение не является программным обеспечением, запускаемым в web-браузере. Оно предназначено для работы на смартфонах, его функционал и дизайн создаются под возможности именно мобильных платформ.

Разрабатываемое web-приложение в свою очередь позволяет получать доступ к данному сервису как с мобильных, так и со стационарных устройств без скачивания отдельного приложения на персональное устройство, что положительно скажется на удобстве пользования.

После сравнения различных видов web-приложений было принято решение о создании серверного одностраничного web-приложения.

Предварительно был проанализирован инструментарий для реализации задачи и произведен выбор более оптимальных инструментов для этого. Языком для создания серверной части web-приложения был выбран Python 3. Этот язык программирования является высокоуровневым, динамически типизированным и интерпретируемым. Он имеет богатую библиотеку пакетов и сообщество разработчиков, которые поддерживают и развивают язык [8]. К неоспоримым достоинствам языка можно отнести его универсальность, простоту синтаксиса и богатые возможности. Код клиентской части пишется с использованием HTML и CSS. Основной средой разработки стал PyCharm, являющийся кроссплатформенной интегрированной средой разработки для языка программирования Python.

Для создания web-приложения был выбран Flask — небольшой и легкий фреймворк, написанный на языке Python [9]. Данный фреймворк предлагает полезные инструменты и функции для облегчения процесса создания web-приложений на языке Python. Инструментарий Bootstrap, в свою очередь, помогает сделать приложение привлекательным для пользователя.

Интерфейс приложения представлен на рисунке 1.

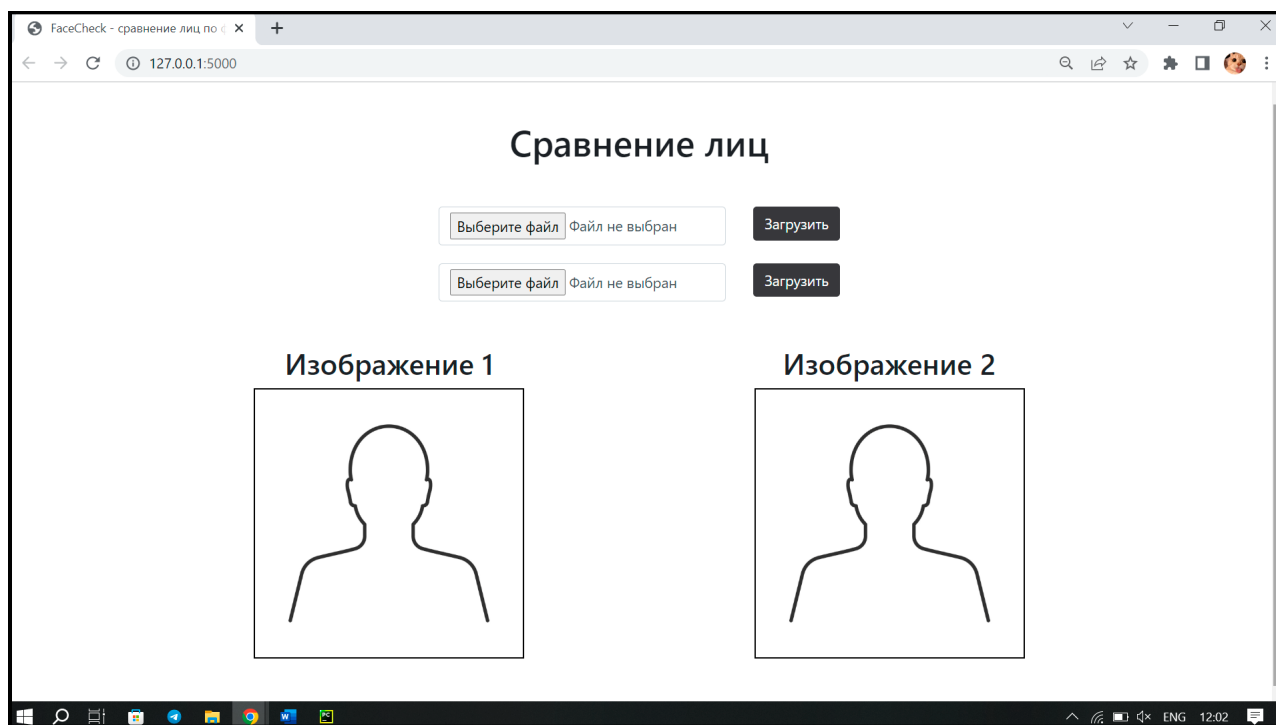


Рисунок 1 — Интерфейс web-приложения

Главная страница выглядит следующим образом. На ней имеются два по-

ля и две кнопки для загрузки изображений пользователем, а также две формы, в которых будут отображаться загруженные фотографии.

Пользовательский интерфейс является минималистичным и интуитивно понятным. Он так же не перегружен избыточными функциональными элементами, что делает его использование достаточно удобным.

ЗАКЛЮЧЕНИЕ

Отраслей, в которые можно внедрить технологии по распознаванию лиц, действительно много. На протяжении длительного времени внедряются технологии, позволяющие производить бесконтактную оплату с помощью не только гаджетов, но и с помощью лица или разблокирование устройства. Однако, стоит помнить о возникающих рисках. Системы для распознавания лиц открывают новые возможности для киберпреступников. Так, злоумышленник способен завладеть конфиденциальной информацией, выдавая себя за другого человека. Данная технология становится всё популярнее и постепенно проникает во все сферы деятельности человека, несмотря на сложность своей реализации и возможные риски. Возможности по внедрению технологии распознавания лиц поражают своими перспективами.

В результате проделанной работы была достигнута поставленная цель, а именно было создано web-приложение для определения сходства людей по фотографиям.

Для достижения поставленной цели были выполнены следующие задачи:

- были систематизированы знания об общих принципах функционирования нейронных сетей, их характеристиках и особенностях;
- были собраны и систематизированы знания о программном обеспечении для разработки нейронной сети;
- были собраны и систематизированы знания о программном обеспечении для разработки web-приложения;
- были выбраны наиболее оптимальные программные средства для реализации;
- была проведена программная разработка нейронной сети, позволяющей сравнивать лица людей;
- была проведена разработка web-приложения для сравнения двух фотографий, загруженных пользователем;
- был интегрирован код нейронной сети в разрабатываемое web-приложение.

Данный продукт может оказаться полезным людям как в быту, так и в профессиональной деятельности, например, государственными органам. С его помощью можно не только убедиться, является ли человек на фото чьим-то старым школьным другом и т.д.

Основные источники информации:

- 1 Как учится машина. Революция в области нейронных сетей и глубокого обучения/ Лекун Я. — М.: Альпина Диджитал, 2019. 351 с.
- 2 Руднева А.А., Искусственный интеллект и нейронные сети / А.А. Руднева // Журн. Информационные технологии в строительных, социальных и экономических системах. 2020 Т.19, №1. С. 48-50.
- 3 Организация и использование нейронных сетей (методы и технологии)/ Аксенов С.В., Новосельцев В.Б. — Томск: НТЛ, 2006. 128 с.
- 4 Нейронные сети: основы теории/ Галушкин А. — М.: Горячая линия — Телеком, 2012. 412 с.
- 5 Нейронные сети. [Электронный ресурс] URL: <http://statsoft.ru/home/textbook/modules/stneunet.html> (дата обращения: 13.05.2023). Загл. с экрана. Яз. рус.
- 6 Глубокие нейронные сети: пути применения. [Электронный ресурс] URL: <https://postnauka.ru/longreads/155983> (дата обращения: 02.04.2023). Загл. с экрана. Яз. рус.
- 7 Детекторы и дескрипторы ключевых точек. Алгоритмы классификации изображений. Задача детектирования объектов на изображениях и методы её решения. [Электронный ресурс] URL: <https://intuit.ru/studies/courses/10621/1105/lecture/17983> (дата обращения: 11.03.2023). Загл. с экрана. Яз. рус.
- 8 Почему Python такой популярный. [Электронный ресурс] URL: <https://tproger.ru/articles/pochemu-python-takoj-populjarnyj/> (дата обращения: 11.03.2023). Загл. с экрана. Яз. рус.
- 9 Создание веб-приложений с помощью Flask на Python 3. [Электронный ресурс] URL: <https://www.digitalocean.com/community/tutorials/how-to-make-a-web-application-using-flask-in-python-3-ru> (дата обращения: 06.05.2023). Загл. с экрана. Яз. рус., англ.