

# «Детекция эмоций человека»

Итоговый проект по курсу «Data Scientist»

**Александр Дубовик**  
Студент группы: DS-72



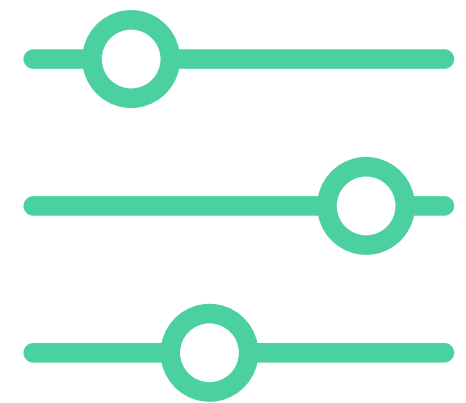
# Содержание

- 1 Постановка задачи
- 2 Анализ данных
- 3 Методика реализации
- 4 Итоги обучения модели и реализации детекции
- 5 Выводы
- 6 Список источников



# Постановка задачи

1



# Постановка задачи

1

## Исходная задача

Первая часть - реализация искусственной нейронной сети с использованием компьютерного зрения для классификации и детекции объектов.

Вторая часть — это определение на видео в режиме реального времени эмоции субъектов. Объекты детекции лица людей, классификация — определение 7 эмоций.

Архитектура модели – CNN.

2

## Актуальность задачи, её место в предметной области

Распознавание (детекция) эмоций и возможность предугадывания тех или иных эмоций человека, в настоящее время очень востребовано. В медицине, психологии – идентификация эмоций при эмоциональных расстройствах. В торговых точках для получения эмоционального отклика клиента. В рамках безопасности общественных мест, выявление эмоционально неустойчивых субъектов, в том числе потенциальных нарушителей. Для роботов с ИИ для идентификации эмоционального отклика оппонента и корректировки генерируемых ответов, либо вопросов.

3

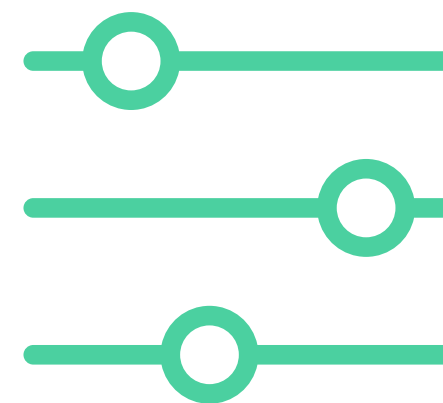
## Целевая метрика

Метрика качества: F1-мера (*F1-score*) для многоклассовой классификации.



# Анализ данных

2



# Анализ данных

**Датасет:** «FER2013»

<https://www.kaggle.com/datasets/ashishpatel26/facial-expression-recognitionferchallenge/data>

**Количество изображений:** 35887 шт.

**Количество изображений в обучающей выборке:** 28709 шт.

**Количество изображений в тестовой выборке:** 7178 шт.

**Количество классов:** 7

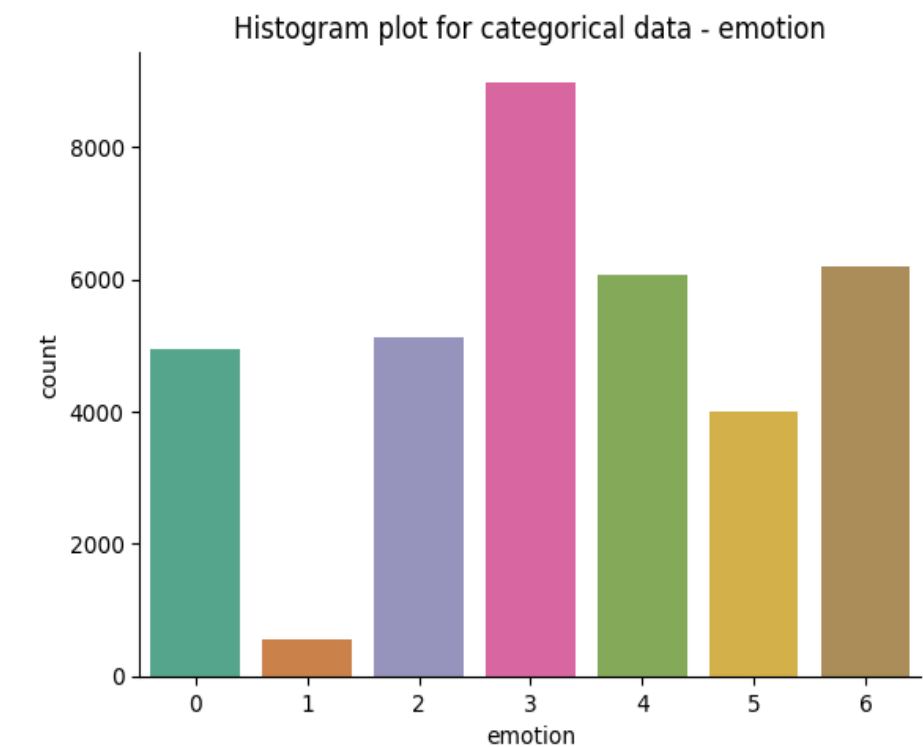
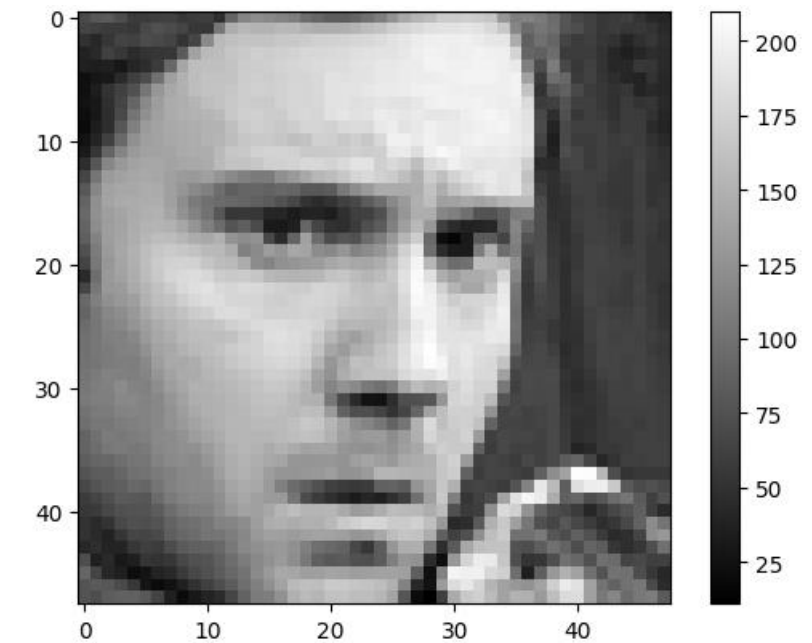
**Классы:** anger, disgust, fear, happiness, sadness, surprise, neutral

**Описание:** Данные состоят из изображений лиц размером 48x48 пикселей в оттенках серого. Столбец «emotion» содержит числовой код от 0 до 6 включительно для эмоции, присутствующей на изображении. Столбец «pixels» содержит строку, заключенную в кавычки для каждого изображения. Содержимое этой строки представляет собой значения пикселей, разделенные пробелами, в порядке возрастания строк.

**Особенности:** Лица были автоматически зарегистрированы таким образом, чтобы лицо располагалось более или менее по центру и занимало примерно одинаковое пространство на каждом изображении.

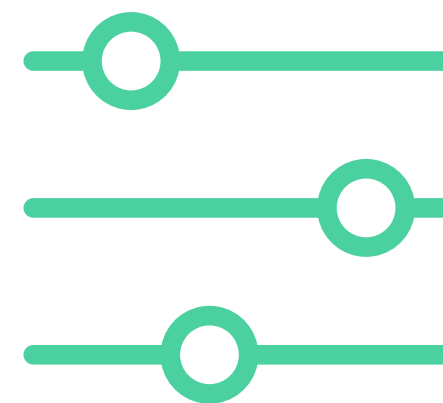
В датасете несбалансированное распределение классов, т. е. классы, которые мы пытаемся предсказать, представлены неравномерно.

[Анализ данных](#)



# Методика реализации

3





# Методика реализации

## ***Шаги реализации модели:***

Предварительная обработка данных: Кодирование меток, one-hot кодирование, создание словаря отображения, преобразование данных изображений в трехмерный массив, стекинг изображений, разделение данных на обучающую и валидационную выборки, нормализация данных.

1. Обучение модели с учетом весов класса: Вычисление весов классов, разбиение данных на обучающую и тестовую выборки, преобразование меток классов, восстановление исходного целевого массива, вычисление веса классов, конвертация веса в словарь.
2. Обучение модели с переобучением данных с использованием SMOTE: Преобразование изображений в векторы, проверка совпадения количества образцов, применение SMOTE, изменение формы данных после SMOTE, приведение типов данных для `y_resampled`.
3. Определение и обучение модели: Конструкция модели CNN, включая сверточные слои, слои максимального пула, плотные слои, методы регуляризации, компиляция модели, настройка обратных вызовов, обучение модели.
4. Обнаружение лиц на видео с веб-камеры и определение эмоций: Инициализация классификатора Haar cascades, открытие видеопотока с веб-камеры, создание директории для сохранения изображений лиц с распознанными эмоциями, цикл чтения кадров с веб-камеры, обнаружение лиц, определение эмоций, отображение исходного кадра с возможной отметкой эмоции на лице.

## ***Функции активации используемые в исследовании ReLu и ELU:***

1. ReLU (Rectified Linear Unit): Нелинейная функция активации, преобразует входное значение в значение от 0 до положительной бесконечности. Если входное значение меньше или равно нулю, то ReLU выдает ноль, в противном случае - входное значение.
2. ELU (Exponential Linear Unit): Измененная версия ReLU, которая помогает ускорить обучение глубоких нейронных сетей и справляется с проблемой "мертвых нейронов". ELU использует экспоненциальную функцию, чтобы получить значение, которое ближе к нулю, чем значение, возвращаемое ReLU.

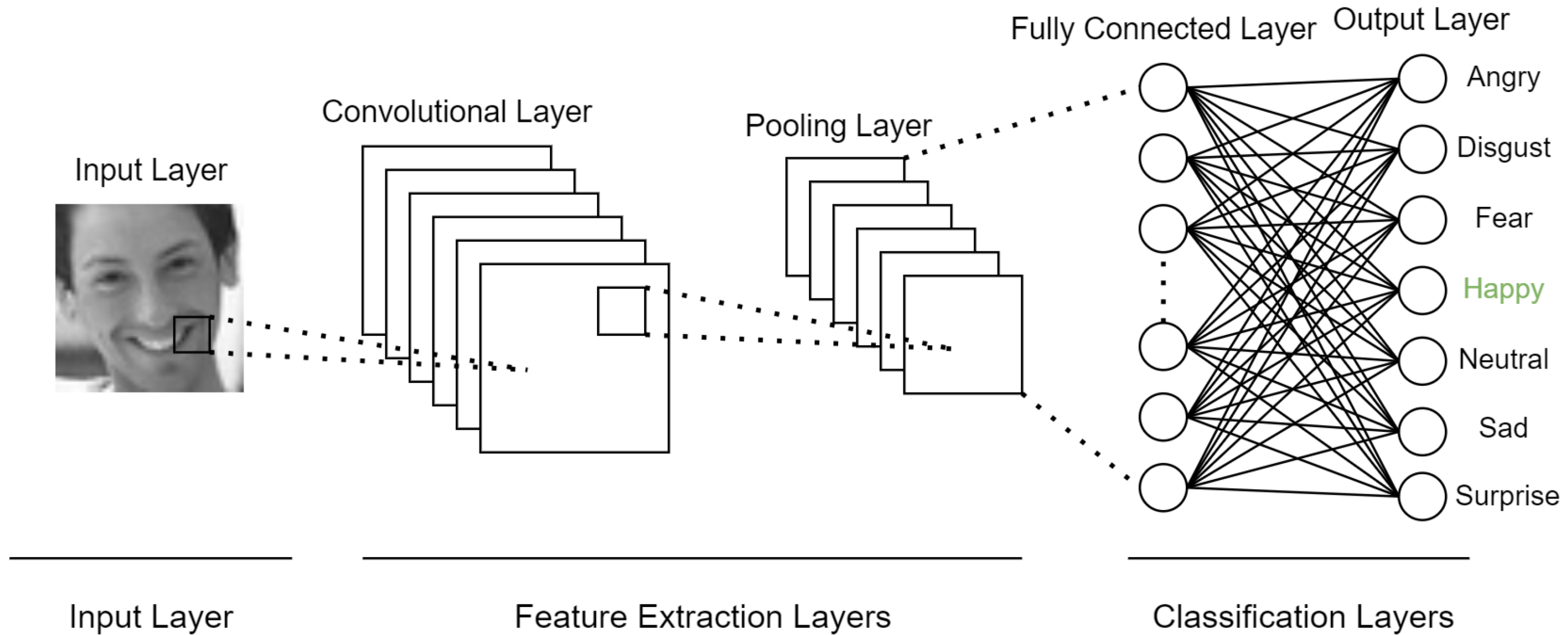
Ссылка на репозиторий с проведенными исследованиями:

[https://github.com/dubovik-alexander/Netology-DS/tree/main/Diploma/final/py\\_files](https://github.com/dubovik-alexander/Netology-DS/tree/main/Diploma/final/py_files)

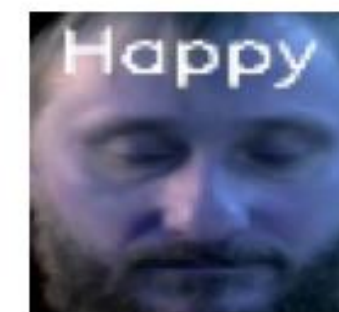
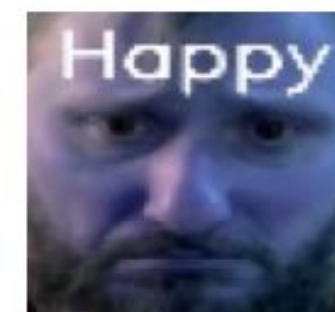
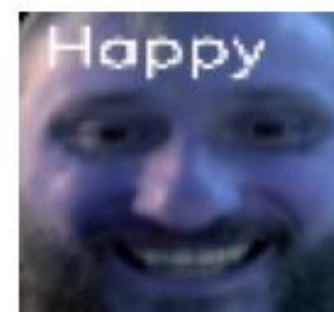
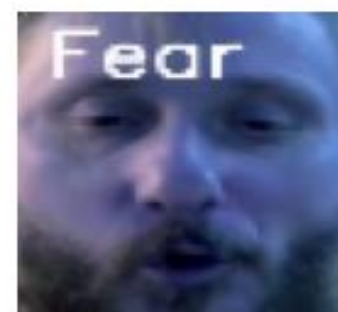




# Методика реализации

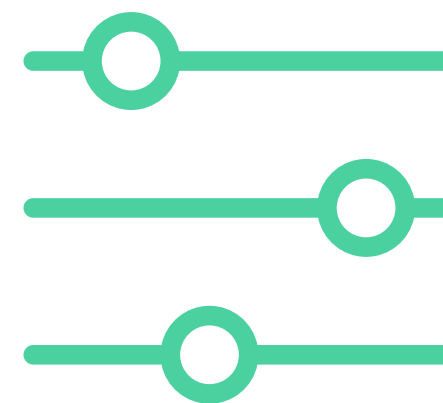


# Методика реализации



# Итоги обучения модели и реализации детекции

4



# Итоги обучения модели и реализации детекции

## *Итоги обучения модели:*

По результатам исследований, обучение модели CNN с активацией «ELU» с процессом переобучения данных с использованием метода Synthetic Minority Over-sampling Technique (SMOTE) для решения проблемы дисбаланса классов в наборе данных показала себя лучше остальных, хоть и не значительно. Во всех обучении было использовано 50 эпох, лучшая модель достигла метрик micro avg - 0.62, macro avg - 0.62, weighted avg - 0.62. Метрики по отдельным классам получились соответственно: 0 – 0.53, 1 – 0.67, 2 – 0.50, 3 – 0.79, 4 – 0.50, 5 – 0.76, 6 – 0.57. Модель лучше обучилась распознавать те объекты, которые больше представлены в обучающей выборке.

## *Обнаружение лиц на видео с веб-камеры и определение эмоций на этих лицах с помощью предварительно обученной модели.*

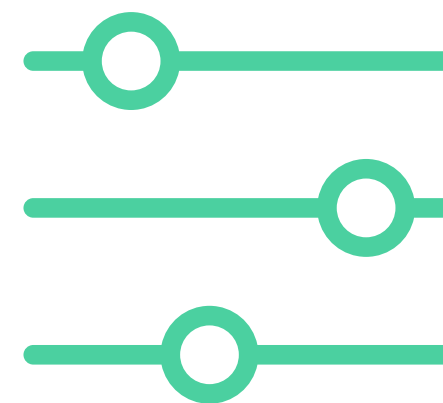
По итогам применения модели к видеопотоку, с большой долей вероятности, можно сделать вывод о том, что модель выучила основные паттерны, для определения эмоций. В тоже время остается еще много возможностей для улучшения модели. Например, можно заметить ложные срабатывания, когда модель определяет разные эмоции на одной и той же картинке. Возможно необходимо доработать не только модель, но и код по определению эмоции в режиме реального времени. Возможные варианты решения данной проблемы предложены далее.

Ознакомиться с результатами работы модели можно в репозитории: [face\\_images\\_final](#)



# Выводы

5





# Выводы

В ходе проведенного исследования была получена модель детекции эмоций, а так же написан код для обнаружения лиц на видео с веб-камеры и определение эмоций на этих лицах с помощью предварительно обученной модели. Данная модель основана на архитектуре CNN. Целевые метрики достигли следующих результатов: micro avg - 0.62, macro avg - 0.62, weighted avg - 0.62. Метрики по отдельным классам : 0 – 0.53, 1 – 0.67, 2 – 0.50, 3 – 0.79, 4 – 0.50, 5 – 0.76, 6 – 0.57.

## ***Для дальнейшей настройки модели рекомендуется:***

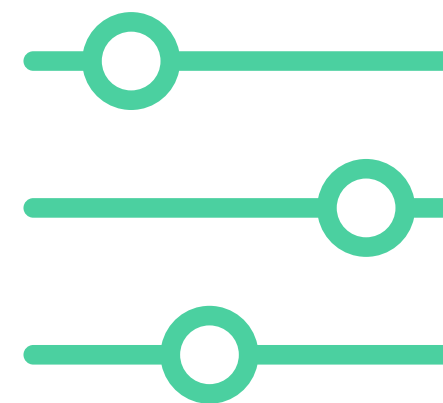
- Увеличить размер обучающей выборки. 35887 изображений не достаточный объем для обучения нейронной сети;
- Добавление (выравнивание) данных по тем классам, которые имеют самый маленький набор данных, физическим методом, т.е. добавление изображений.
- Проводить обучение на изображениях с более высоким разрешением;
- Разнообразить изображения в обучающей выборке, добавить объекты с разными масштабами, относительно размеров изображения;
- Обучить модель на большем количестве эпох;
- Использование [предварительно обученных моделей](#): таких как VGG19, ResNet, MobileNet или EfficientNet, которые уже были обучены на больших наборах данных ImageNet. Эти модели могут служить хорошим стартом и сократить время обучения.
- Data Augmentation: применение техник увеличения объема данных, таких как повороты, масштабирование, перевороты и другие трансформации, чтобы увеличить разнообразие обучающих примеров и сделать модель более устойчивой к изменениям.
- Fine-tuning: после использования предварительно обученной модели, можно продолжить обучение на специальном наборе данных для задачи определения эмоций. Это позволяет модели адаптироваться к новым данным без полного обучения с нуля.
- Ensemble Methods: использование комбинации нескольких моделей для повышения точности прогнозов. Например, можно обучить несколько различных архитектур CNN и усреднить их прогнозы.
- Regularization Techniques: использовать параметры L1 и L2 в обучении моделей сверточных нейронных сетей (CNN) для регуляризации весов сети. Регуляризация помогает предотвратить переобучение, уменьшая сложность модели путем добавления штрафных функций к функции потерь. Так же можно исследовать другие методы регуляризации, такие как dropout или batch normalization, для дальнейшего улучшения обобщающей способности модели.

В целом обнаружение лица и детекция эмоции соответствует ожиданиям, но для получения более точных результатов обнаружение лиц на видео с веб-камеры и определение эмоций на этих лицах с помощью предварительно обученной модели, помимо дальнейшей настройки модели, доработать кода так, чтобы была задержка по времени обработки или определение эмоции только при смене эмоции (выражения лица). Либо детекцией изменения положения лица.



# Список источников

6





# Список источников

- Датасет, используемый в работе - <https://www.kaggle.com/datasets/ashishpatel26/facial-expression-recognitionferchallenge/data>
- Исследование возможности детектирования эмоций, А. Г. Лойко кафедра инженерной кибернетики НИТУ «МИСиС» Москва, Россия. - [https://www.researchgate.net/publication/377271090\\_Issledovanie\\_vozmoznosti\\_detektirovania\\_emocij](https://www.researchgate.net/publication/377271090_Issledovanie_vozmoznosti_detektirovania_emocij)
- Выбор слоя активации в нейронных сетях: как правильно выбрать для вашей задачи <https://habr.com/ru/articles/727506/>
- Классификация изображений CNN в TensorFlow с шагами и примерами <https://www.guru99.com/ru/convnet-tensorflow-image-classification.html>
- Распознавание лиц при помощи Python и OpenCV <https://pythonist.ru/raspoznavanie-licz-pri-pomoshhi-python-i-opencv/>
- Лекции Нетология. Курс «Data Scientist»



**Спасибо за  
внимание!**

