



САМАРСКИЙ УНИВЕРСИТЕТ
SAMARA UNIVERSITY

Применение свёрточных нейронных сетей в задачах по определению типов лиц

Выполнила Пчелкина Ю.Ж.
Научный консультант Литвинов В.Г.

Самара 2021



Цель работы: автоматизировать процесс распознавания типов лиц по фотоснимкам при помощи интеллектуального анализа изображений с применением свёрточных нейронных сетей.

Задачи:

- сбор и предварительная обработка данных;
- разработка программного модуля для анализа и классификации имеющихся данных;
- подготовка данных для обучения;
- применения свёрточных нейронных сетей для определения типа лица на изображении.

Исходные данные: Изображения с различными характеристиками (размер, разрешение, качество, цветовая модель).

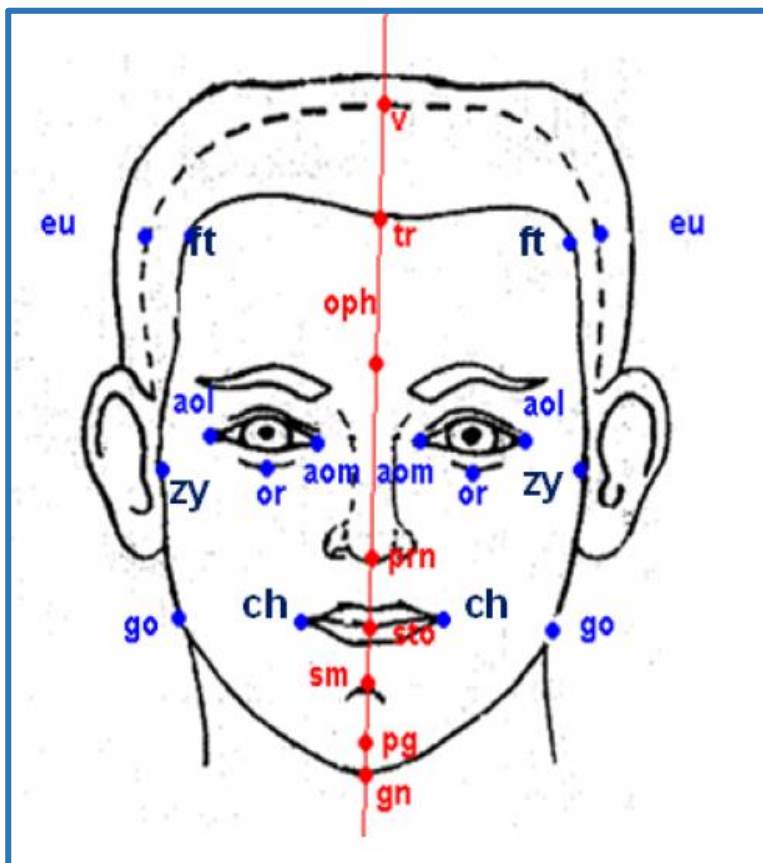


Антропометрический метод (anthropos – человек; metron – мера) – метод исследования, основанный на измерении морфологических и функциональных признаков человека.

В ортодонтии антропометрический метод включает в себя исследование эстетики лица (формы, размеров, угловых и линейных параметров лица, их соотношение)

Условные обозначения	Меднальные точки	Условные обозначения	Латеральные точки	Условные обозначения	Меднальные точки	Условные обозначения	Латеральные точки
op	Опистокранион – наиболее выступающая кзади точка затылочной кости.	eu	Эурион – латерально выступающая точка на голове	n	Назион – наиболее выраженное углубление между лбом и носом	t	Трагус – точка на верхнем крае наружного слухового прохода
tr	Трихион – точка передней границы волосистой части лба	ft	Фронтотемпорале – наиболее выступающая точка лобно-височного соединения	prn	Проназале – наиболее выступающая точка кончика носа при ориентировании головы во франкфуртской	aom	Ангулюс окули медиалис – точка внутреннего угла галаза
gl	Глабелла – наиболее впереди расположенная точка лобной кости.	or	Орбитале – точка нижнего края глазницы (находится на пересечении перпендикуляра, опущенного на касательную к нижнему краю глазницы из зрачка при смотрящих вперед глазах)	sn	Субназале – точка соединения кожной перегородки носа с верхней губой	aol	Ангулюс окули латералис – точка наружного угла глаза
oph	Офрион – точка пересечения срединно-сагиттальной плоскости с линией, соединяющей надбровные дуги	zy	Зигион – наиболее выступающая снаружи точка скуловой дуги	sto	Стомион – точка пересечения линии смыкания губ с спелинно-сагиттальной	ch	Хейлион – точка угла рта
				pg	Погонион – наиболее выступающая точка подбородка	go	Гонион – наиболее выступающая снаружи и кзади точка угла нижней челюсти
				gn	Гнатион – наиболее нижняя точка подбородка		

Цефалометрические измерительные точки головы и лица. Морфологический фациальный индекс Izard



1. **Прямоугольная форма** – когда морфологическая высота лица превосходит его скуловую ширину, а касательные к боковому контуру головы (eu-go) параллельны между собой или конвергируют незначительно до 30 %.

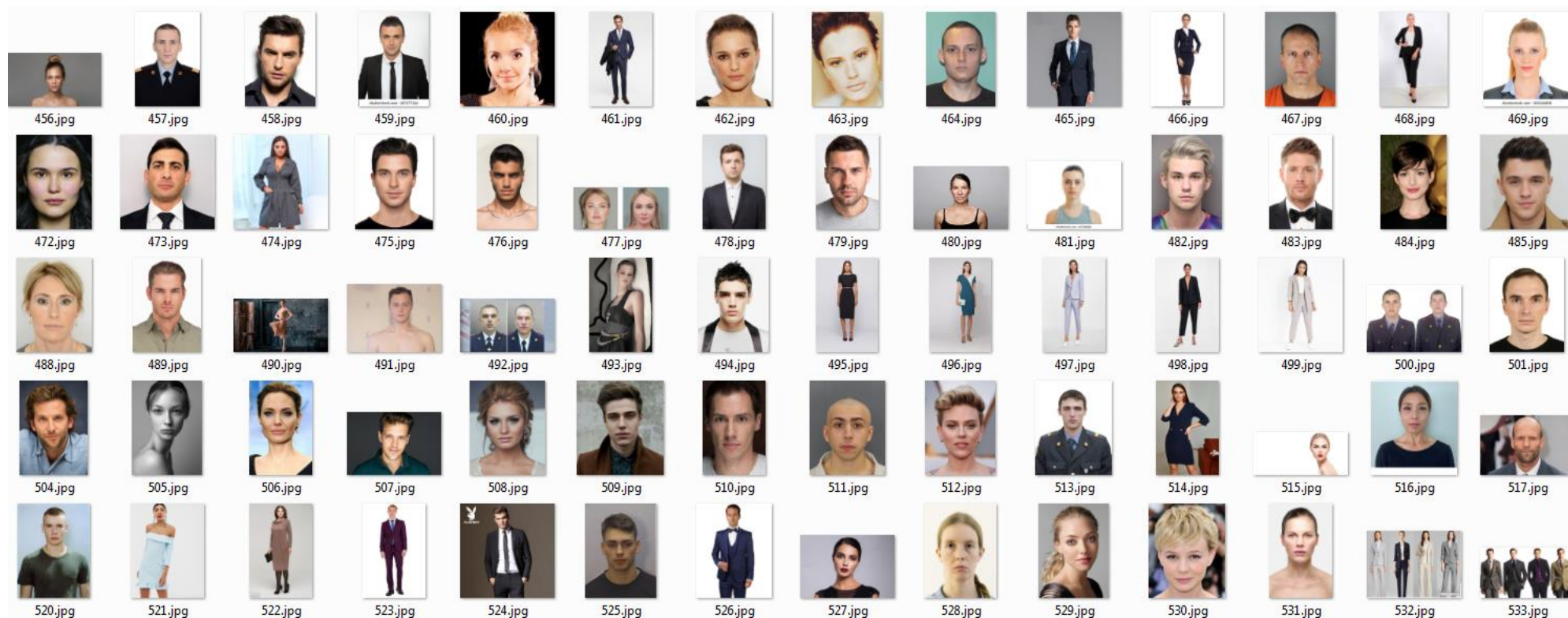
2. **Квадратная форма** – при равном или меньшем значении морфологической высоты по отношению к скуловой ширине, при параллельных касательных к боковому контуру головы.

3. **Треугольная форма** – при которой касательные к боковым контурам головы резко конвергируют книзу (угол наклона больше 30%).



Исходные данные

В работе использовались фотоснимки с различными характеристиками (размер, разрешение, качество, цветовой режим).





исходное
изображение
Размер 800x800



выделение контура
лица каскадным
классификатором Хаара



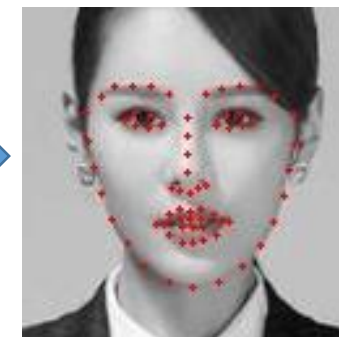
Преобразование
в оттенки серого,
обрезка по контуру
Размер 145x145



масштабирование
изображения
Размер 128x128



нахождение
ключевых
точек лица





Определяем ключевые точки:

$$N(x_n; y_n), G_n(x_{gn}; y_{gn}), Zy_1(x_{z1}; y_{z1}), Zy_2(x_{z2}; y_{z2}), Go_1(x_{go1}; y_{go1}), Go_2(x_{go2}; y_{go2})$$

$$\text{Расстояния: } NG = \sqrt{(x_n - x_{gn})^2 + (y_n - y_{gn})^2} \quad ZZ = \sqrt{(x_{z1} - x_{z2})^2 + (y_{z1} - y_{z2})^2}$$

Прямые задаются каноническими уравнениями:

$$Zy_2Go_2: \frac{x - x_{z2}}{x_{go2} - x_{z2}} = \frac{y - y_{z2}}{y_{go2} - y_{z2}} \quad \text{или} \quad \frac{x - x_{z2}}{m_1} = \frac{y - y_{z2}}{n_1}$$

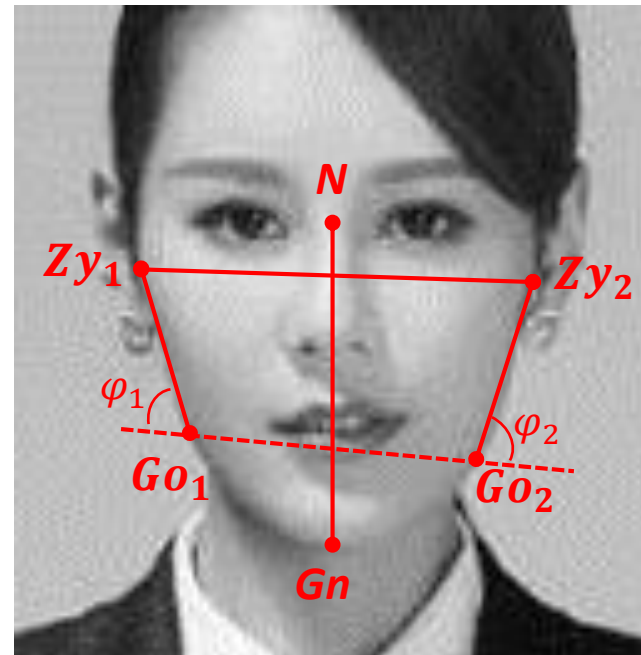
$$Go_1Go_2: \frac{x - x_{go1}}{x_{go2} - x_{go1}} = \frac{y - y_{go1}}{y_{go2} - y_{go1}} \quad \text{или} \quad \frac{x - x_{go1}}{m_2} = \frac{y - y_{go1}}{n_2}$$

Тогда угол наклона прямой Zy_2Go_2 :

$$\varphi_2 = \arccos \left(\frac{m_1 m_2 + n_1 n_2}{\sqrt{m_1^2 + n_1^2} \cdot \sqrt{m_2^2 + n_2^2}} \right)$$

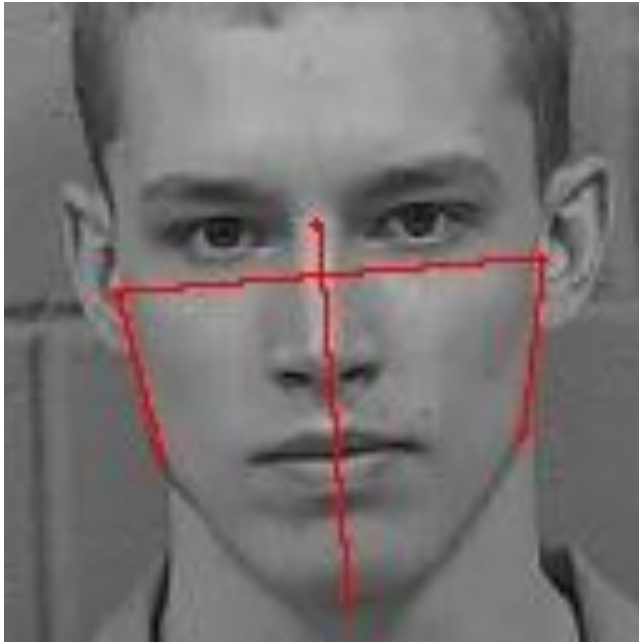
Определяем тип лица:

1. **Прямоугольная форма:** $NG > ZZ$ и прямые Zy_1Go_1 и Zy_2Go_2 параллельны или угол наклона меньше 30°
2. **Квадратная форма:** $NG \leq ZZ$ и прямые Zy_1Go_1 и Zy_2Go_2 параллельны или угол наклона меньше 30°
3. **Треугольная форма:** угол наклона касательных Zy_1Go_1 и Zy_2Go_2 больше 30°





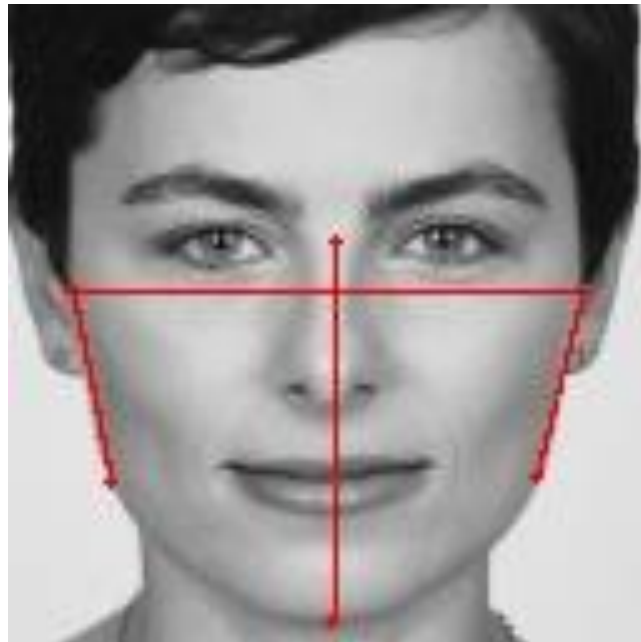
1 тип.
Прямоугольная форма



угол между наклона боковых линий $< 30^\circ$

длина срединного вертикального
отрезка больше горизонтального

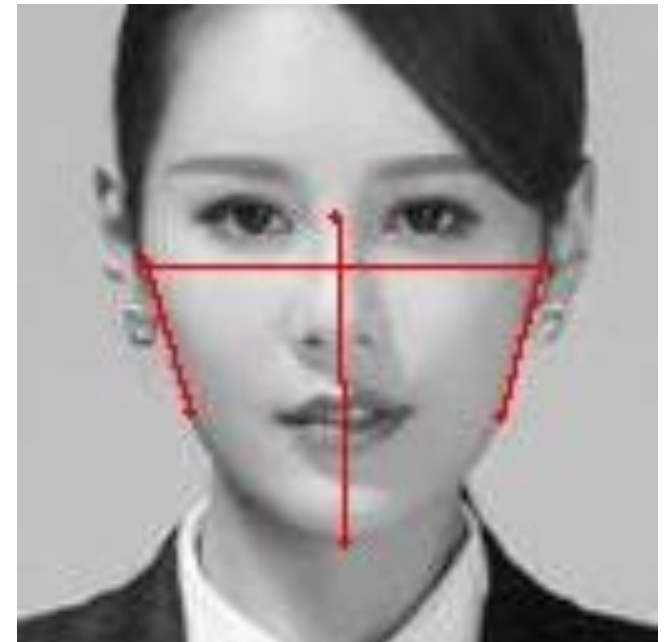
2 тип.
Квадратная форма



угол между наклона боковых линий $< 30^\circ$

длина срединного вертикального
отрезка меньше горизонтального

3 тип.
Треугольная форма



угол между наклона боковых линий $> 30^\circ$



Предобработка данных. Создание DataFrame.

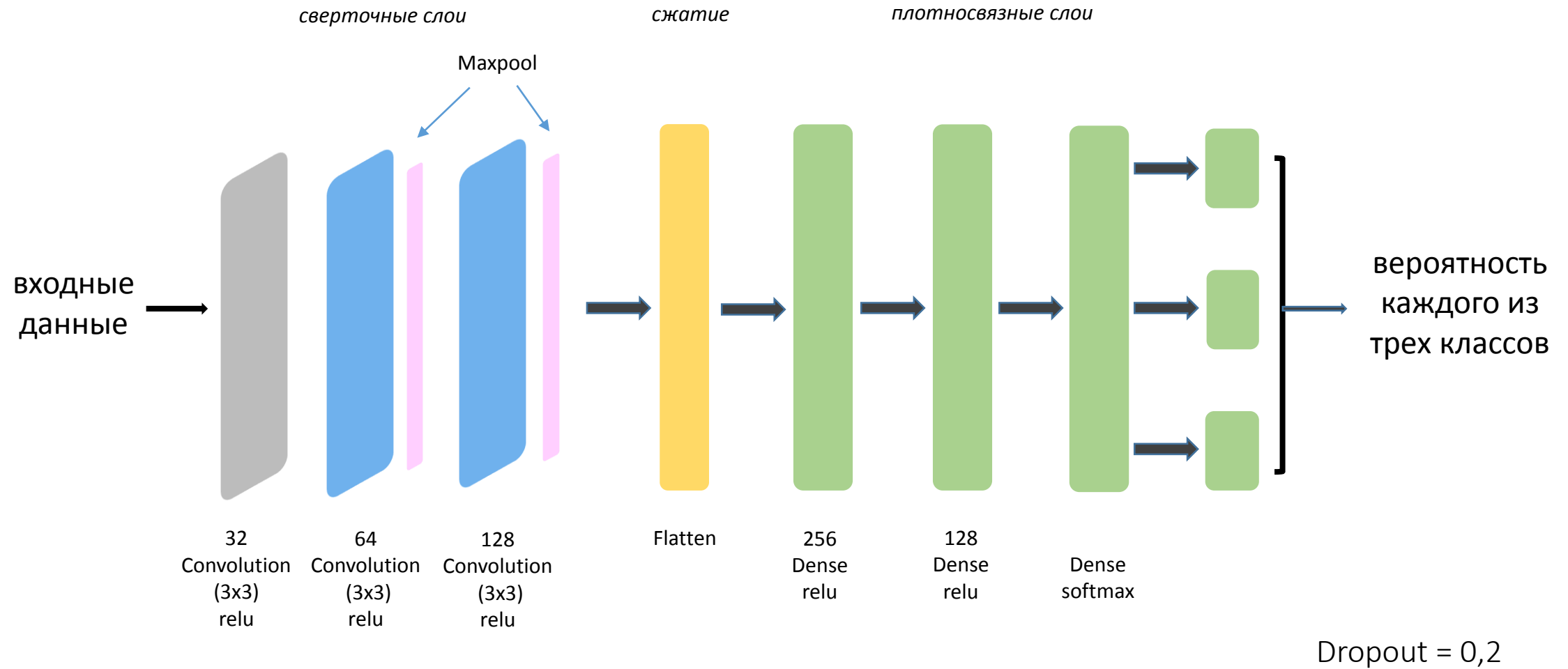


foto	kind
r1.jpg	3
r10.jpg	2
r100.jpg	2
r1000.jpg	2
r1004.jpg	2
r1006.jpg	2
r1008.jpg	2
r1009.jpg	2
r1011.jpg	3
r1013.jpg	3
r1015.jpg	1
r1017.jpg	2

размер 1703 × 2
статистика:
3 тип – 421
2 тип – 668
1 тип – 614



Свёрточная нейронная сеть. Архитектура.





- Размер обучающей выборки 75 %
- Функция потерь – перекрестная энтропия
- Оптимизатор – алгоритм Adam
- Количество партий (bath) = 32

Для оценки точности обучения нейронной сети использовалась метрика:

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

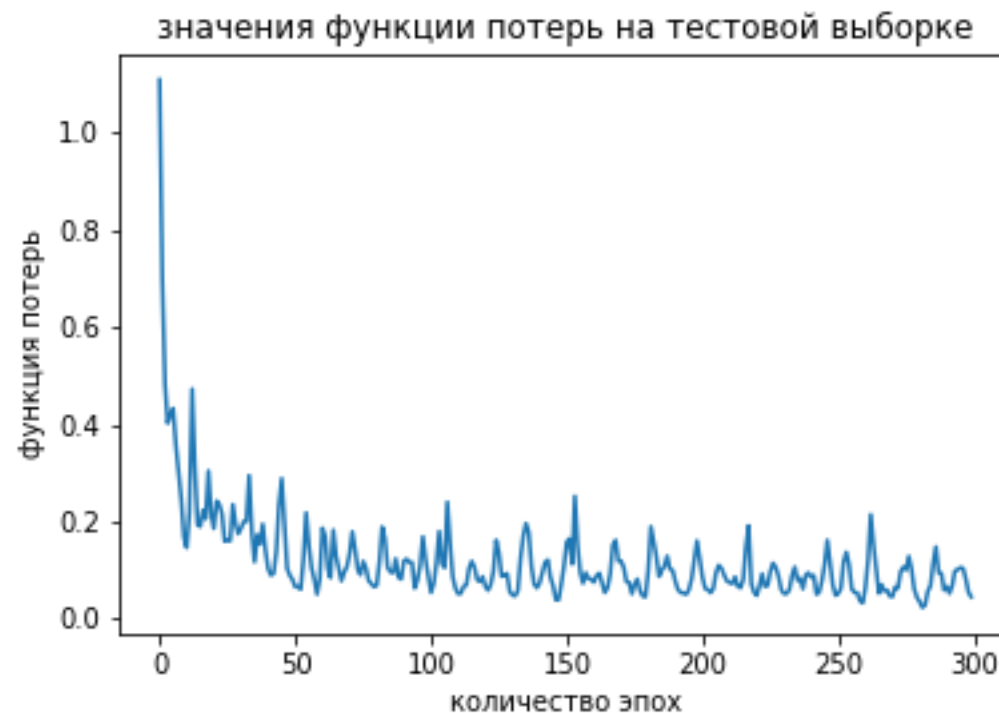
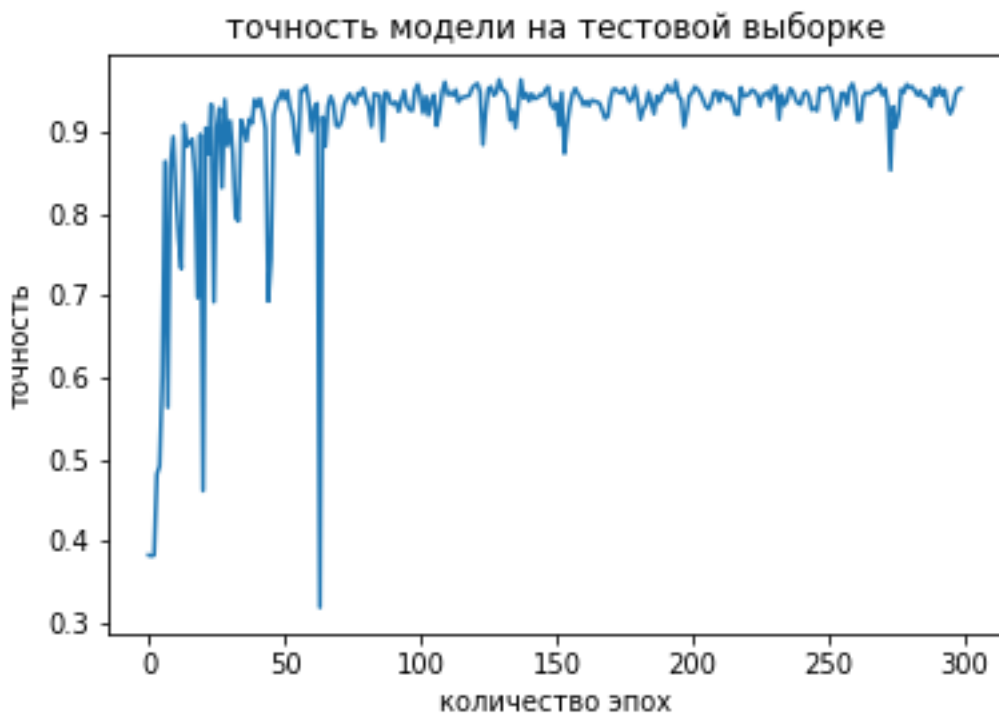
- tp (true positive, истинно положительные) – объект верно отнесён к данному классу;
- tn (true negative, истинно отрицательные) – верно определено, что объект не принадлежит к данному классу;
- fp (false positive, ложно положительные) – объект неверно отнесён к данному классу;
- fn (false negative, ложно отрицательные) – неверно определено, что объект не принадлежит к данному классу.



Объем выборки	Разделение по классам	Количество эпох обучения	Точность обучения
1 703	1 тип – 614 2 тип – 668 3 тип – 421	250	64,87 %
3 318	1 тип – 1584 2 тип – 1278 3 тип – 456	250	76,03%
4 110	1 тип – 1584 2 тип – 1278 3 тип – 1248	250	94,55%



количество эпох обучения	3	10	50	250	300
точность	11 %	65,6 %	87,3 %	92,71 %	95,43 %





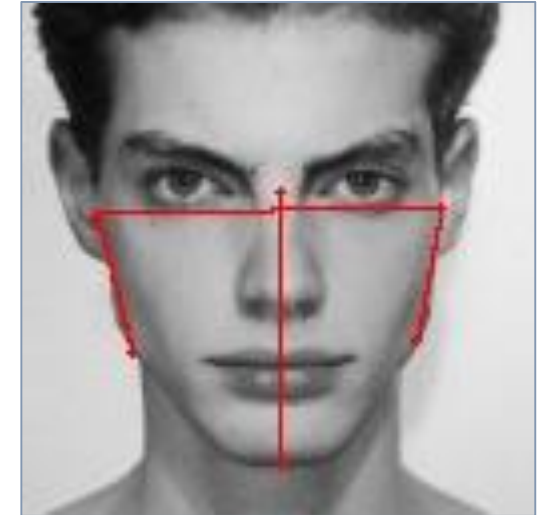
Тип лица «3» - треугольный
Опознан как тип «2» - квадратный



Тип лица «1» - прямоугольный
Опознан как тип «2» - квадратный



Тип лица «2» - квадратный
Опознан как тип «3» - треугольный



Тип лица «2» - квадратный
Опознан как тип «3» - треугольный



В ходе выполнения работы были решены следующие задачи:

- Описан метода определения типов и признаков лица в ортодонтии
- Проведена предобработка данных, создан набора данных
- Разработана программная реализация метода определения типов лиц
- Реализована модель нейронной сети
- Проведены обучение модели по исходным данным, оценка точности модели:
 - на 300 эпохах обучения точность выше 95%.
 - На величину ошибки обучения влияют:
 - ✓ объем исходных данных;
 - ✓ сбалансированность классов;
 - ✓ качество исходных данных;
 - ✓ количество эпох обучения.

Полученные алгоритмы применимы при автоматизации процесса заполнения части обязательной медицинской документации при первичном осмотре пациента врачом ортодонтом.



САМАРСКИЙ УНИВЕРСИТЕТ
SAMARA UNIVERSITY

БЛАГОДАРЮ
ЗА ВНИМАНИЕ!