

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

Кафедра инноватики

СОРТИРОВКА И АНАЛИЗ ДАТАСЕТОВ ФОТОГРАФИЙ ПО
ОДНОРОДНЫМ И ПОДОБНЫМ ОБЪЕКТАМ

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 4111 группы
направления 27.03.05 «Инноватика»
института физики

Исаева Романа Сергеевича

Научный руководитель

к.ф.-м.н., доцент

должность, уч. степень, уч.
звание

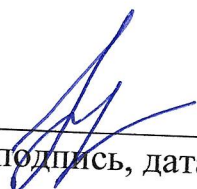

подпись, дата

Е.М. Ревзина
инициалы, фамилия

Научный консультант

Ассистент кафедры

должность, уч. степень, уч.
звание


подпись, дата

Н.Б. Макарова
инициалы, фамилия

Зав. кафедрой

к.ф.-м.н., доцент

должность, уч. степень, уч.
звание


подпись, дата

Е.М. Ревзина
инициалы, фамилия

Саратов 2023

Введение. Методы глубокого обучения эффективно применяются в различных областях, например, в медицине, в разработке самоуправляемых транспортных средств и во многих промышленных средах с потрясающими результатами в задаче идентификации объектов. Применение этих методов для сортировки датасетов фотографий могут увеличить количество перерабатываемого материала и, таким образом, упростить деятельность, связанную с профессиональной фотографией, а также повысить эффективность труда.[1, 11]

Целью выпускной квалификационной работы разработка собственного программного обеспечения на базе языка программирования Python, способного анализировать и сортировать датасеты фотографий (серии снимков) по лицам, находящимся в кадре (фотографии).

На основе поставленной цели необходимо решить следующие задачи:

1. Выявление принципов и оптимальных алгоритмов для создания кода машинного обучения.
2. Обсуждение проблем работников сферы, в которой применяется данный алгоритм, и их решение с помощью машинного обучения.
3. Изучение и анализ программных обеспечений, представленных на рынке с похожим или подобным функционалом.
4. Определение преимущества разработки и применения программы на практике работы специалиста.
5. Разработка программного обеспечения, реализующего предложенную технологию.

Дипломная работа занимает 51 страницу и имеет 26 рисунков.

Обзор составлен по 23 информационным источникам.

Во введение рассматривается актуальность работы, устанавливается цель и выдвигаются задачи для достижения поставленной цели.

Основное содержание работы

1. Анализ перспектив развития алгоритмов машинного обучения

1.1. Анализ тенденций в области машинного обучения

Обратимся к основным понятиям, необходимым для дальнейшего рассмотрения и разработки программного обеспечения с функциями анализа и сортировки фотографий по датасетам на основе верификации лиц, изображённых на них. Главенствующим понятием в этом вопросе является глубокое обучение.

Глубокое обучение – это класс алгоритмов машинного обучения, который является подмножеством искусственного интеллекта, использующего несколько уровней представления данных и извлечения признаков [2, 6, 9]. В мире уже разработано множество приложений, где глубокое обучение играет основную роль, - эти алгоритмы как правило связаны с распознаванием речи, изображений, лиц, преобразованием устной речи в текст, прогнозом погоды и т.д. Глубокое обучение позволяет подготовить с помощью вычислительных моделей множество слоев для изучения представлений данных с абстрагированием множества слоев.

Но в наших интересах рассмотреть одно из самых востребованных направлений в мире - обнаружение объектов. Если подробно разобраться в вопросе, то мы выясним, что обнаружение объектов — это составная деталь в механизме машинного зрения, при которой программное обеспечение способно обнаружить, отследить и определить местонахождение объекта по любому входящему видео или изображению. Главной особенностью обнаружения объектов является распознавание класса объекта (например, человек, стол, стул и т.д.), а так же координаты местоположения объекта на данном изображении. Создавая зонирующую рамку вокруг исследуемого объекта, алгоритм определяет регион местоположения с высокой точностью. Возможность определения местоположения объекта на изображении описывает точность выполнения алгоритма, который используется для обнаружения. Верификация

лица является одной из основополагающих функций в системе обнаружения объектов.[13, 14]

Обнаружение объектов происходит в три этапа:

1. На входном изображении создаются небольшие прямоугольные участки.
2. Завершается извлечение признаков для каждой части прямоугольной области, чтобы предсказать, содержит ли прямоугольная форма существующий объект.
3. Наконец, прямоугольники, перекрывающие друг друга, объединяются в одну ограничивающую рамку, показывающую результат обнаружения объекта.

Программа, разработке которой посвящена эта выпускная квалификационная работа, поможет быстро упорядочить несортированную базу изображений. Основная идея заключается в том, чтобы предоставлять изображения в качестве входных данных программе, в последующем на них проводить обнаружение лица. Затем изображение будет перемещено или скопировано в папку с аналогичным изображением. Это позволит быстро просмотреть и упорядочить большое количество фотографий.

Входное изображение передается в программу, которая определяет, есть это лицо на других снимках фотосета или нет. Для распознавания лиц используется "каскадный алгоритм Хаара". После распознавания лица создаётся каталог по имени заданной метки класса, а для распознавания лиц используется алгоритм "Линейной бинарной гистограммы"

Алгоритм каскада Хаара использует два фундаментальных и трудоемких процесса. На изображении большая часть области не является лицевой областью. Следовательно, мы проверяем, является ли окно лицевой областью, если оно не удаляет это окно за один снимок.

Модуль распознавания лиц используется для распознавания лиц на изображении, которое получается с помощью модуля распознавания лиц. После того, как лица людей на изображении будут распознаны, модуль затем присвоит

изображению уникальный идентификатор, который будет именоваться. Затем он передаст уникальный идентификатор человека в модуль сопоставления лиц в качестве входных данных. Распознавание лиц достигается с помощью алгоритма LBPH, реализованного через OpenCV.

Модуль сопоставления лиц это следующий за распознаванием лиц. Для этого потребуется идентификатор человека, сгенерированный модулем распознавания лиц, и выполнить поиск по текущему каталогу, чтобы найти изображения лиц с таким же идентификатором. Если совпадающие идентификаторы не найдены, то эта проблема возвращается пользователю в виде сообщения. Если найдены совпадающие изображения, то текущий путь к каталогу и имена файлов изображений передаются либо в модуль организации изображений, либо в модуль отображения результатов изображений на основе выбранного флага параметров.[6, 18]

1.2. Изучение и анализ программных обеспечений, представленных на рынке, с похожим или подобным функционалом.

Качественный анализ программных обеспечений с похожим или подобным функционалом, уже представленных на рынке, является важным этапом в разработке собственного программного обеспечения. Предлагаем рассмотреть самые интересные «кейсы» конкурентов.

1.2.1 Распознавания лиц на снимках в смартфонах.

Функция сортировки фотографий по лицам в смартфонах является уже неотъемлемой частью жизни каждого пользователя современных гаджетов, а работа самой функции начинается с первого снимка, сделанного на телефон. Алгоритмы, встроенные в такие сервисы как Google фото, галерея на iPhone, уже имеют синхронизацию с серверами, на которых хранятся предобученные нейросети, на базе которых работает распознавание и сравнение лиц. Но, к сожалению, встроенные алгоритмы не предназначены для работы с большими сериями съемок, так как крайне неудобно загружать объемные архивы для пакетной сортировки снимков.

1.2.2 Распознавание лиц на фотографиях, загруженных на сервера ВКонтакте.

Основной смысл работы распознавания лиц в социальной сети раскрывается таким образом: "Искусственный интеллект подсказывает, а не навязывает». Дело в том, что в момент загрузки изображения в социальную сеть, алгоритмы самостоятельно определяют, кто из друзей представлен на фотоснимке. После этого система может подсказать пользователю отметить вручную друзей на фотографии, а также предоставляет возможность друзьям, присутствующим на снимке отметить самих себя. Финальное решение, подтвердить или отклонить свою отметку на фотографии, принимает сам пользователь.

1.2.3 Технологии для сортировки фотографий на платформе Облако mail.ru

Обучаемая модель способна сравнивать снимки людей – в профиль, вполоборота, три четверти, анфас, местами скрытые лица. Чем больше лиц какого-либо человека она видит, тем выше результат сортировки, это называется принципом кластерного анализа.

Принцип работы платформы таков: пользователь заходит по открытой ссылке в папку, алгоритм находит на фотоснимках в ней лица, анализирует и сортирует файлы, а потом передаёт результат пользователю. Полученный результат дублируется в кеш, и следующие пользователи, переходящие по внешней ссылке, будут сразу его получать.

Нужно иметь в виду, что файлы в папках могут подвергаться изменениям, добавляться и удаляться, для этого разработчики добавили автоматическую инвалидацию кеша с чётко выстроенной периодичностью. Даже в случае, если в папку, для которой ещё не создан кеш, зайдёт сразу несколько пользователей, то сортировка всё равно будет сгенерирована один раз и закеширована. [12]

2. Анализ существующих методов и технологий распознавания лиц с помощью технологии машинного зрения.

Обнаружение лиц является системой особенностей технологий, состоящей из ряда трудностей, которые нужно структурированно описать:

1. Нужно проанализировать изображение и распознать все присутствующие лица на снимке.

2. Вторым пунктом реализации нашего алгоритма будет анализ каждого отдельного взятого лица. Компьютеру необходимо понять, что на изображении находится один и тот же человек независимо от поворота угла лица или неравномерного освещения.

3. Для характерного отличия человека от других людей алгоритму нужно выявить ряд особенностей присутствующих на каждом отдельно взятом лице, в решении этой задачи нам помогут различные антропометрические данные – расстояние между глазами, ширина подбородка и т.д.

4. В итоге алгоритму предстоит проанализировать и сопоставить уникальные антропометрические данные с данными других прорабатываемых нами личностями, чтобы вывести результат с конкретным человеком.

Вышеописанные операции человеческий мозг проводит мгновенно и не требует полной вовлеченности сознания, то есть автоматически. Подводя итоги, люди определяют лица с непревзойденной точностью и, зачастую, наблюдают очертания лиц в бытовых и не только объектах.

Когда основные принципы работы распознавания лиц машинным зрением определены, можно приступить к практической реализации программного обеспечения.

3. Проведение тестирования предложенных методов

3.1. Программная реализация предложенных методов.

Реализация предложенных подходов в среде Google Colab. В реализованном программном обеспечении применялся следующий ряд модулей:

- OS – модуль, предоставляющий возможность обращаться к множеству функций для надежной работы с операционной системой, с учётом того, что стабильность действия функций не зависит от самой системы, поэтому программы становятся переносимыми.

- Operator – модуль, который экспортирует набор эффективных функций, которые соответствуют внутренним операторам Python.

- shutil – предоставляет возможность множеству функций для исполнения высокоуровневых как с коллекциями файлов, так и с самими файлами.

- DeepFace – это легкая платформа для распознавания лиц и анализа атрибутов лица (возраст, пол, эмоции и раса).

Использование этих модулей в Google Colab позволяет более эффективно реализовывать и тестировать предложенные подходы в машинном обучении и глубоком обучении. Кроме того, Google Colab обеспечивает доступ к более мощным вычислительным ресурсам и GPU, что позволяет быстрее и эффективнее работать с большими объемами данных и моделями глубокого обучения.

Подключаем все доступные методы сравнения 128 сгенерированных характеристик, которые уже предустановлены в библиотеку DeepFace.

Теперь определим функцию `face_v`, которая будет верифицировать лица.

- `img.1` – оригинальное изображение сравниваемого лица

- `img.2` – изображение лица, с которым сравнивают

Вся база сравниваемых фотографий, будет проходить через определяющую функцию циклом, что позволит проверить каждую фотографию с оригинальным изображением.

Мы допускаем возможность того, что программа не считает лица по следующим возможным причинам, например: глубокая тень на лице, случайный объект перекрывающий рассматриваемое алгоритмом лицо, или солнечный блик. Если это случится, то цикл остановится и весь алгоритм придётся запускать заново. Чтобы этого не произошло функция `enforce` выносит все не считываемые картинки в ошибку, о которой сообщает пользователю.

В том случае, если все условия оказались выполнены, то реализуемое нами программное обеспечение создает директорию, в которой и будет отображен результат верификации.

Чтобы привести код в исполнение, осталось только задать диапазон работы цикла, в зависимости от количества проверяемых нами фотографий.

3.2. Проведение вычислительных экспериментов

Проведем вычислительный эксперимент по классификации данных, а также продемонстрируем работу программного обеспечения.

1. Загрузка данных: Пользователь может выбрать набор данных из доступных вариантов и загрузить их в программу.
2. Предобработка данных: Набор данных может быть предварительно обработан для удаления выбросов, заполнения пропущенных значений и приведения данных к единому формату.
3. Разбиение данных: Набор данных разбивается на тренировочный и тестовый наборы, используя метод перекрёстной проверки (cross-validation).
4. Тестирование модели: Разработанная модель используется для классификации тестового набора данных, и точность ее работы оценивается на основе заранее известных классов в данных.
5. Визуализация результатов: Подобласть интерфейса программы содержит вкладки, которые отображают результаты точности работы модели, а также важность каждого признака данных для конкретного классификатора.
6. Оценка точности классификатора: В программном обеспечении доступен анализ метрик, таких как матрица ошибок и метрики качества, которые помогают определить, насколько хорошо работает классификатор на данном наборе данных и какие можно предпринять действия для улучшения производительности.

Таким образом, предоставленный в интерфейсе инструментарий предоставляет возможность достаточно гибко настроить параметры классификатора и анализировать его производительность на конкретных наборах данных.

Для проверки исправности программного обеспечения загрузим в новую директорию 30 изображений разных медиа-персон.

Обозначим диапазон проверки циклом от 1 до 31, таким образом мы сможем верифицировать каждое изображение. За эталон, для проверки снимков мы возьмем изображение под номером 2 и запустим код.

8 из 8 изображений Президента Российской Федерации были верно определены и распознаны. При добавлении в датасет ещё 10 случайных изображений других медиа-персон, результат стал 7 из 8, что даёт нам возможность и перспективу к дальнейшей и более детальной разработке собственного программного обеспечения оффлайн верификации лиц.

Для достижения наилучшего качества производительности классификатора на наборе данных необходимо проводить предварительную обработку и анализ данных, включая удаление выбросов и заполнение пропущенных значений.

Предложенные подходы могут использоваться для решения широкого спектра задач классификации данных, включая медицинские, финансовые, маркетинговые и другие приложения, где необходимо быстро и точно классифицировать большие объемы данных.

Заключение. Представленные в выпускной квалификационной работе исследования позволяют сформировать следующие выводы:

- Исследования на основе алгоритмов машинного обучения показали высокую эффективность в предсказании и оптимизации различных систем. Например, построение моделей машинного обучения позволяет сократить время поиска оптимальных параметров, повышает точность прогнозирования и улучшает качество принимаемых решений в динамических системах.

- Использование алгоритмов машинного обучения в науке и технике позволяет решать задачи, которые ранее не могли быть решены при помощи классических методов анализа и моделирования.

- Важным фактором успеха при применении методов машинного обучения является правильный выбор алгоритмов и признаков, которые используются для обучения модели.

- Вычислительные эксперименты подтвердили жизнеспособность автоматизированной системы обнаружения лиц и организации их в отдельную папку с использованием алгоритмов глубокого обучения и методов обработки изображений. Наша программа работала с большим набором данных изображений, алгоритмами обучения и прогностическими шаблонами для обнаружения и классификации объектов. В нашей программе используются новые технологии глубокого и машинного обучения, благодаря чему нам удалось по-новому взглянуть на сортировку и группировку изображений.

Перспективой нашего исследования является усовершенствование описанного программного обеспечения, с целью уточнения его анализа других наборов данных, включающих множество объектов (лиц) на изображении. Также необходима разработка и внедрение этой технологии в приложение со своим интерфейсом, чтобы пользователю было комфортно пользоваться ей без вторжений в кодовое тело программы.

Таким образом, исследования в области машинного обучения имеют большой потенциал для оптимизации различных процессов и решения сложных

задач, и их использование будет продолжать расширяться и вносить новые возможности в науку и технику.

Список использованных источников

1. Арзаманов, Н.А. Технология машинного обучения и ее практическое применение / Н. А. Арзаманов, Н. И. Ематина // Исследование различных направлений современной науки – материалы XXI Международной научно-практической конференции. В 2-х частях. 24 апреля 2017. – Научный центр "Олимп" (Астрахань), 2017. – С. 7-10. – Текст : непосредственный.
2. Аусабаев, Д.М. Использование машинного обучения в поддержке принятия решений / Д. М. Аусабаев, О. П. Волобуев // Прикладная математика и информатика: современные исследования в области естественных и технических наук – материалы III научно-практической всероссийской конференции (школы-семинара) молодых ученых. Тольятти, 24–25 апреля 2017 года. – Издатель Качалин Александр Васильевич, 2017. – С. 43-47.
3. Власов, А.В. Машинное обучение применительно к задаче классификации семян зерновых культур в видеопотоке / А. В. Власов, А. С. Федеев // Молодежь и современные информационные технологии – сборник трудов XIV Международной научно-практической конференции студентов, аспирантов и молодых учёных, 07–11 ноября 2016. – Национальный исследовательский Томский политехнический университет (Томск), 2016. – С. 133-135.
4. Жуков, Д.А. Формирование контрольных выборок при технической диагностике объекта с применением машинного обучения / Д. А. Жуков, А. С. Хорева, Ю. Е. Кувайскова, В. Н. Клячкин // Математические методы и модели: теория, приложения и роль в образовании – международная научно-техническая конференция : сборник научных трудов, 28–30 апреля 2016 года. – Ульяновский государственный технический университет (Ульяновск), 2016. – С. 44-48.
5. Иванников Ю.Ю. Применение методов машинного обучения для выявления бот-трафика среди запросов к веб-приложению / Ю. Ю. Иванников,

Е. Ю. Митрофанова // Сборник студенческих научных работ факультета компьютерных наук ВГУ, Факультет компьютерных наук, 2017. – ФГБОУ ВО «Воронежский государственный университет», 2017. – С. 119123.

6. Клячин В.Н. Использование агрегированных классификаторов при технической диагностике на базе машинного обучения / В. Н. Клячин, Ю. Е. Кувайскова, Д.А. Жуков // Информационные технологии и нанотехнологии (ИТНТ-2017) – сборник трудов III международной конференции и молодежной школы. Самарский национальный исследовательский университет имени академика С.П. Королева. 2017. – Предприятие "Новая техника" (Самара), 2017. – С. 1770-1773.

7. Кононова, Н.В. Исследование подсистемы контентной фильтрации с использованием методов машинного обучения / Н. В. Кононова, Ю. А. Андрусенко, Т. А. Самокаева // Студенческая наука для развития информационного общества – сборник материалов VI Всероссийской научно-технической конференции. 22–26 мая 2017. – Северо-Кавказский федеральный университет (Ставрополь), 2017. – С. 268-270.

8. Мелдебай, М.А. Анализ мнений покупателей на основе машинного обучения / М. А. Мелдебай, А. К. Сарбасова // Прикладная математика и информатика: современные исследования в области естественных и технических наук – материалы III научно-практической всероссийской конференции (школы-семинара) молодых ученых. 24–25 апреля 2017 года. – Издатель Качалин Александр Васильевич, 2017. – С. 360363.

9. Наумов, Д.П. Регулятор САР на основе машинного обучения / Д. П. Наумов, Д. П. Стариков // Информационные технологии в управлении, автоматизации и мехатронике – сборник научных трудов Международной научно-технической конференции. 06–07 апреля 2017 года. – ЗАО "Университетская книга" (Курск), 2017. – С. 106-114.

10. Ткач, Т. Ч. Машинное обучение и обработка больших данных - обучение в основной и средней школе / Т. Ч. Ткач // Актуальные проблемы методики обучения информатике и математике в современной школе –

материалы международной научно-практической интернет-конференции. Московский педагогический государственный университет, Москва, 24 апреля 2020 года. – Московский педагогический государственный университет (Москва), 2020. – с. 217-223.

11. Распознавание лиц с InsightFace или как CatBoost имена угадывал [Электронный ресурс] // Сайт ХАБР [Электронный ресурс] : [Электронный ресурс] : [сайт], – URL : <https://habr.com/ru/articles/699232/> (дата обращения 17.04.2023).– Загл. с экрана.. – Яз. Рус

12. Разложить по полочкам: как мы научили Облако сортировать фотографии по лицам [Электронный ресурс] // Сайт ХАБР [Электронный ресурс] : [сайт]. – URL: <https://habr.com/ru/companies/vk/articles/678770/> (дата обращения 17.04.2023).– Загл. с экрана.. – Яз. Рус

13. Обучение машины – забавная штука: современное распознавание лиц с глубинным обучением [Электронный ресурс] // Сайт ХАБР [Электронный ресурс] : [сайт]. – URL: <https://habr.com/ru/articles/306568/> (дата обращения 17.04.2023).– Загл. с экрана. – Яз. рус

14. Машинное зрение. Что это и как им пользоваться? Обработка изображений оптического источника [Электронный ресурс] // Сайт ХАБР [Электронный ресурс] : [сайт]. – URL: <https://habr.com/ru/articles/350918/> (дата обращения 17.04.2023).– Загл. с экрана. – Яз. рус.

15. Шогунова, Н.М. Разработка программы компьютерного зрения для локализации объектов на изображении и видеопотоке / Н.М. Шогунова, В.Я. Олексюк // Сборник статей Всероссийской студенческой научнопрактической междисциплинарной конференции «Молодежь. Наука.

16. Якимчук, А.А. Глубокое обучение как эффективный метод машинного обучения / А.А. Якимчук // Научное сообщество студентов XXI столетия. Технические науки – сборник статей по материалам ХСП студенческой международной научно-практической конференции. 2020. – ООО “Сибирская академическая книга” (Новосибирск), 2020. – С. 40-43.

17. Ярыгин, А.А. Актуальные вопросы машинного обучения с подкреплением интеллектуальных агентов в задачах принятия решений / А.А. Ярыгин // Автоматизация: проблемы, идеи, решения - сборник статей по итогам Международной научно-практической конференции 2017. – ООО "Агентство международных исследований", 2017. – С. 62-68.
18. Dhanya, S. An Enhancement of the MapReduce Apriori Algorithm Using Vertical Data Layout and Set Theory Concept of Intersection / S. Dhanya, M. Vysaakan, A. S. Mahesh // Intelligent Systems Technologies and Applications. – Springer International Publishing Switzerland, 2016. – P. 225-233.
19. FaceNet — пример простой системы распознавания лиц с открытым кодом Github [Электронный ресурс] // Сайт Neurohive [Электронный ресурс] : [сайт]. – URL: // <https://neurohive.io/ru/tutorial/raspoznvanie-lica-facenet/> (дата обращения 17.04.2023). – Загл. с экрана. – Яз. англ.
20. Хауди Хо – Просто о мире IT [Электронный ресурс] // Сайт YouTube.ru [Электронный ресурс] : [сайт]. – URL: // <https://www.youtube.com/@HowdyhoNet> (дата обращения 17.04.2023). – Загл. с экрана.. – Яз. рус
21. Краткая история машинного зрения [Электронный ресурс] // Научно – производственная компания «НИРМЕХ» [Электронный ресурс] : [сайт]. – URL: <https://inlnk.ru/xv4lv6> (дата обращения 17.04.2023). – Загл. с экрана. – Яз. рус
22. Глазами компьютера: как работает машинное зрение [Электронный ресурс] // Сайт ФОРМАТ КОДА [Электронный ресурс] : [сайт]. – URL: <https://formatkoda.ru/blog/glazami-kompyutera-kak-rabotaet-mashinnoe-zrenie/> (дата обращения 17.04.2023). – Загл. с экрана. – Яз. рус