МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

(МИНОБРНАУКИ РОССИИ)

федеральное государственное бюджетное образовательное учреждение

высшего образования

«Рыбинский государственный авиационный технический университет

имени П. А. Соловьева»

(РГАТУ имени П. А. Соловьева)

Институт информационных технологий и систем управления

Кафедра вычислительных систем

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА**

**Дипломный проект**

Разработка программы для распознавания человеческих эмоций по лицу на основе машинного обучения

на соискание квалификации бакалавра

по направлению 09.03.01 Информатика и вычислительная техника

**Пояснительная записка**

Соискатель, студент группы ИВБ2–19 Иванов Д. А.

Руководитель канд. техн. наук, доцент Сизов П. В.

Нормоконтролер канд. техн. наук, доцент Гусаров А. В.

К защите допустить

Зав. кафедрой канд. техн. наук, профессор Комаров В. М.

ВКР передана в ГЭК \_\_ июня 2023 г.

Секретарь ГЭК канд. техн. наук, доцент каф. ВС Малышев Р. А.

Рыбинск 2023

Содержание

[Введение 3](#_Toc135399990)

[1 Информационно-патентный поиск 4](#_Toc135399991)

[1.1 EmoReact 4](#_Toc135399992)

[1.2 Affectiva 5](#_Toc135399993)

[1.3 FaceReader 6](#_Toc135399994)

[2 Анализ технического задания 7](#_Toc135399995)

[3 Техническое проектирование 8](#_Toc135399996)

[3.1 Анализ предметной области 8](#_Toc135399997)

[3.1.1 Основные эмоции 9](#_Toc135399998)

[3.1.2 Обзор методов распознавания эмоций по лицу 13](#_Toc135399999)

[3.2. Средства проектирования программного обеспечения 17](#_Toc135400000)

[3.3 Формирование данных для обучения модели 20](#_Toc135400001)

[3.4 Архитектура нейросети 24](#_Toc135400002)

[3.5 Декомпозиция 26](#_Toc135400003)

[3.6 Алгоритмизация программы 28](#_Toc135400004)

[Заключение 30](#_Toc135400005)

[Список использованных источников 31](#_Toc135400006)

# Введение

Современный уровень развития науки и техники предполагает необходимость создания и реализации программного обеспечения, способного распознавать и анализировать эмоциональное состояние людей. Оценка эмоций играет важную роль в множестве сфер, включая маркетинг, психологию, медицину и т.д.

Предлагается разработка программы, основанной на алгоритмах машинного обучения, которая будет способна распознавать и классифицировать различные эмоциональные состояния на основе анализа физических выражений лица, жестов и речи. Программа будет использовать алгоритмы машинного обучения и глубокого обучения, чтобы определить эмоциональное состояние с высокой точностью.

Такое программное обеспечение сможет быть применено в различных сферах, включая социальные исследования, медицину, разработку интерфейсов и даже виртуальную реальность. Оно поможет улучшить понимание эмоций людей, повысить эффективность коммуникации и создать более персонализированные и адаптивные системы.

Таким образом, разработка программного обеспечения для распознавания эмоций имеет большой потенциал в современном обществе, открывая новые возможности и применения для улучшения качества жизни людей. В результате использования программного обеспечения для распознавания эмоций достигается повышение эффективности, снижение субъективности и увеличение объективности в оценке эмоций. Это в свою очередь способствует более точному анализу и принятию соответствующих решений, основанных на эмоциональных данных.

# 1 Информационно-патентный поиск

Существует множество программ, которые способны распознавать эмоции по лицу человека. Однако, не все они одинаково эффективны и точны в своей работе. Некоторые из них могут работать только с определенными типами лиц, что делает их непригодными для использования в различных сферах, например, в медицине или в сфере общения.

Рассмотрим несколько программ, которые способны точно распознавать эмоции по лицу человека.

## 1.1 EmoReact

Программное обеспечение, разработанное специально для анализа эмоциональной реакции людей на различные стимулы. Оно предоставляет возможность исследования и обработки эмоциональных данных, а также управления устройством для сбора информации.

Одной из ключевых функций EmoReact является возможность обнаружения и классификации эмоций. Программа использует различные алгоритмы и методы машинного обучения, чтобы распознавать и интерпретировать эмоциональные состояния человека на основе его физиологических показателей, таких как выражение лица, тон голоса, пульс и другие.

EmoReact также предлагает набор инструментов для анализа эмоциональных данных. Например, программа позволяет визуализировать результаты анализа в виде графиков или диаграмм, что помогает исследователям и профессионалам в области психологии и маркетинга лучше понять эмоциональные реакции и предпочтения аудитории.

EmoReact поддерживает различные методы исследования, включая проведение экспериментов, опросов и наблюдений. Программа предоставляет удобный интерфейс для настройки экспериментов и сбора данных, а также позволяет анализировать полученные результаты и извлекать статистическую информацию о проявлении эмоций в определенных ситуациях.

Благодаря EmoReact и его возможностям анализа эмоций, исследователи, маркетологи и разработчики могут получить ценные инсайты о восприятии и реакции людей на различные продукты, услуги или контент. Это помогает улучшить качество продуктов и повысить их привлекательность для целевой аудитории, создавая более эмоционально привлекательные и востребованные решения.

## 1.2 Affectiva

Программное обеспечение, которое специально разработано для анализа и изучения эмоциональной выразительности людей с использованием технологий компьютерного зрения и машинного обучения. Оно предоставляет мощные инструменты для анализа эмоций на основе видео- и аудиоданных.

Главная особенность Affectiva заключается в его способности распознавать и классифицировать эмоции, выражаемые лицом людей на видео. Программа использует алгоритмы компьютерного зрения, чтобы автоматически анализировать выражение лица, мимику, движения глаз и другие физиологические показатели, связанные с эмоциональной выразительностью.

Affectiva также обеспечивает возможность анализа и классификации эмоциональной выразительности по голосу. Путем анализа тона, интонации и других акустических характеристик голоса, программа способна определить эмоциональное состояние человека, а также выделить особенности произнесенных слов и фраз.

Помимо этого, Affectiva предлагает широкий спектр инструментов для анализа эмоций и визуализации данных

Программа Affectiva также предоставляет возможность измерения эмоциональной архитектуры и взаимодействия людей в реальном времени. Благодаря интеграции с различными устройствами, такими как веб-камеры, мобильные устройства и сенсоры, Affectiva может непрерывно мониторить эмоциональные реакции пользователей в различных ситуациях. Это позволяет оптимизировать пользовательский опыт, улучшить коммуникацию и создать более эмоционально привлекательные взаимодействия.

Одним из сценариев использования Affectiva является эмоциональный анализ в области здравоохранения и психологии. Программа может помочь в диагностике и оценке эмоциональных расстройств, облегчая работу специалистов и повышая точность диагностики. Также она может быть использована в образовательных целях, помогая ученикам и студентам лучше понять свои эмоции и развивать навыки эмоционального интеллекта.

## 1.3 FaceReader

Программа FaceReader разработана компанией Noldus Information Technology и использует алгоритмы машинного обучения для распознавания семи базовых эмоций: радость, грусть, злость, страх, удивление, презрение и отвращение. Кроме того, программа может определять интенсивность выражения каждой эмоции на основе микроэкспрессий, а также FaceReader работает с помощью веб-камеры, которая подключается к компьютеру. После запуска программы, она начинает считывать данные с камеры и анализировать изображение лица человека. Затем программа автоматически распознает выражение лица и определяет эмоции, которые человек выражает в данный момент. определять пол и возраст человека.

FaceReader требует определенного уровня технических знаний и навыков для его установки, настройки и использования. Пользователям без опыта работы с программным обеспечением компьютерного зрения может потребоваться время и обучение, чтобы эффективно использовать все возможности FaceReader.

В целом, программа FaceReader предоставляет полезные инструменты для анализа эмоций на основе выражений лица, ее недостатки в точности распознавания, ограниченном применении, отсутствиисложности использования должны быть учтены при ее применении и интерпретации результатов.

Таблица 1 – Данные об аналогах разрабатываемой программы

|  |  |  |  |
| --- | --- | --- | --- |
| Параметр | EmoReact | Affectiva | FaceReader |
| Поддержка различных форматов | Да | Да | Да |
| Работа с камерой | Нет | Да | Да |
| Анализ изображения | Да | Да | Да |
| Машинное обучение | Да | Да | Да |
| Построение графиков | Да | Нет | Да |
| Формирование отчета | Да | Да | Да |
| Требуется высокопроизводительное оборудование | ДА | Нет | Да |

Рассмотренные программы аналоги также выполняют различные алгоритмы, и в каждой программе есть существенные недостатки, которые возможно избежать при разработке.

2 Анализ технического задания.

Из анализа технического задания следует, что программа должна:

- использовать видеопоток с вебкамеры;

- обеспечивать графическое отображение результатов работы;

В качестве среды разработки будет использоваться PyCharm 2023.

В качестве среды разработки будет использоваться Microsoft Visual Studio Community 2015.

Python-это высокоуровневый, интерпретируемый язык программирования. PyCharm - это среда разработки, созданная компанией JetBrains, которая предоставляет широкий набор инструментов для разработки на Python, включая редактор кода, отладчик, автодополнение, анализ кода и многое другое.

Библиотека стандартных модулей Python содержит множество полезных функций и классов, которые упрощают написание кода и расширяют возможности языка. Python также имеет богатый экосистему сторонних библиотек, которые позволяют решать широкий спектр задач, включая научные вычисления, машинное обучение, веб-разработку и многое другое.

К программным средствам предъявляется требование обязательно установленной операционной системы не ниже Windows 7.

К аппаратным средствам предъявляются следующие требования:

- процессор не ниже 1100 МГц;

- оперативная память, не менее 4 ГБ;

- наличие вебкамеры.

# 3 Техническое проектирование

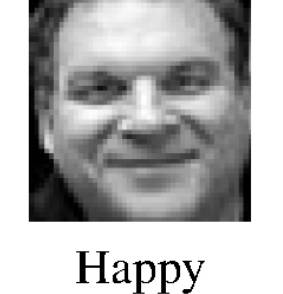
## 3.1 Анализ предметной области

Эмоции на лице являются выражением внутреннего эмоционального состояния человека. Лицо играет важную роль в межличностном взаимодействии, поскольку оно является главным средством выражения эмоций и передачи информации о состоянии человека окружающим.

Существует широкий спектр эмоций, которые могут быть выражены на лице. Основные эмоции, изученные в контексте распознавания эмоций, включают:

### 3.1.1 Основные эмоции

1. Радость. Радость выражается улыбкой, поднятыми уголками губ, морщинами возле глаз (так называемыми "морщинами глаз" или "морщинами улыбки"). Глаза могут быть более широко открытыми, а брови поднятыми.



* 1. Рисунок 3.1 – Радость

1. Грусть. Грусть выражается опущенными уголками губ, опущенными бровями и плоскостью между бровями. Глаза могут выглядеть взглядом, а рот может быть слегка открытым.



Рисунок 3.2 – Грусть

1. Злость.Злость выражается сжатыми губами, скрученными бровями и напряженностью мышц в области лба и вокруг глаз. Глаза могут быть суженными или смотреть с яростью.



Рисунок 3.3 – Злость

1. Страх. Страх выражается широко открытыми глазами, поднятыми бровями, расширенными ноздрями и открытым ртом. Кожа на лице может быть натянутой или бледной.

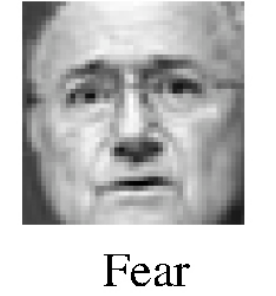


Рисунок 3.4 – Страх

1. Удивление. Удивление выражается широко открытыми глазами, поднятыми бровями и раскрытым ртом. Часто сопровождается взглядом и ожиданием.

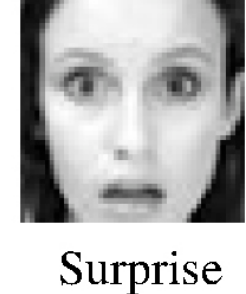


Рисунок 3.5 – Удивление

1. Отвращение. Отвращение выражается сжатыми губами, скрученными носом и опущенными бровями. Может сопровождаться морщинами на носу и между бровями.



Рисунок 3.6 – Отвращение

1. Нейтральное выражение. Нейтральное выражение лица не является явным выражением эмоции и характеризуется отсутствием ярко выраженных эмоциональных признаков. Лицо в нейтральном состоянии может быть расслабленным и спокойным.

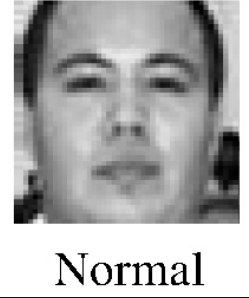


Рисунок 3.7 – Нормальное выражение

Однако стоит отметить, что эмоции на лице не всегда являются однозначными и могут иметь нюансы и вариации в выражении в зависимости от культурных, индивидуальных и контекстуальных факторов. Например, выражение эмоции радости может отличаться в различных культурах или у разных людей.

Исследования показывают, что распознавание эмоций по лицу является сложной задачей, поскольку выражение эмоций может быть динамичным и субъективным. Однако развитие компьютерного зрения и алгоритмов машинного обучения позволяет создавать модели и системы, способные автоматически распознавать и классифицировать эмоции на лице с высокой точностью.

Определение распознавания эмоций по лицу - это процесс автоматического определения эмоционального состояния человека на основе анализа его лица. Это включает в себя обнаружение и анализ различных физических признаков на лице, таких как выражение глаз, рта, бровей и других мимических движений. Распознавание эмоций по лицу может использоваться для определения широкого спектра эмоций, таких как радость, грусть, страх, злость, удивление и т.д.

Этот процесс становится все более популярным и актуальным во многих областях, таких как психология, медицина, реклама, безопасность и т.д. Например, в психологии, распознавание эмоций по лицу может помочь улучшить эмоциональную диагностику и лечение, в медицине - определить состояние больного, в рекламе - улучшить таргетирование рекламных кампаний, а в безопасности - обеспечить дополнительный уровень защиты путем обнаружения и анализа эмоциональных реакций людей.

Существуют различные методы для распознавания эмоций по лицу, такие как традиционный анализ мимики, выражения лица, глаз и т.д., а также использование компьютерного зрения и алгоритмов машинного обучения.

### 3.1.2 Обзор методов распознавания эмоций по лицу

Обзор существующих методов распознавания эмоций по лицу позволяет ознакомиться с различными подходами и техниками, используемыми в этой предметной области.

Традиционные методы:

Анализ мимики лица: Этот метод основан на изучении различных выражений лица, таких как улыбка, грусть, гнев и др. Анализируются мимические движения глаз, бровей, рта и других частей лица. Например, поднятые брови и улыбка могут указывать на радость, а скрученные брови и сжатые губы могут указывать на злость. Однако этот метод имеет ограниченную точность и требует определения и использования набора правил, что может быть сложно и субъективно.

Извлечение геометрических признаков: Данный метод основан на измерении геометрических параметров лица, таких как расстояние между ключевыми точками на лице. Например, можно измерять расстояние между глазами, ширину губ и другие параметры. По этим параметрам можно определить эмоциональное состояние человека. Однако этот метод также имеет свои ограничения, так как не учитывает текстуры и детали лица, которые могут быть важными для распознавания эмоций.

Преимущества:

- Относительно простая реализация и интерпретируемость.

- Некоторые мимические выражения имеют явные связи с определенными эмоциями.

Ограничения:

- Ограниченная точность, особенно при сложных или неоднозначных выражениях.

- Требует определения явных правил и шаблонов для классификации эмоций, что может быть сложным и субъективным.

Компьютерное зрение и алгоритмы машинного обучения:

Методы основанные на признаках: В этом методе используются алгоритмы компьютерного зрения для выделения и анализа различных признаков на лице. Например, можно извлекать текстурные характеристики или цветовые характеристики изображения лица и использовать их для обучения классификаторов эмоций. Этот подход требует ручного выбора и извлечения признаков, что может быть сложно и требует экспертных знаний.

Сверточные нейронные сети (CNN): CNN-архитектуры являются мощным инструментом в распознавании эмоций по лицу. Они автоматически извлекают признаки из изображений лиц с различных уровней детализации, что помогает лучше представить эмоциональные характеристики. CNN-модели обучаются на больших объемах размеченных данных, что позволяет им обнаруживать сложные шаблоны и закономерности в данных. Они способны достичь высокой точности в распознавании эмоций.

Рекуррентные нейронные сети (RNN): RNN-архитектуры применяются для анализа последовательности временных данных, например, для анализа временных изменений выражений лица. RNN-модели могут учитывать контекст и динамику в данных, что позволяет им лучше моделировать эмоциональные состояния. Этот подход особенно полезен при анализе видео, где можно наблюдать изменения выражений лица во времени.

Преимущества:

- Более высокая точность в распознавании эмоций по лицу по сравнению с традиционными методами.

- Способность автоматически извлекать и обрабатывать сложные признаки и шаблоны, которые могут быть сложными для человека.

Ограничения:

- Требуется большой объем размеченных данных для обучения моделей машинного обучения.

- Могут возникать проблемы с переносимостью моделей на различные условия освещения, позы или разрешения изображений.

Использование дополнительных источников данных:

Интеграция голосовых сигналов: В некоторых методах распознавания эмоций по лицу используется комбинация анализа лица и анализа голосовых сигналов. Голос может предоставить дополнительную информацию о состоянии эмоций человека, такую как тон голоса, интонации и ритм речи. Интеграция данных о лице и голосе может повысить точность распознавания эмоций.

Использование данных сенсоров: Некоторые исследования включают использование дополнительных сенсорных данных, таких как пульс, электрокожные реакции и другие биологические показатели. Эти данные могут предоставить дополнительную информацию о состоянии эмоций человека и помочь улучшить точность распознавания.

Обзор этих методов позволяет понять различные подходы и техники, используемые в области распознавания эмоций по лицу. Каждый метод имеет свои преимущества и ограничения, и выбор метода зависит от конкретной задачи и доступных ресурсов.

Преимущества:

- Повышение точности распознавания эмоций благодаря интеграции информации из различных источников.

- Дополнительные данные сенсоров или голосовых сигналов могут предоставить дополнительную информацию о состоянии эмоций.

Ограничения:

- Необходимость доступа к дополнительным источникам данных, таким как микрофон или сенсоры.

- Возможна сложность в синхронизации данных разных источников и их корректной интерпретации.

В целом, компьютерное зрение и алгоритмы машинного обучения, включая сверточные нейронные сети и рекуррентные нейронные сети, демонстрируют высокую точность и широкий потенциал в распознавании эмоций по лицу.

## 3.2. Средства проектирования программного обеспечения

Поскольку разработка будет вестись с использованием машинного обучения, то в качестве языка разработки будет использоваться Python 3.11. Одними из самых главных преимуществ данного языка являются:

Улучшенная производительность: В Python 3.11 произошли значительные улучшения производительности благодаря оптимизации интерпретатора и улучшению работы с памятью. Это позволяет значительно ускорить выполнение программ, особенно при работе с большими объемами данных.

Расширяемость: Python 3.11 по-прежнему поддерживает возможность подключения библиотек, написанных на языке C, что позволяет использовать оптимизированные функции и ускорить выполнение программ. Большое количество библиотек и фреймворков для машинного обучения, таких как TensorFlow, PyTorch и scikit-learn, доступны в Python и предоставляют мощные инструменты для разработки моделей.

Удобство использования: Python имеет низкий порог вхождения и понятный синтаксис, что делает его привлекательным для новичков в программировании. Большое количество ресурсов, документации и сообщества разработчиков делает процесс изучения и разработки на Python более доступным.

Динамическая типизация: Python 3.11 поддерживает динамическую типизацию, что означает, что переменные могут автоматически адаптироваться к типу данных, который им присваивается. Это облегчает и ускоряет процесс разработки, особенно при работе с большими объемами кода.

Кроссплатформенность: Python является кроссплатформенным языком программирования, что означает, что код, написанный на Python, может быть запущен на различных операционных системах, таких как Windows, macOS и Linux, без необходимости внесения изменений в код.

Широкая область применения: Python широко используется в машинном обучении и анализе данных благодаря своей простоте, гибкости и мощным библиотекам. Он также применяется во многих других областях, таких как веб-разработка, научные исследования, автоматизация и разработка игр.

В качестве среды разработки будет использоваться PyCharm 2023.1, разработанная компанией JetBrains на основе IntelliJ IDEA. PyCharm предоставляет мощный и удобный редактор кода со всеми полезными функциями, включая:

Подсветка синтаксиса: PyCharm обладает расширенной подсветкой синтаксиса для языка Python, что облегчает чтение и понимание кода.

Поддержка рефакторинга: С помощью PyCharm можно выполнять различные рефакторинговые операции, такие как переименование переменных, методов и классов, извлечение методов, инлайнинг и многое другое. Это помогает улучшить качество кода и упростить его поддержку.

Анализ кода "на лету": PyCharm осуществляет анализ кода в реальном времени и предлагает подсказки, предупреждения об ошибках, автозаполнение и другие полезные функции, которые помогают улучшить производительность и точность разработки.

Интеграция управления версиями: PyCharm интегрируется с популярными системами управления версиями, такими как Git, Mercurial, Subversion и другими. Вы можете выполнять коммиты, обновления, слияния и другие операции непосредственно из среды разработки.

Удаленная отладка: PyCharm позволяет подключаться к удаленным интерпретаторам Python и выполнять отладку удаленного кода. Это особенно полезно, когда ваши приложения выполняются на удаленном сервере или в виртуальной среде.

PyCharm 2023.1 обеспечивает комфортную и продуктивную среду разработки для проектов на языке Python, сделав процесс разработки более эффективным и удобным.

Основными библиотеками, которые будут использоваться в данной работе для разработки и обучения моделей распознавания эмоций по лицу, являются:

TensorFlow - это открытая программная библиотека для машинного обучения, разработанная компанией Google. Она предоставляет широкий спектр инструментов и возможностей для построения и тренировки нейронных сетей. TensorFlow позволяет решать различные задачи, включая классификацию образов, в том числе распознавание эмоций на лицах. Библиотека обладает мощными инструментами для создания и обучения глубоких нейронных сетей, что позволяет достичь высокой точности и качества распознавания.

NumPy - это библиотека с открытым исходным кодом для языка Python, предоставляющая вычислительные алгоритмы и удобные структуры данных для работы с большими объемами данных. В контексте распознавания эмоций по лицу, NumPy может использоваться для эффективной обработки изображений и манипуляции пикселями. Она предоставляет мощные функциональности для работы с многомерными массивами данных, что облегчает выполнение различных математических операций, необходимых при обработке и анализе изображений.

OpenCV (Open Source Computer Vision Library) - это библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения. Она предоставляет множество функций и инструментов для работы с изображениями, включая загрузку, обработку, анализ и визуализацию. В контексте распознавания эмоций по лицу, OpenCV может быть использована для обнаружения и выделения лиц на изображениях, а также для извлечения характеристик лица, таких как расположение глаз, носа и рта. Библиотека также предоставляет алгоритмы для обработки изображений, фильтрации, сегментации и многое другое.

Keras - это высокоуровневый API нейронной сети, работающий поверх TensorFlow. Он предоставляет простой и интуитивно понятный интерфейс для создания и обучения нейронных сетей. Keras упрощает процесс разработки моделей машинного обучения, включая модели для распознавания эмоций по лицу. Он предлагает различные предварительно обученные модели и слои, которые можно использовать для построения и настройки моделей быстро и эффективно.

Эти библиотеки являются популярными инструментами в области машинного обучения и компьютерного зрения, и их использование в разработке моделей для распознавания эмоций по лицу обеспечивает высокую гибкость, эффективность и точность работы.

## 3.3 Формирование данных для обучения модели

Данные для обучения взяты с сервиса Kaggle.  
Kaggle - это онлайн-платформа, предназначенная для соревнований и совместной работы в области анализа данных и машинного обучения. Сервис предоставляет доступ к разнообразным наборам данных реального мира. Это позволяет участникам практиковаться на реальных проблемах и данных, что является ценным опытом. Многие наборы данных на платформе предоставлены компаниями и организациями, которые ищут помощи в анализе и решении своих задач. Kaggle предлагает обширную библиотеку учебных материалов, руководств, которые помогают новичкам в мире анализа данных и машинного обучения. Это отличный ресурс для изучения новых концепций, алгоритмов и практического применения моделей машинного обучения.

Загрузка данных: Сначала необходимо собрать и загрузить изображения из kaggle. Получилось загрузить 35887 файлов.

Изучение структуры данных: после загрузки данных необходимо изучить их структуру и разбить их на разные наборы по эмоциям. Получившиеся наборы: ”angry”, ”disgust”, “fear”, ”happy”, ”neutral”, ”sad”, ”surprised”.

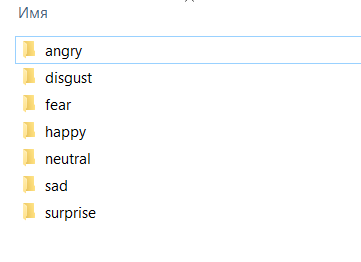


Рисунок 3.8 – Наборы данных для обучения модели



Рисунок 3.9 – Пример данных в папке «happy»

Далее требуется предварительная обработка данных. Предварительная обработка данных является важным этапом в обучении модели машинного обучения. Она включает ряд шагов, направленных на обеспечение качественного обучения модели. Рассмотрим подробнее каждый из этих шагов:

Изменение размера изображений: проверка размеров изображений в датасете и приведение их к одному размеру является важным шагом. Это необходимо для обеспечения согласованности данных перед подачей их на вход модели. Различные алгоритмы машинного обучения требуют одинаковый размер входных данных, поэтому изменение размера изображений помогает унифицировать данные.

Нормализация значений пикселей: нормализация значений пикселей является важной частью предварительной обработки изображений. Целью нормализации является приведение значений пикселей к определенному диапазону, чтобы обеспечить стабильность и согласованность данных. Обычно значения пикселей нормализуются в диапазоне от 0 до 1 или от -1 до 1. Это может быть особенно полезно для моделей, использующих функции активации, которые действуют в определенном диапазоне.

Устранение шума: если в данных присутствует шум, например, артефакты или искажения, необходимо применить методы фильтрации или обработки изображений для его удаления. Это может включать применение различных фильтров, сглаживание изображений, удаление выбросов и другие техники для устранения нежелательных эффектов, которые могут повлиять на процесс обучения модели.

Все эти шаги предварительной обработки данных помогают улучшить качество обучения модели и повысить ее способность обобщать данные.

После предварительной обработки требуется разбить данные на обучающую и тестовые выборки. Разделение данных на обучающую и тестовую выборки является важным шагом при разработке моделей машинного обучения. Основная цель разделения данных заключается в оценке производительности модели на новых, ранее не виденных ей данных. Вот несколько причин, почему разделение данных является важным:

Оценка обобщающей способности модели: Разделение данных позволяет оценить, насколько хорошо модель способна обобщать свои знания на новые данные. Обучение модели на обучающей выборке позволяет ей изучить общие закономерности в данных, в то время как тестирование модели на независимой тестовой выборке позволяет оценить ее способность делать предсказания на новых данных. Если модель показывает хорошие результаты как на обучающей, так и на тестовой выборках, можно сделать вывод о ее способности обобщать и обнаруживать паттерны.

Предотвращение переобучения: Разделение данных помогает предотвратить переобучение модели. Переобучение возникает, когда модель слишком хорошо запоминает обучающие данные, вместо того чтобы обобщать их. Если бы мы не разделяли данные и использовали бы все доступные данные для обучения модели, она могла бы просто запомнить обучающие примеры, не улавливая общие закономерности. Разделение данных позволяет оценить, насколько хорошо модель обобщает данные и предотвратить переобучение.

Выбор модели и настройка параметров: Разделение данных позволяет проводить эксперименты с различными моделями и настраивать их параметры на основе производительности на тестовой выборке.

Обнаружение проблем и улучшение модели: Разделение данных также позволяет обнаружить проблемы в модели и улучшить ее. Если модель показывает низкую производительность на тестовой выборке, это может указывать на проблемы в обобщающей способности модели или на несбалансированность данных. Это может быть сигналом для внесения изменений в модель или проведения дополнительных манипуляций с данными.

Разделение данных на обучающую и тестовую выборки позволяет оценить производительность модели на новых данных, предотвратить переобучение, выбрать лучшую модель и настроить ее параметры. Разделение проведено случайным образом в примерном соотношении 80% и 20% (28709 и 7178 файлов).



Рисунок 3.10 – Выборки для работы с моделью

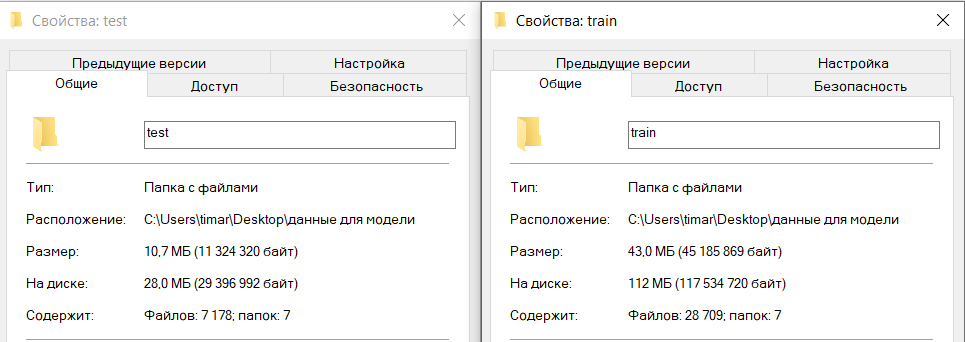


Рисунок 3.11 – Количество файлов в выборках

## 3.4 Архитектура нейросети

Сверточные нейронные сети (Convolutional Neural Networks, CNN) широко используются для задач компьютерного зрения, включая распознавание эмоций по видео. Необходимо изучить архитектуру такого вида нейронных сетей.

Входные данные: Видеофайл, содержащий последовательность кадров, служит входными данными для модели. Каждый кадр представляет собой изображение.

Предварительная обработка: Перед подачей на вход модели, каждый кадр может быть подвергнут предварительной обработке для улучшения качества данных. Это может включать изменение размера кадров, нормализацию яркости и контраста и применение фильтров для сглаживания или усиления определенных особенностей.

Сверточные слои: Сверточные слои играют ключевую роль в архитектуре сверточной модели. Они используются для извлечения визуальных признаков из каждого кадра. Каждый сверточный слой содержит набор фильтров (ядер), которые сканируют изображение и выделяют различные визуальные паттерны. Например, низкоуровневые фильтры могут обнаруживать границы и текстуры, а более высокоуровневые фильтры могут обнаруживать более сложные объекты или лица.

Пулинговые слои: После каждого сверточного слоя обычно следует пулинговый слой. Пулинг (субдискретизация) используется для уменьшения пространственных размеров карт признаков и повышения инвариантности к малым трансляциям объектов. Наиболее распространенным методом пулинга является слой максимального пулинга (MaxPooling), который выбирает наибольший элемент из каждой области.

Полносвязные слои: После нескольких сверточных и пулинговых слоев, полученные признаки подаются на полносвязные слои. Полносвязные слои соединяют все признаки вместе и выполняют классификацию. Каждый нейрон в последнем полносвязном слое представляет собой класс эмоции, который модель пытается распознать.

Функция активации и выходной слой: Каждый нейрон в сверточной модели обычно имеет функцию активации, которая вводит нелинейность в модель. Для задачи классификации эмоций часто используется функция активации Softmax в выходном слое, которая представляет вероятности принадлежности каждому классу эмоции.

Обучение: Модель обучается на размеченном наборе данных, где для каждого видео указана соответствующая эмоция. В процессе обучения модель корректирует веса своих слоев, минимизируя функцию потерь между предсказанными эмоциями и истинными эмоциями в обучающем наборе.

После обучения модель может быть применена для распознавания эмоций на новых видео, где она классифицирует эмоции для каждого кадра и может сделать вывод о эмоциональном состоянии в течение всего видео.

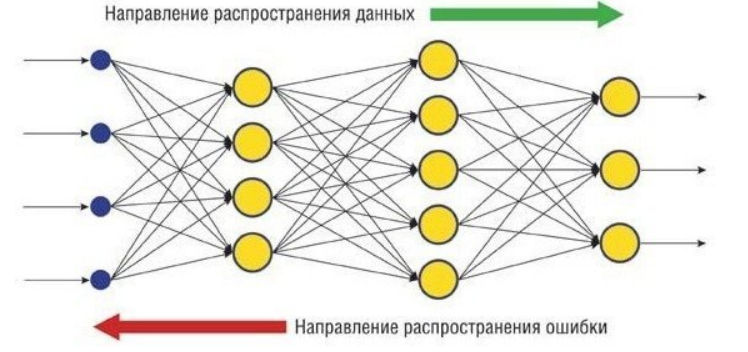


Рисунок 3.12 – пример архитектуры нейросети

## 3.5 Декомпозиция

Задача декомпозиции состоит в том, чтобы разбить сложную программу на ряд более простых задач. Декомпозиция происходит в несколько этапов. На первом этапе вся программа делится на две задачи: интерфейс и математическое ядро. Интерфейс связывает внешние данные и внутренние, а математическое ядро выполняет необходимые операции над внутренними данными.

На рисунке 3.12 представлена исходная схема программы распознавания эмоций.

В качестве входных данных будет использоваться изображение с камеры и обученная модель в формате \*.h5, а выходными данными является: кадр с данными, полученными об эмоции на входном изображении.

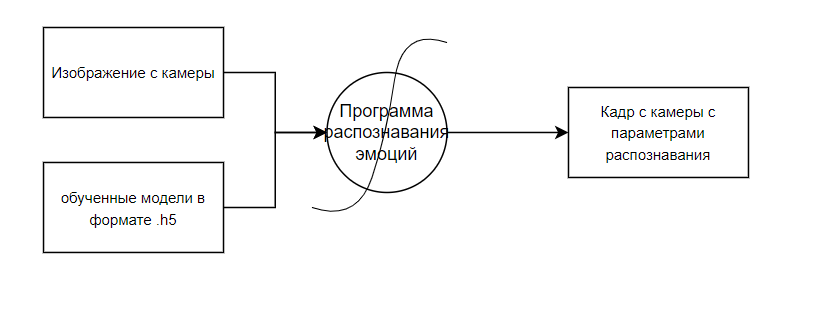


Рисунок 3.13 - Исходная схема программы распознавания эмоций

На этапе черного ящика задача «Программа распознавания эмоций» разбивается на более простые. Статическая модель программы после первого этапа декомпозиции представлена на рисунке 3.12.

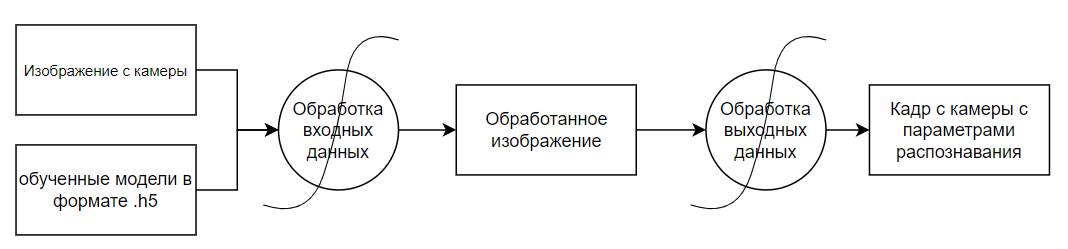


Рисунок 3.14 – Статическая модель программы после первого этапа декомпозиции

Подзадачи «Обработка входной информации» и «Обработка выходной информации» требуют дальнейшей декомпозиции. Статическая модель программы после второго этапа декомпозиции представлена на рисунке 3.13.

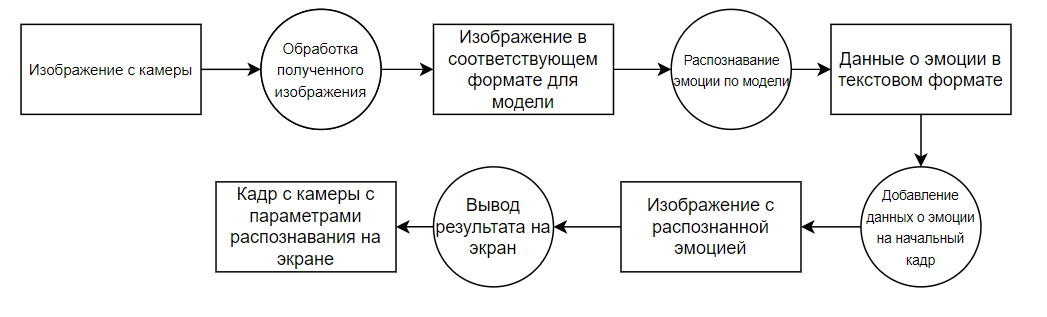


Рисунок 3.15 – Статическая модель программы после второго этапа

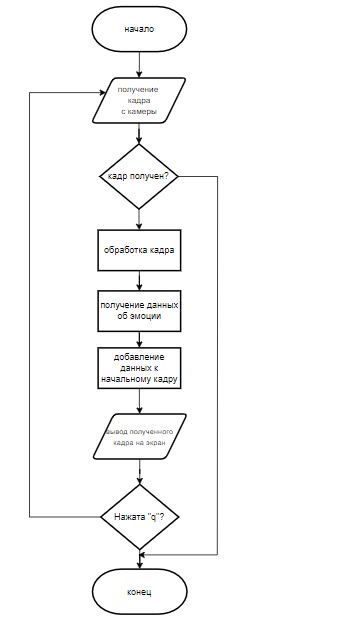
декомпозиции

После второго этапа декомпозиции вся программа разбита на логически завершенные подзадачи. На данном этапе можно закончить процесс декомпозиции.

## 3.6 Алгоритмизация программы

В данном подразделе описаны основные алгоритмы программы.

На рисунке 3.16 представлена блок-схема алгоритма функции main(), в которой выполняется получение изображения с камеры и выводится результат работы программы. Создается поток для доступа к камере и программа получает кадр из потока. Производится проверка на наличие значения в переменной, если значения нет выход из функции, иначе загрузка происходит обработка полученного изображения. Далее кадр попадает в нейросеть, которая возвращает данные об эмоции в текстовом формате. Эти данные добавляются к начальному кадру и выводятся на экран. Происходит проверка на нажатие кнопки “Q”: если кнопка была нажата – завершение работы программы, нет – алгоритм начинается с начала.



Нет

Нет

Да

Да

Рисунок 3.16 - Блок-схема работы программы

# Заключение

В ходе выполнения преддипломной практики был произведен патентно-информационный поиск, были найдены патенты и программы-аналоги, схожие по функционалу, а также выделены их достоинства и недостатки. Была начата разработка программного обеспечения, в ходе которой были выполнены некоторые функции согласно техническому заданию:

- углубленное изучение теоретических и практических сведений, необходимых для проектирования объекта ВКР;

- декомпозиция программы;

- составление алгоритма работы программы;

- сбор и подготовка данных для обучения модели;

Следующим шагом является обучение нейронной сети по собранным данным для дальнейшего использования в программе.

# Список использованных источников

1. Noldus / [Электронный ресурс] / Режим доступа: - URL: https://www.noldus.com// (дата обращения: 05.05.2023 г.).

2. Жегалло А. Перцептивная категоризация выражений лица: Пособие/ Издатель: Litres, 2022 – 375 с.

3.  Affectiva [Электронный ресурс] / Режим доступа: URL: https://www.affectiva.com/ (дата обращения: 16.05.2023 г.)

4. Гусаров А. В. Оформление текстовой документации ВКР: Пособие / Под ред. Э. И. Семенова. – Рыбинск: РГАТУ имени П. А. Соловьева, 2012. – 112 с.