

# Алгоритмы распознавания невербальной информации

Эфендиев Камал Вугар оглу

Научный руководитель: м.н.с., кандидат физико-математических наук  
Дергач Пётр Сергеевич

ФИЛИАЛ МГУ имени М.В. ЛОМОНОСОВА в городе БАКУ

09.06.2022

# Постановка задачи

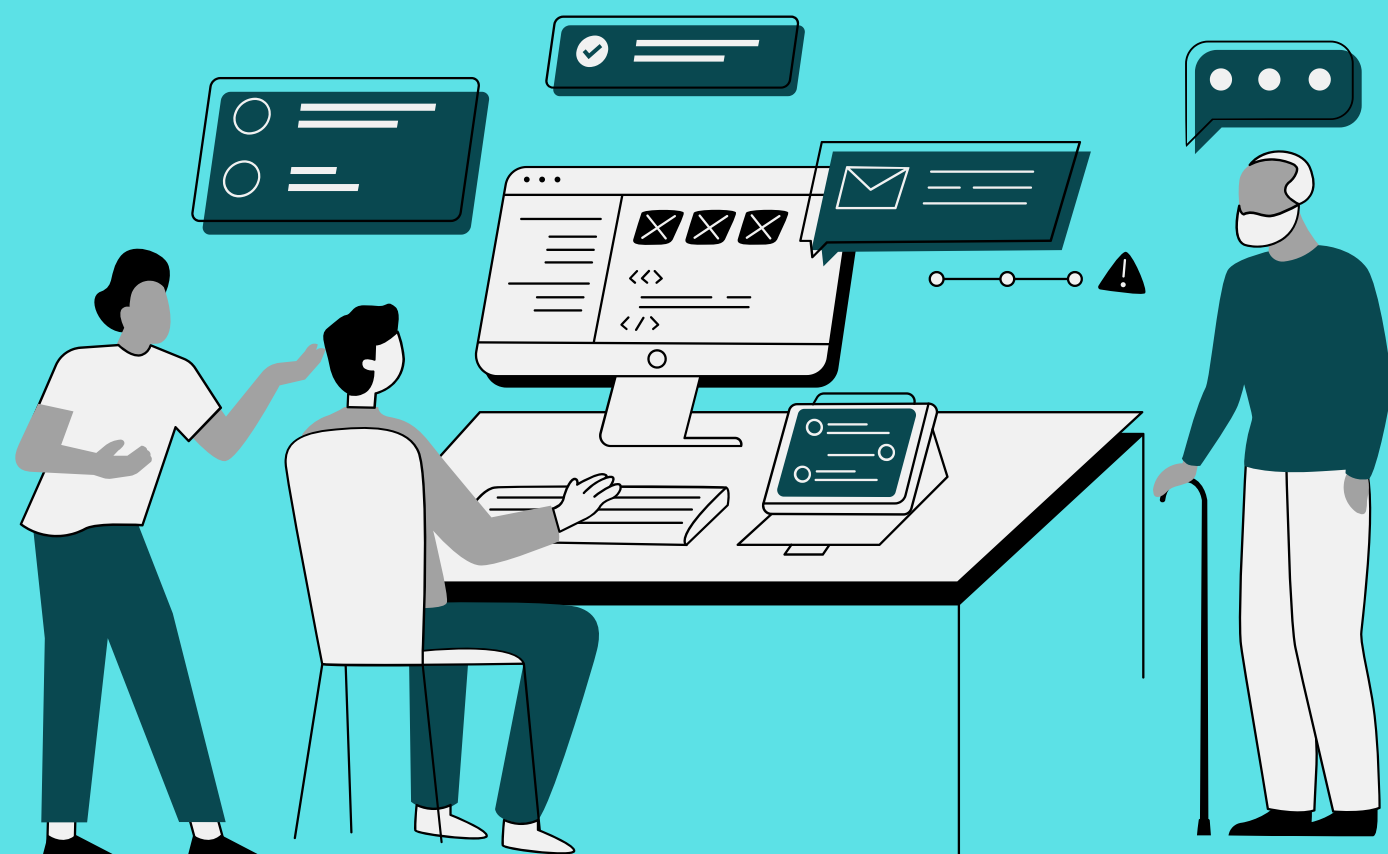
- Необходимо разработать метод на основе нейронных сетей, который смог бы классифицировать эмоции по изображению лица.
- В нашей модели мы занимались классификацией четырех эмоций, а именно : отвращение, радость, страх и равнодушие.

# Введение

Распознавание эмоций – это задача, решение которой может служить как для создания мультимедиа приложений и игр, так и для более сложных и востребованных задач, таких как анализ психического и эмоционального состояния людей, эффективности дискуссий, переговоров, рекламы и т. д.



# Область приложения



Продажи и реклама

Банки

Сфера развлечений

Образование

Работа с персоналом

Автомобили

Медицина

# Основные трудности

Задача распознавания эмоций является трудно формализуемой, и поэтому при ее постановке возникает ряд проблем, главными из которых являются отсутствие стандарта мимических картин и субъективность мнения эксперта, поскольку человек не может в полной мере воспринимать все микродвижения собеседника.

Также проблемой является наличие избыточной информации на изображениях, используемых для обучения сети таких как позиция головы и неравномерное освещение.

Помимо проблем связанных с самой задачей, были проблемы связанные с подходом, а именно:

- Время

- Сложность в подборе параметров и оптимизации гиперпараметров

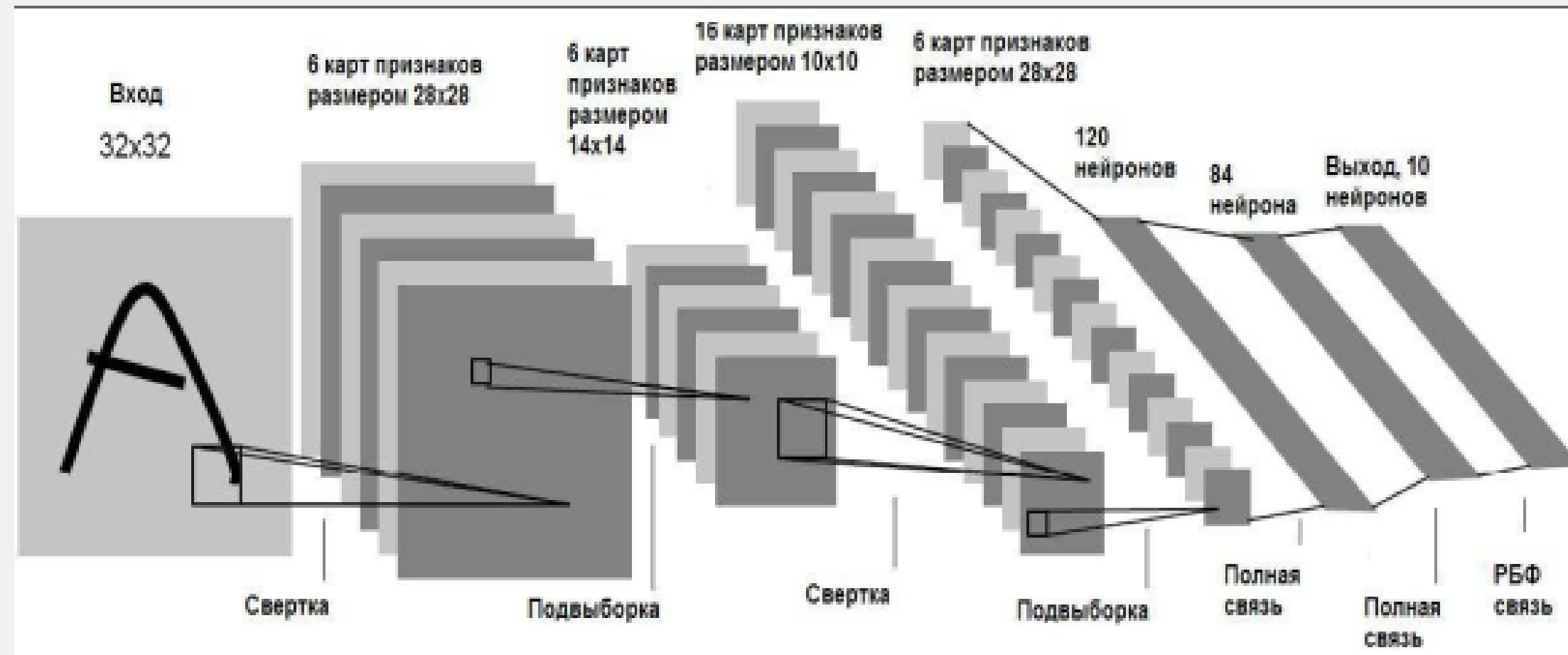
- Вычислительные мощности

# Метод

- Нейронная сеть — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма.
- В задачах где требуется обработать изображения наилучшие результаты показывает свёрточная нейронная сеть.
- Обучение проводилось методом обратного распространения ошибки (обучение с учителем).

# Сверточные нейронные сети

- Один из лучших алгоритмов по распознаванию и классификации изображений.
- По сравнению с полносвязной нейронной сетью (типа перцептрона) — гораздо меньшее количество настраиваемых весов, так как одно ядро весов используется целиком для всего изображения.
- Относительная устойчивость к повороту и сдвигу распознаваемого изображения.
- Обучение при помощи классического метода обратного распространения ошибки.



# Архитектура сверточной нейронной сети

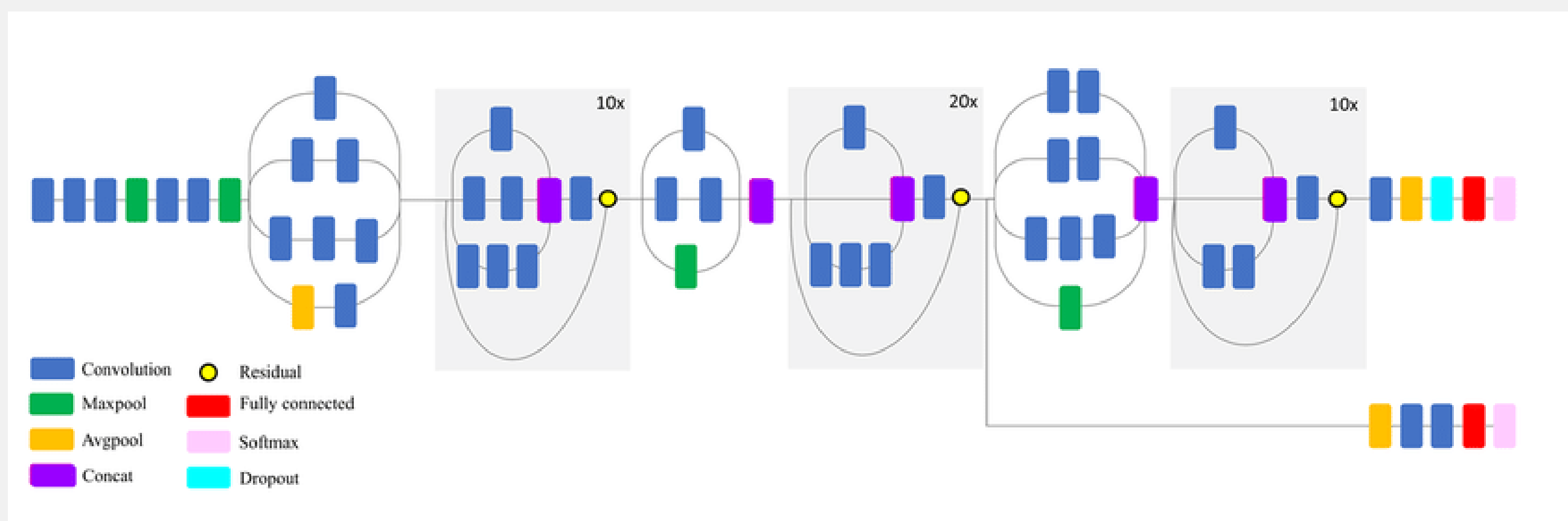
1. Входной слой
2. Сверточный слой
3. Слой подвыборки
4. Полносвязный слой
5. Слой нормализации
6. Исключающий слой



# Архитектура

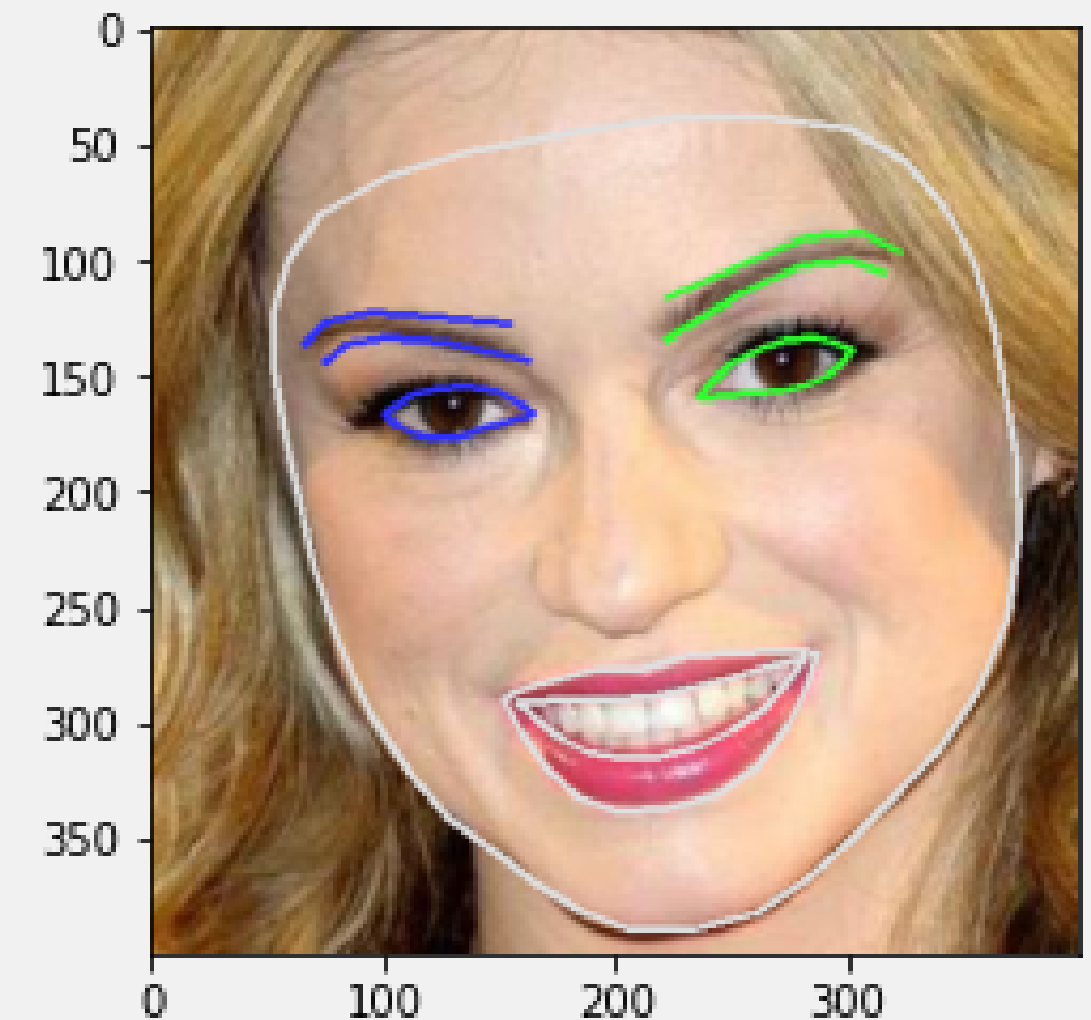
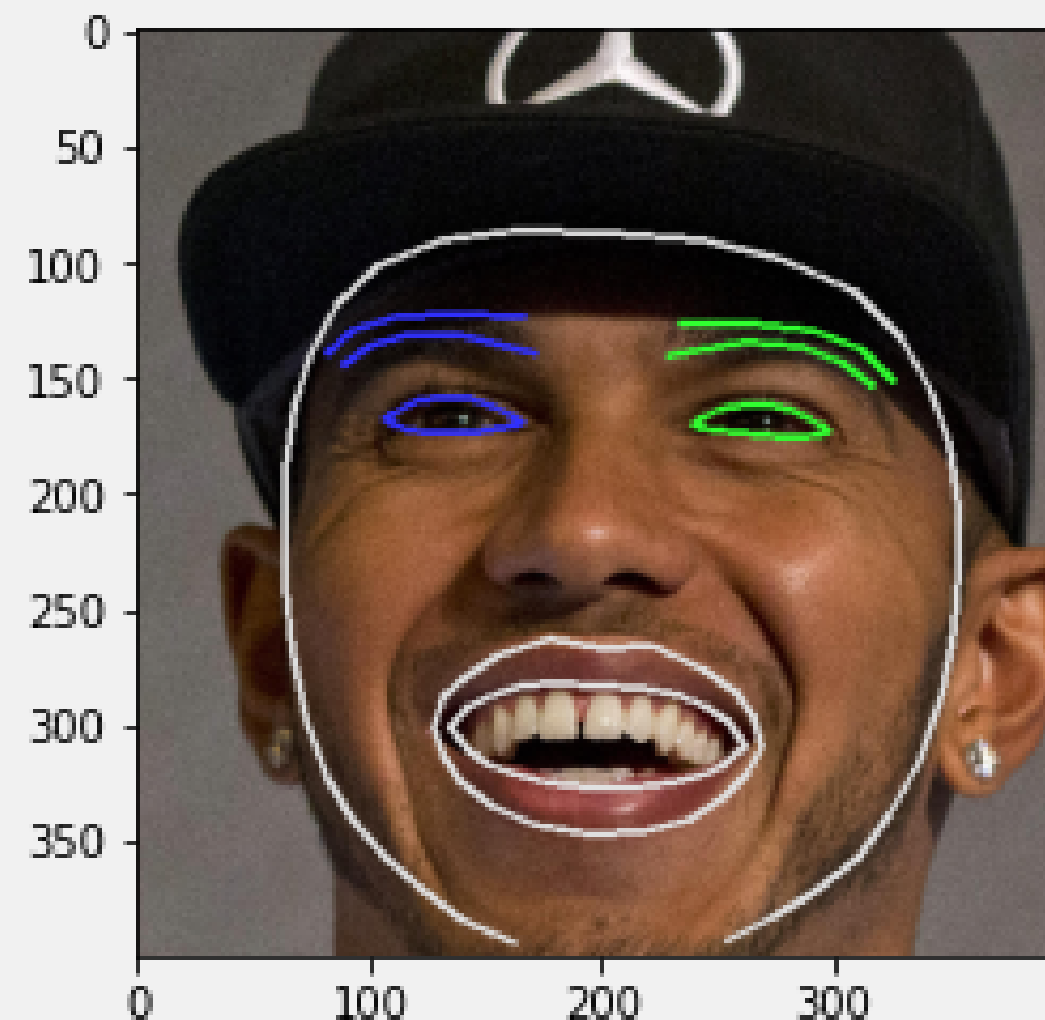
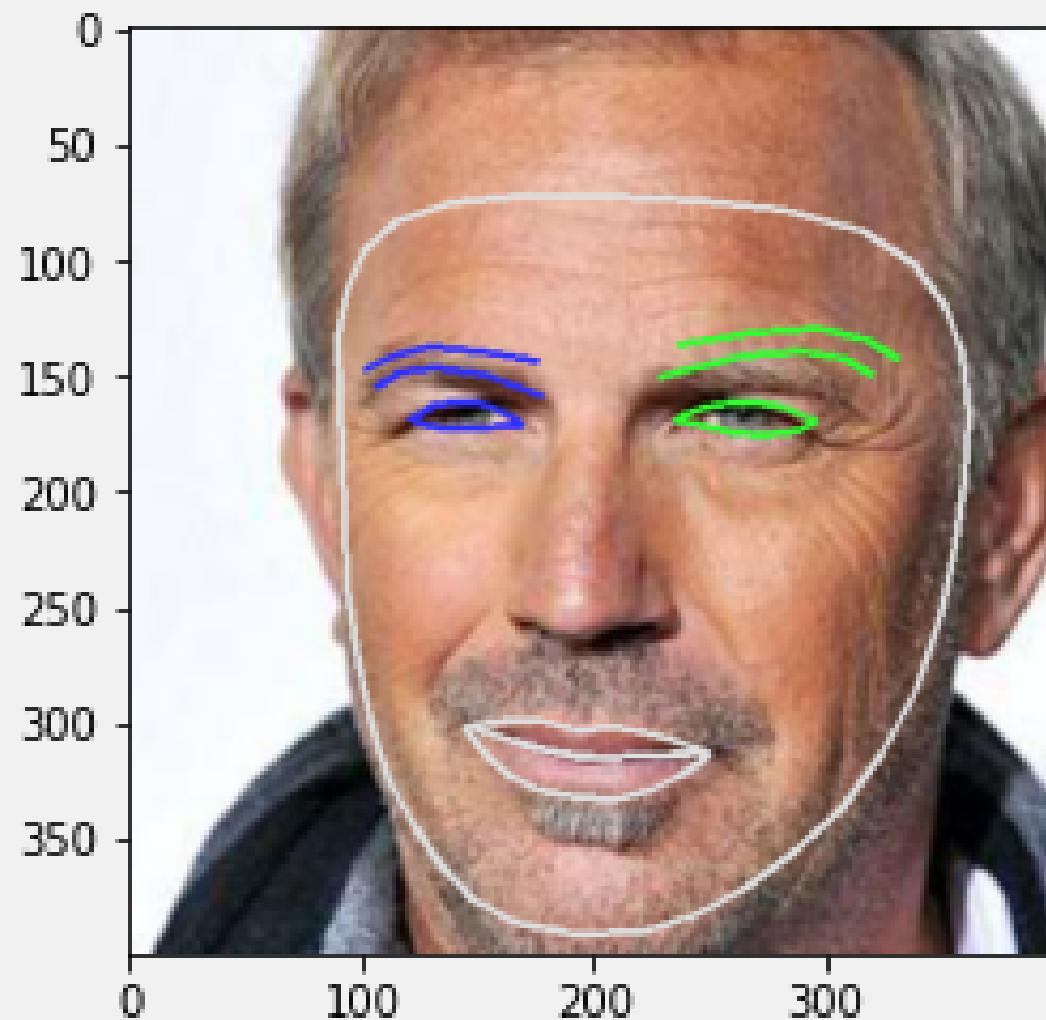
- В попытках улучшить результаты были опробованы некоторые предобученные модели, доступные в Keras, такие, как: VGG19 (которая использовалась в предыдущей задаче распознавания языка жестов), Xception, ResNet50, InceptionV3, DenseNet201. Наилучшие результаты показала InceptionResNetV2.

- InceptionResNetV2 — это сверточная нейронная сеть, обученная на более чем миллионе изображений из базы данных ImageNet.



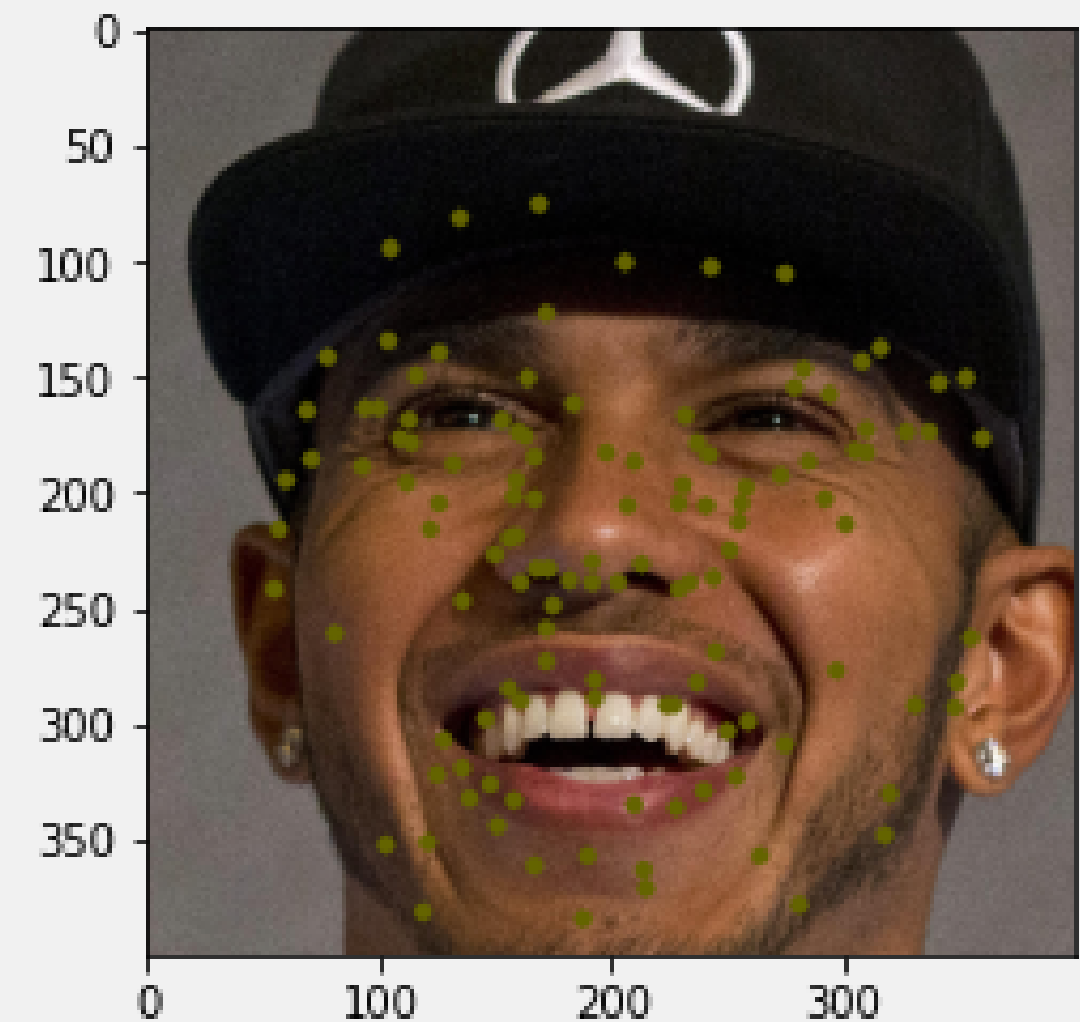
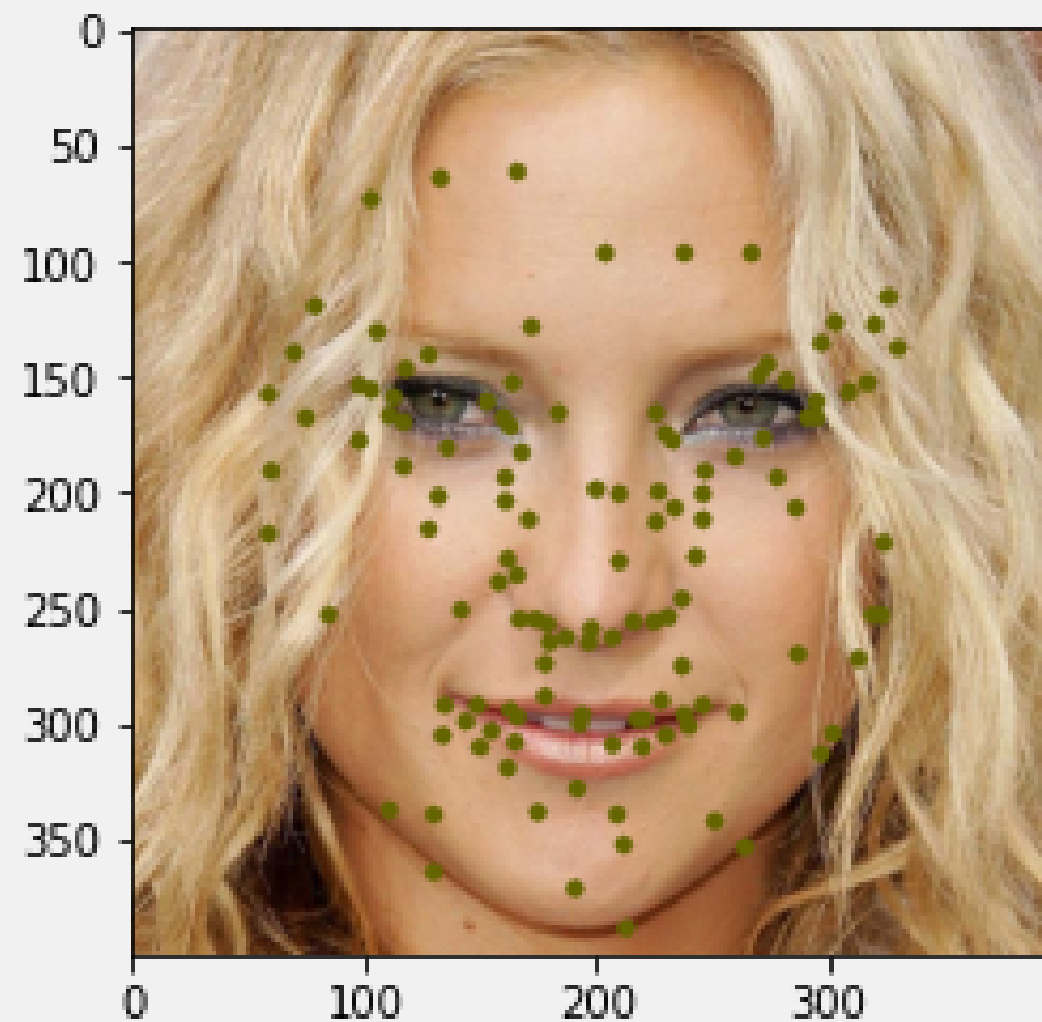
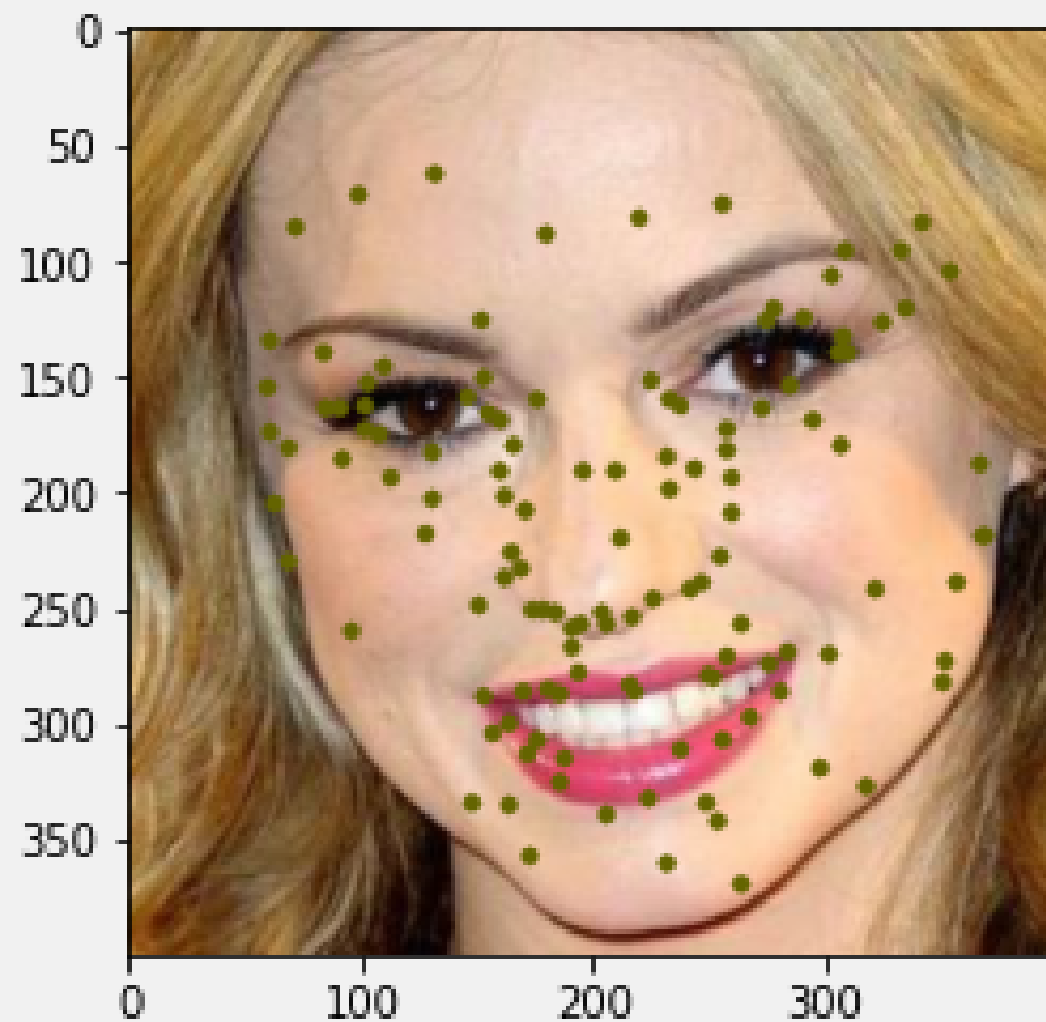
# Инструменты для выделения признаков

◦ В поисках возможности увеличения точности классификации была рассмотрена возможность предварительной обработки изображений с помощью наложений на них "маски", которая должна была выделить общий контур каждого лица и подчеркнуть опознавательные признаки каждой эмоции. В этом на помощь приходит Mediapipe Face Mesh.



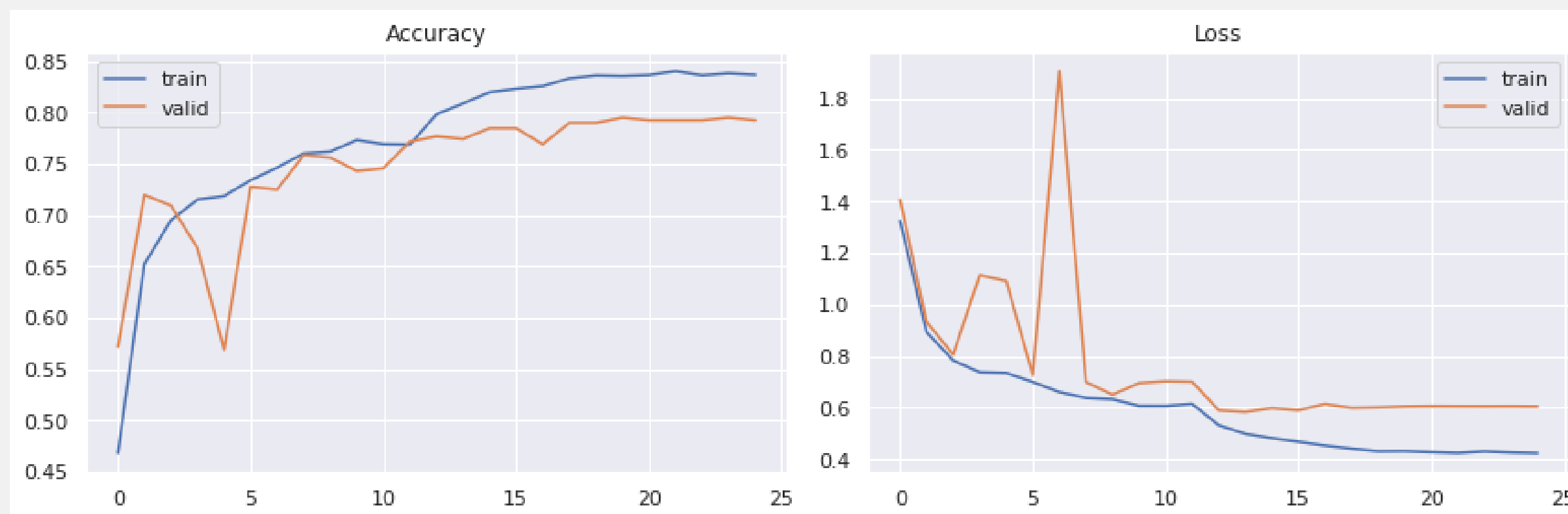
◦MediaPipe Face Mesh — это "потомок" MediaPipe, который используется для геометрии лица, которое оценивает 468 трехмерных ориентиров лица в режиме реального времени.

Данные методы могут быть полезны не только для классификации, но и для обнаружения лица на изображении.



# Иллюстрация процесса обучения

Предварительно обработанные изображения подавались в модель с размером 112 на 112, в трех каналах, пакетами по 64 штуки. Обучение длилось 25 эпох.



# Результаты

Сравнение нашей модели с чужими:

	Точность	Полнота
Модель	78.28%	77.31%
CAER-Net	74.4%	74.5%
EmoPy	68.6%	68.4%
НГТУ	64.6%	64.2%
БГУ	59.8%	59.3%

Матрица ошибок:



# Заключение (выводы)

В ходе исследования была решена задача обнаружения и классификации эмоции по выражению лица.

Точность классификации эмоций составила 78.3%

Было выявлено, что предварительная обработка изображений с помощью Mediapipe помогает получить более высокую точность.

# Список литературы

- Компьютерные технологии и анализ данных (CTDA'2020) : материалы II Междунар. науч.-практ. конф., Минск, 23–24 апр. 2020 г. / Белорус. гос. ун-т ; редкол.: В. В. Скакун (отв. ред.) [и др.]. – Минск : БГУ, 2020. – С. 245-248.
- Тимофеева Ольга Павловна, Неимущев Сергей Андреевич, Неимущева Любовь Ивановна, Тихонов Иван Александрович Распознавание эмоций по изображению лица на основе глубоких нейронных сетей // Труды НГТУ им. Р. Е. Алексеева. 2020. №1 (128).
- Гафаров Ф.М Г12 Искусственные нейронные сети и приложения: учеб. пособие / Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Изд-во Казан. ун-та, 2018. – 121 с.
- Предобученные модели распознавания эмоций EmoRu выложили в открытый доступ ([neurohive.io](https://neurohive.io))
- Нейросеть научили распознавать эмоции по окружающей ситуации ([nplus1.ru](https://nplus1.ru))

**Благодарю за внимание!**