

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра физики открытых систем

**Проектирование и частичная реализация системы организации
фотоматериалов с распознаванием образов**
АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 2241 группы

направления

09.04.02 «Информационные системы и технологии»

института физики

Хмырова Артема Михайловича

Научный руководитель
доцент кафедры математической
теории упругости и биомеханики,
к.ф.-м.н., доцент

Л.В. Бессонов

Заведующий кафедрой
физики открытых систем,
д.ф.-м.н., профессор

А.А. Короновский

Саратов 2024 г.

Проектирование и частичная реализация системы организации фотоматериалов с распознаванием образов

Введение

В эпоху цифровизации и информационных технологий, управление большим объемом фотоматериалов становится все более сложной задачей. Системы организации фотоматериалов, основанные на распознавании образов, предоставляют мощный инструмент для эффективного и быстрого поиска нужных изображений среди тысяч файлов.

Система распознавания объектов на фотографии имеет широкое применение в современном мире. Правоохранительные органы используют распознавание лиц для идентифицирования террористов и преступников, а компании мобильной связи используют распознавание лиц, чтобы предоставить потребителям новые уровни биометрической безопасности.

Искусственный интеллект эволюционирует с каждым годом, модернизируются методы распознавания, разрабатываются новые алгоритмы для решения задач в любых областях науки.

Различают три вида искусственного интеллекта: слабый (Narrow AI), сильный (AGI) и супер-ИИ (Super AI).

Первый вид используется повсеместно: голосовые помощники, распознавание лиц, поиск собеседника и другое.

Сильный ИИ максимально приближен к способностям человеческого интеллекта, обычно применяется в медицине и способен заменить некоторые функции человеческого тела такие как, бионические руки и ноги.

Супер-ИИ может стать не просто подобным людям, но и превзойти лучшие умы человечества во всех областях, при этом программируя самого себя, продолжая совершенствоваться и, вероятно, разрабатывая новые системы и алгоритмы самостоятельно.

Современные средства поиска объектов на фотографиях модернизируются. Множество сервисов могут определять людей, животных, здания и прочие объекты на фото и видеопотоках, однако для комбинированного поиска таких средств мало, а возможности существующих сервисов ограничены.

Одной из главных проблем при хранении и сортировке данных, является беспорядочность единой кучи фотографий, содержащей в себе множество неотсортированных данных. Поиск среди множества фотоэлементов для человека занимает длительное время.

Актуальность исследования заключается в геометрической прогрессии использования нейронных сетей и их возможностей, что говорит о непрерывном спросе искусственного интеллекта на данный момент и в будущем.

Объект исследования – фотоархивы и системы организации фотоматериалов, основанные на распознавании образов.

Предмет исследования – разработка нейронной сети для сортировки и поиска лиц или группы лиц на изображениях, направленной на массовое использование.

Целью работы является реализация программного обеспечения с клиентским интерфейсом для поиска и сортировки фотоэлементов среди множества фотографий.

Значимость магистерской работы заключается в решении представленной выше проблемы, путем написания нейронной сети для сортировки и поиска лиц или группы лиц среди множества фотоэлементов. Данная программа направлена на массовое пользование, что в будущем позволит отслеживать работу нейронной сети и обучать на более высоких мощностях.

Задачи исследования включают в себя:

1. Анализ методов распознавания лиц;
2. Анализ рынка программ – конкурентов;
3. Проектирование пользовательского интерфейса, обеспечивающего легкость в использовании и функциональность;
4. Разработка ядра программы, способного эффективно обрабатывать фотоматериалы и проводить распознавание образов;
5. Тестирование системы и анализ полученных результатов.

В ходе работы будет продемонстрировано, как современные методы программирования и искусственного интеллекта могут быть применены для решения задач организации фотоматериалов, делая процесс более эффективным и доступным для широкого круга пользователей.

Структура работы содержит в себе 62 страницы, состоящих из титульного листа, содержания, введения, основную часть, состоящую из теоретической и практической частей, заключения и списка литературы.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первая глава магистерской работы «**Анализ предметной области и методов распознавания лиц для организации фотоархива**» посвящена изучению проблематики имеющихся нейронных сетей для организации фотоэлементов. При анализе существующих программ были выявлены преимущества каждой из программ, удовлетворяющие основным задачам программы. Также выявлены недостатки, в следствии чего была поставлена цель работы и ряд задач, способствующих решению общей проблемы нейронных систем для организации фотоэлементов и поиску лиц по фото. По ходу данной работы были проанализированы следующие сервисы по поиску лиц и организации фотоматериалов:

- Google Photo;

- Search4faces;
- Яндекс картинки;
- Adobe Bridge;
- Corel AfterShot Pro;
- Picasa;
- Digikam;
- Tonfotos.

Каждая из рассмотренных программ отлично работает для поставленных задач, которые описывает разработчик. Стоит отметить, что наиболее удобная в использовании и автоматически находит лица на фотографиях – программа Tonfotos. В ходе работы, программа подвергается более глубокому анализу, что позволяет выявить больше недостатков и учесть их при разработке аналога.

Для разработки системы поиска и организации фотоматериалов, первая глава посвящена подробному описанию методов распознавания лиц, сравнения скорости работы и других характеристик. Анализируются следующие методы:

- Метод главных компонент;
- Линейный дискриминантный анализ;
- Метод гибкого сравнения на графах;
- Метод Виолы–Джонса;
- Сверточные нейронные сети;
- Метод опорных векторов;
- Метод k-ближайших соседей.

Метод главных компонент позволяет уменьшить размерность данных, сохраняя при этом наибольшее количество информации, а также выделять основные характеристики данных. Однако он не всегда эффективен для

нелинейных данных и может потерять часть информации при сжатии данных.

Линейный дискриминантный анализ эффективен для разделения классов данных и учитывает внутриклассовую и межклассовую дисперсию, но плохо работает с несбалансированными данными и чувствителен к выбросам.

Метод гибкого сравнения на графах учитывает нелинейные связи между данными и эффективен для анализа сложных структур данных, но требует большого количества вычислительных ресурсов и может быть сложен в интерпретации результатов.

Метод Виолы-Джонса подходит для обнаружения объектов на изображениях и обладает высокой точностью распознавания, но может быть чувствителен к изменениям в освещении и фоне и требует большого объема обучающих данных.

Сверточные нейронные сети эффективны для анализа изображений и видео, а также автоматического извлечения признаков, однако требуют большого объема данных для обучения и много вычислительных ресурсов.

Метод опорных векторов хорошо работает с высокоразмерными данными и является эффективным для классификации, но требует правильного выбора параметров и ядра, а также может быть чувствителен к шумам в данных.

Метод k -ближайших соседей прост в реализации и понимании, эффективен для классификации, однако требует хранения всего обучающего набора данных и чувствителен к выбору параметра k .

Для реализации программы, выявляются критерии для успешного прогона нейронной сети, которые обеспечивают выбор метода распознавания

в качестве главного. Таблица 1 демонстрирует сравнение характеристик методов распознавания лиц.

	Шум	Ракурс	Мимика	Свет	Качество	Помехи	Время	Память
Метод главных компонент	-	+	+	+	+	+	-	-
Линейный дискриминантный анализ	-	+	-	-	-	-	-	-
Метод гибкого сравнения на графах	+	-	-	-	-	+	+	+
Метод Виолы–Джонса	-	-	-	-	+	+	-	-
Сверточные нейронные сети	-	-	+	-	+	+	+	-
Метод опорных векторов	+	-	-	+	+	+	-	-
Метод k-ближайших соседей	-	-	-	-	+	-	+	+

Таблица 1 – влияние факторов на работу алгоритмов

Табличные обозначения:

- «-» - Незначительное влияние на работоспособность алгоритма;
- «+» - Значительное влияние на работоспособность алгоритма.

Как видно из приведенной выше таблицы, метод k-ближайших соседей, наиболее подходящий для использования в программе по поиску лиц, за счет устойчивости к шуму, помехам, а также ракурсу и мимике.

В качестве вспомогательных инструментов рассматриваются встроенные библиотеки для распознавания лиц:

- Библиотека OpenCV;
- Библиотека Face Recognition.

Библиотека OpenCV предоставляет широкие возможности для обработки изображений и видео, включая распознавание объектов, трекинг, сегментацию и многое другое. Это популярное решение для компьютерного зрения, которое поддерживает различные языки программирования.

Библиотека face recognition специализируется на распознавании лиц и предоставляет инструменты для обучения моделей и их использования в приложениях. Она обладает высокой точностью распознавания и широкими возможностями настройки параметров для достижения оптимальных результатов.

Благодаря анализу рынка программ-аналогов и изучения характеристик методов распознавания лиц, можно разработать систему организации фотоэлементов, опираясь на все аспекты, изложенные первой главой.

Вторая глава «Разработка цифрового ресурса методического обеспечения для изучения web-программирования в старших классах» направлена на описание разработки программного приложения на языке Python.

Front-end и back-end - это две основные части веб-приложений, которые взаимодействуют друг с другом для обеспечения работоспособности и удобства использования приложения.

Front-end – это часть приложения, с которой взаимодействует пользователь. Здесь размещается пользовательский интерфейс, который включает в себя дизайн, макет и интерактивные элементы, отвечающие за отображение информации пользователю, его взаимодействие с приложением, а также за удобство использования приложения.

Back-end – это часть приложения, которая отвечает за обработку данных, бизнес-логику, взаимодействие с базой данных и обеспечение функциональности приложения. Back-end выполняет запросы пользователя, обрабатывает их, взаимодействует с базой данных, обеспечивает безопасность и контролирует работу всего приложения.

Данные сущности приложений работают вместе, обмениваясь данными и командами для обеспечения полноценного функционирования и удобства использования приложения пользователем.

Программа состоит из двух основных этапов разработки. Back-end часть или ядро программы содержит основную логику программы. На данном этапе программируется нейронная сеть, которая обучается на заранее подготовленных фотоэлементах. Разрабатывается функция, которая проходит по массиву с данными и сравнивает каждый элемент с обученной моделью, на основе чего создает отдельную директорию и перемещает копию картинки в эту папку.

В случае возникновения ошибки сети, перебоев с электричеством или других факторов, влияющих на работоспособность программы, предусмотрена возможность синхронизации с деревом, в случае повторного прогона, программа сообщит пользователю, что обучение на данной фотографии уже проводилось и нет необходимости заново анализировать данный объект.

После написания ядра программы, необходимо реализовать пользовательский интерфейс. Данный этап подразумевает верстку всех

необходимых элементов и разделов, предусматривающий дальнейшее использование данных разделов и элементов.

Реализуются такие разделы, как: «Главная», «Галерея», «Методы», «О программе», «Тех. поддержка».

- Раздел «Главная» описывает возможности данной программы;
- Раздел «Галерея» содержит в себе выходные данные программы, где генерируются копии изображений, структурированные по папкам;
- Раздел «Методы» предоставляет пользователю выбор метода прогона и является страницей запуска программы;
- Разделы «О программа» и «Тех. поддержка» содержат в себе контакты и более подробную информацию о работе программы.

Дополнительной возможностью данной программы, является возможность смены темы программы, что позволяет пользователю настроить цветовые элементы под себя.

После реализации back-end и front-end частей, интерфейс программы интегрируется с ядром, путем переназначение строгих функций на уникальные названия интерфейсных элементов. Успешная работа программы обуславливается корректным проходом нейронной сети по фотоэлементам, где выходные данные соответствуют действительности.

Тестирование программы происходит на базе СГУ при управлении медиакоммуникаций, что позволило решить ряд проблем, а также получить обратную связь по улучшению работы программы.

Заключение

Разработка нейронной сети для сортировки и поиска лиц, стала реальностью благодаря проведенному анализу методов распознавания лиц, рынка программ-конкурентов, проектированию пользовательского

интерфейса, разработке ядра программы и проведению тестирования системы. Эффективность и удобство использования разработанного программного обеспечения позволят пользователям легко и быстро обрабатывать большие объемы фотоматериалов, основываясь на распознавании образов. Полученные результаты демонстрируют, как современные методы программирования и искусственного интеллекта могут быть применены для решения актуальных задач организации фотоархивов.

Данная работа имеет значимость и практическую ценность в контексте растущего интереса к искусственному интеллекту и системам распознавания образов на фотографиях. Прделанная работа позволит упростить и сделать более эффективным процесс поиска и сортировки фотоэлементов среди множества фотографий, что является актуальной задачей в современном мире цифровизации и информационных технологий. В заключение, результаты этой работы могут быть использованы в различных областях, где требуется эффективная организация фотоматериалов на основе распознавания образов, и приносить пользу как для бизнес-сектора, так и для государственных учреждений, исследовательских центров и обычных пользователей, нуждающихся в удобных инструментах для работы с фотографиями.

В ходе работы были достигнуты все поставленные цели:

1. Проанализированы методы распознавания лиц;
2. Проанализирован рынок программ – конкурентов;
3. Спроектирован пользовательский интерфейс, обеспечивающего легкость в использовании и функциональность;
4. Разработано ядра программы, способное эффективно обрабатывать фотоматериалы и проводить распознавание образов;
5. Протестирована система и анализ полученных результатов.

Дальнейшие труды могут быть направлены на улучшение производительности программы, доработку визуальных частей, а также на поддержку разных операционных систем.

Цель и задачи в ходе данной работы были достигнуты.