НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Методы идентификации лиц с автоматическим определением редких данных во входном видеопотоке

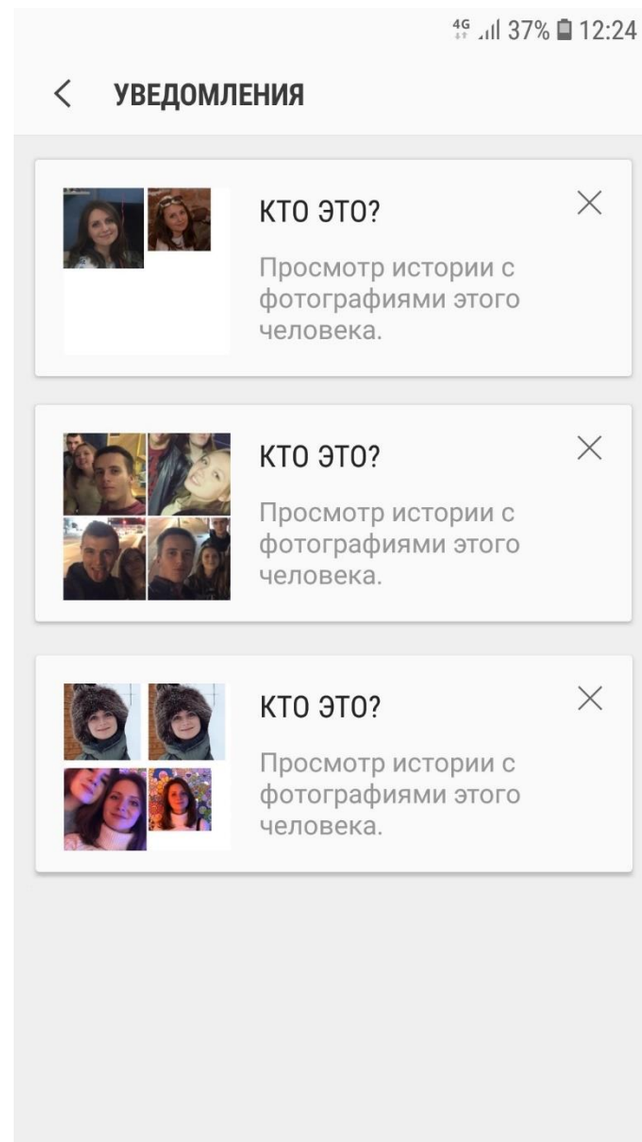
Студент группы 18 МАГ ИАД:
Соколова Анастасия Дмитриевна

Научный руководитель:
д.т.н., профессор
Савченко Андрей Владимирович

Нижний Новгород, 2020

АКТУАЛЬНОСТЬ РАБОТЫ

- Системы навигации по галереи изображений
- Группировка видеоданных для ведения статистики
- Распознавание объектов в сфере общественной безопасности
- и т.д.



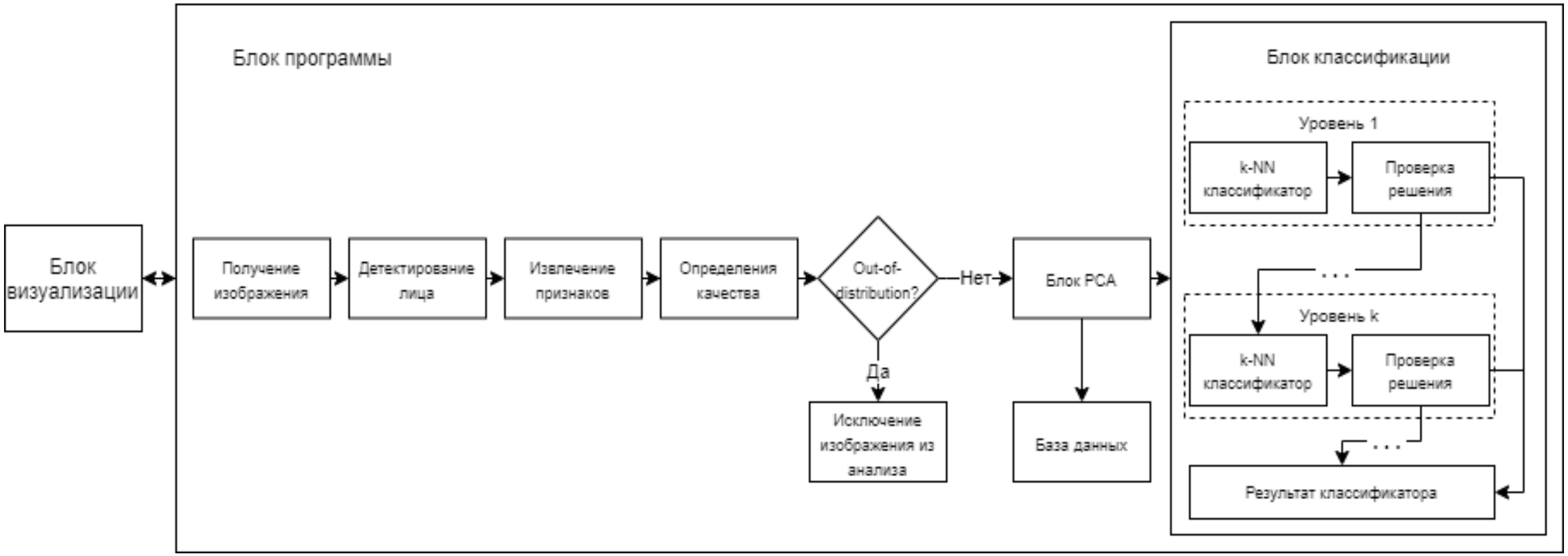
ЦЕЛЬ И ЗАДАЧИ РАБОТЫ

Цель: создание метода идентификации лиц по изображениям на основе определения редких данных во входном потоке: дети, пожилые люди, макияж, плохое качество

Задачи:

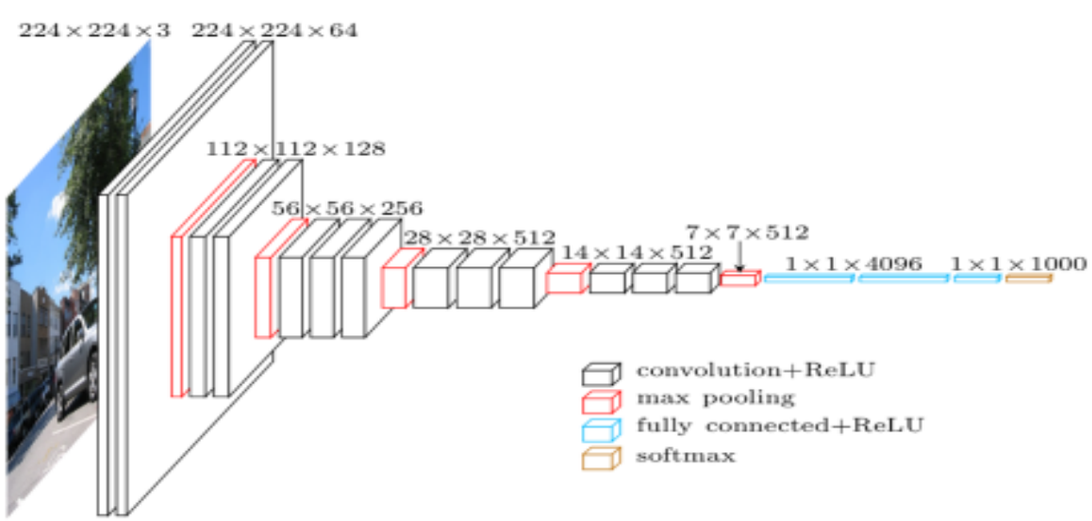
- Провести аналитический обзор литературы по алгоритмам, обрабатывающим большие данные
- Собрать набор данных, содержащий аномалии
- Описать схему разработанного алгоритма
- Разработать прототип системы
- Провести экспериментальное исследование разработанного алгоритма, выявить его достоинства и недостатки

СХЕМА СИСТЕМЫ



ИЗВЛЕЧЕНИЕ ПРИЗНАКОВ ИЗОБРАЖЕНИЯ

- VggFace2 (Resnet50)
  1x2048 вектор
- MobileNet2
  1x1024 вектор
- InsightFace (LResNet100E-IR)
  1x512 вектор
- FaceNet
  1x512 вектор



НАБОРЫ ДАННЫХ

- VGGFace2
  - 9 000 людей
  - > 3.3 млн изображений лиц
  - ~362 изображений на человека
- MS-Celeb-1M (Microsoft Celeb)
  - 100 000 людей
  - 10 млн изображений лиц
  - ~100 изображений на человека
- LWF (Labeled Faces in the Wild)
  - 1680 людей
  - 13 000 изображений лиц
- All ages faces dataset
  - 13 322 людей
  - 13 322 изображений лиц
  - Возраст от 2 до 80
- Large age gap face verification
  - 1 010 людей
  - 3 828 изображений лиц
  - Для каждой идентичности присутствуют как минимум одно детское/молодое изображение и одно взрослое/старое изображение

http://www.robots.ox.ac.uk/~vgg/data/vgg_face2/
https://megapixels.cc/msceleb/
http://vis-www.cs.umass.edu/lfw/
https://github.com/JingchunCheng/All-Age-Faces-Dataset
http://www.ivl.disco.unimib.it/activities/large-age-gap-face-verification/

СОЗДАНИЕ НАБОРА ДАННЫХ (С «OUT-OF-DISTRIBUTION» ИЗОБРАЖЕНИЯМИ)

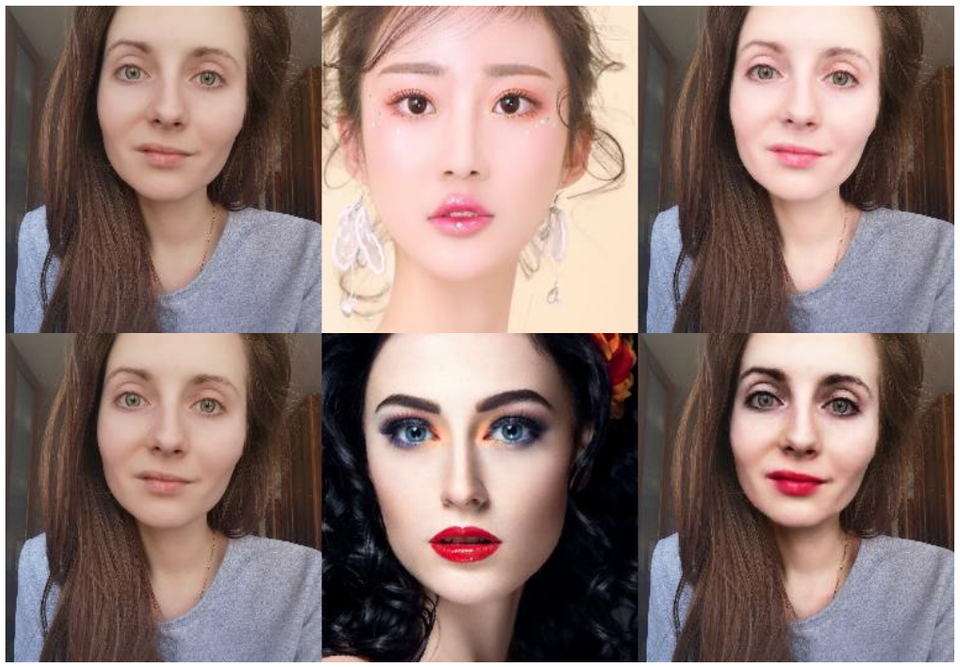
- Для VGGFace2 и MS1M:

Кластеризация изображений одного человека → Выбор одного изображения из каждого кластера → Классификация полученных изображений → Где классификатор ошибается всех чаще - «out-of-distribution» изображения; где классификатор не ошибается – хорошее качество

- Для All ages faces dataset и Large age gap face verification: детские и пожилые изображения были помечены как «out-of-distribution»
- Добавление трансформаций

ДОБАВЛЕНИЕ ТРАНСФОРМАЦИИ

ДОБАВЛЕНИЕ ТРАНСФОРМАЦИИ

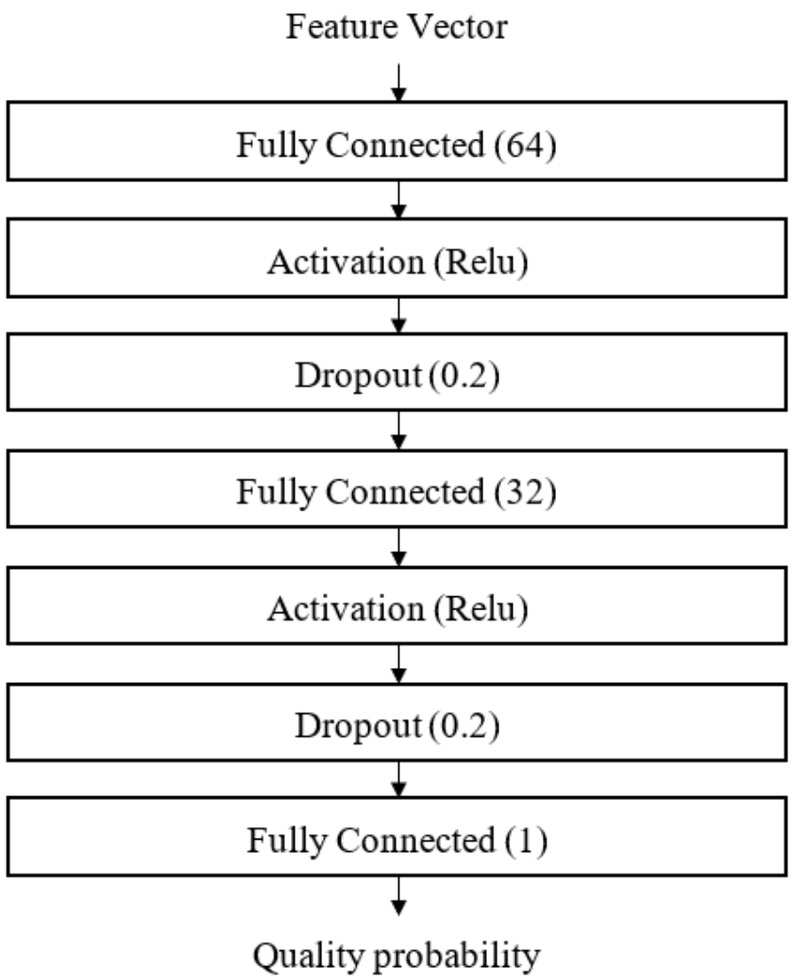


Zhang H. et al. Disentangled Makeup Transfer with Generative Adversarial Network //arXiv preprint arXiv:1907.01144. – 2019

Chen Y. C. et al. Facelet-bank for fast portrait manipulation //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2018. – C. 3541-3549

АЛГОРИТМ КЛАССИФИКАЦИИ

- Определение «out-of-distribution» изображений



АЛГОРИТМ КЛАССИФИКАЦИИ

- Расстояние между признаками с помощью уровней иерархии признаков

$$\rho\left(\tilde{x}^{(l)}, \tilde{x}^{(l)}\right) = \rho\left(\tilde{x}^{(l-1)}, \tilde{x}^{(l-1)}\right) + \sum_{d=d_{(l-1)}+1}^{d_l} \rho(\tilde{x}_d, \tilde{x}_{r;d})$$

- Определение принадлежности к классу

$$\mathrm{c}_l^* = \underset{c \in C_l}{\operatorname{argmin}}\, \rho_{\mathrm{c}}(\tilde{x}^{(l)})$$

- Множество кандидатов

$$C_{l+1} = \left\{ c \in C_l \left| \frac{\rho_{\mathrm{c}}(\tilde{x}^{(l)})}{\rho_{\mathrm{c}_l^*}(\tilde{x}^{(l)})} \leq \delta \right. \right\}$$

ОПРЕДЕЛЕНИЕ «OUT-OF-DISTRIBUTION» ИЗОБРАЖЕНИЙ

- Сверточные нейронные сети использовались с предобученными на ImageNet весами

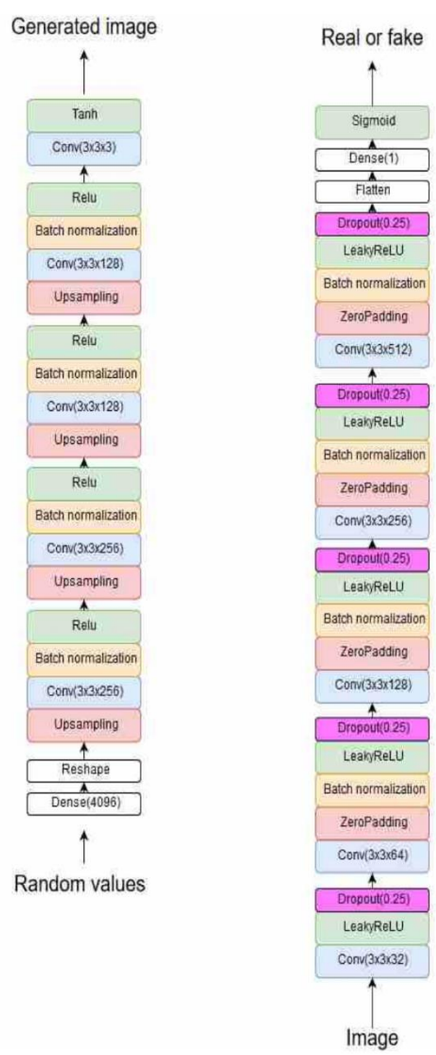
Сеть	Точность	
	Тренировочный набор	Тестовый набор
EfficientNetB3	93.2%	**58.9%**
MobileNetV2	92.2%	56.3%
ResNet50	91.4%	58.8%
Xception	92.4%	49.7%
InceptionV3	90.1%	55.4%

Tan M., Le Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks // arXiv preprint arXiv:1905.11946. – 2019
Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L. C. Mobilenetv2: Inverted residuals and linear bottlenecks // In Proceedings of the IEEE conference on computer vision and pattern recognition. – 2018. – P. 4510-4520
He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition // In Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 770-778
Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computer vision // In Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 2818-2826
Chollet F. Xception: Deep learning with depthwise separable convolutions // In Proceedings of the IEEE conference on computer vision and pattern recognition. – 2017. – P. 1251-1258

ОПРЕДЕЛЕНИЕ OUT-OF-DISTRIBUTION ИЗОБРАЖЕНИЙ

- Генеративно-состязательная сеть

Архитектура сети



Результат генерации изображений лиц

Точность выявления «out-of-distribution» изображения, используя дескриптор

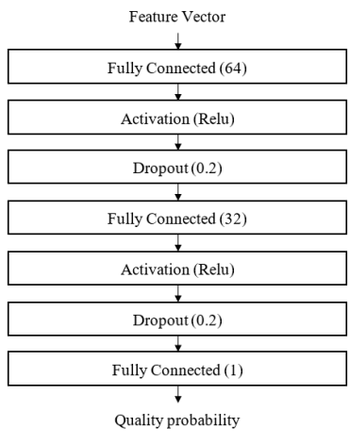
69.4%

ОПРЕДЕЛЕНИЕ «OUT-OF-DISTRIBUTION» ИЗОБРАЖЕНИЙ

- Классификаторы (на вход принимают вектора признаков)

Алгоритм	VGGFace2	MobileNet2	InsightFace	FaceNet
k-NN	89.7%	93.9%	89.6%	90.9%
SVM	69.1%	79.4%	63.0%	64.3%
RandomForest	85.0%	86.5%	80.6%	84.2%
LDA	88.1%	89.0%	83.7%	87.4%
Our NN	92.3%	**94.8%**	89.7%	90.2%

Our NN:



РЕЗУЛЬТАТ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ С ОПРЕДЕЛЕНИЕМ РЕДКИХ ДАННЫХ

- Показатели результатов классификации набора данных, основанного на VGGFace2

VGGFace2

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	88.7	41.63
k-NN + исключение	92.4	20.32
k-NN + преобразование	89.3	43.15
RandomForest + все	36.5	3.89
RandomForest + исключение	82.7	3.54
RandomForest + преобразование	37.3	4.92
LDA + все	88.6	1.89
LDA + исключение	**93.3**	**1.35**
LDA + преобразование	88.8	1.92

MobileNet2

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	94.9	23.39
k-NN + исключение	**97.9**	9.59
k-NN + преобразование	94.9	24.36
RandomForest + все	21.2	3.80
RandomForest + исключение	91.6	2.73
RandomForest + преобразование	21.8	4.02
LDA + все	92.6	1.27
LDA + исключение	97.1	**1.14**
LDA + преобразование	93.0	1.29

InsightFace

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	83.7	13.79
k-NN + исключение	**83.9**	5.94
k-NN + преобразование	83.8	14.49
RandomForest + все	19.9	3.03
RandomForest + исключение	58.5	2.59
RandomForest + преобразование	19.97	3.66
LDA + все	75.7	**0.23**
LDA + исключение	82.0	0.87
LDA + преобразование	75.9	0.89

FaceNet

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	94.1	11.61
k-NN + исключение	**97.7**	5.66
k-NN + преобразование	94.4	12.28
RandomForest + все	42.9	3.02
RandomForest + исключение	93.5	2.54
RandomForest + преобразование	46.1	3.66
LDA + все	92.5	**0.23**
LDA + исключение	96.8	0.87
LDA + преобразование	92.4	0.98

РЕЗУЛЬТАТ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ С ОПРЕДЕЛЕНИЕМ РЕДКИХ ДАННЫХ

- Показатели результатов классификации набора данных, основанного на MS1M

VGGFace2

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	87.6	42.24
k-NN + исключение	**91.9**	21.02
k-NN + преобразование	87.7	43.28
RandomForest + все	50.1	6.34
RandomForest + исключение	79.8	5.21
RandomForest + преобразование	50.1	6.35
LDA + все	88.7	1.94
LDA + исключение	91.5	**1.35**
LDA + преобразование	88.6	1.96

MobileNet2

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	86.3	23.44
k-NN + исключение	90.5	10.13
k-NN + преобразование	86.3	24.29
RandomForest + все	37.4	4.11
RandomForest + исключение	85.7	2.82
RandomForest + преобразование	37.2	4.13
LDA + все	86.4	1.32
LDA + исключение	**91.0**	**1.12**
LDA + преобразование	86.6	1.46

InsightFace

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	83.8	11.51
k-NN + исключение	**84.1**	6.17
k-NN + преобразование	84.0	12.08
RandomForest + все	43.2	4.23
RandomForest + исключение	66.5	2.77
RandomForest + преобразование	41.7	4.28
LDA + все	67.7	0.84
LDA + исключение	67.9	**0.81**
LDA + преобразование	67.5	0.86

FaceNet

Алгоритм	Точность (%)	Время классификации одного изображения (мс)
k-NN + все	84.2	11.59
k-NN + исключение	**89.3**	5.68
k-NN + преобразование	84.4	12.04
RandomForest + все	32.3	4.37
RandomForest + исключение	79.9	2.12
RandomForest + преобразование	32.3	4.49
LDA + все	79.8	0.93
LDA + исключение	82.3	**0.84**
LDA + преобразование	79.7	0.91

РЕЗУЛЬТАТ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ С ОПРЕДЕЛЕНИЕМ РЕДКИХ ДАННЫХ

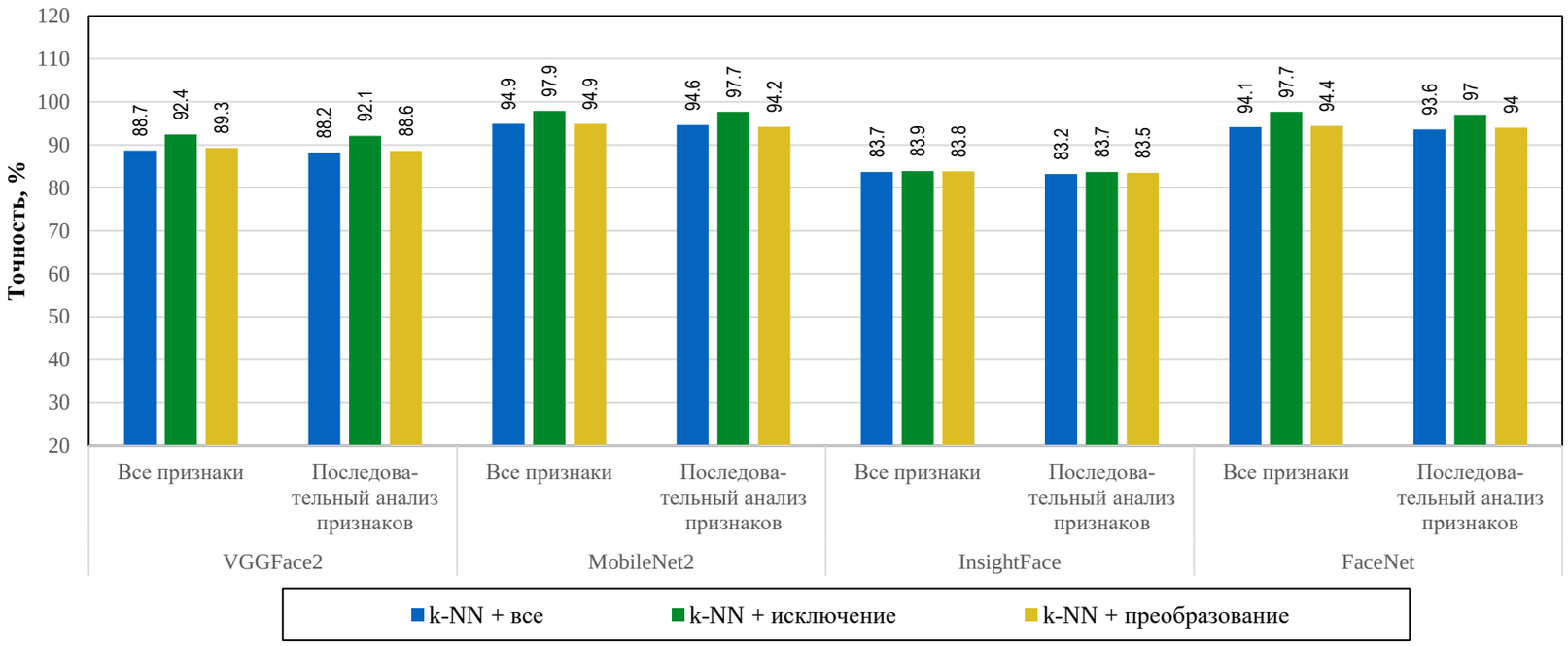
- Матрица ошибок определения «out-of-distribution» изображений (VGGFace2 набор данных)

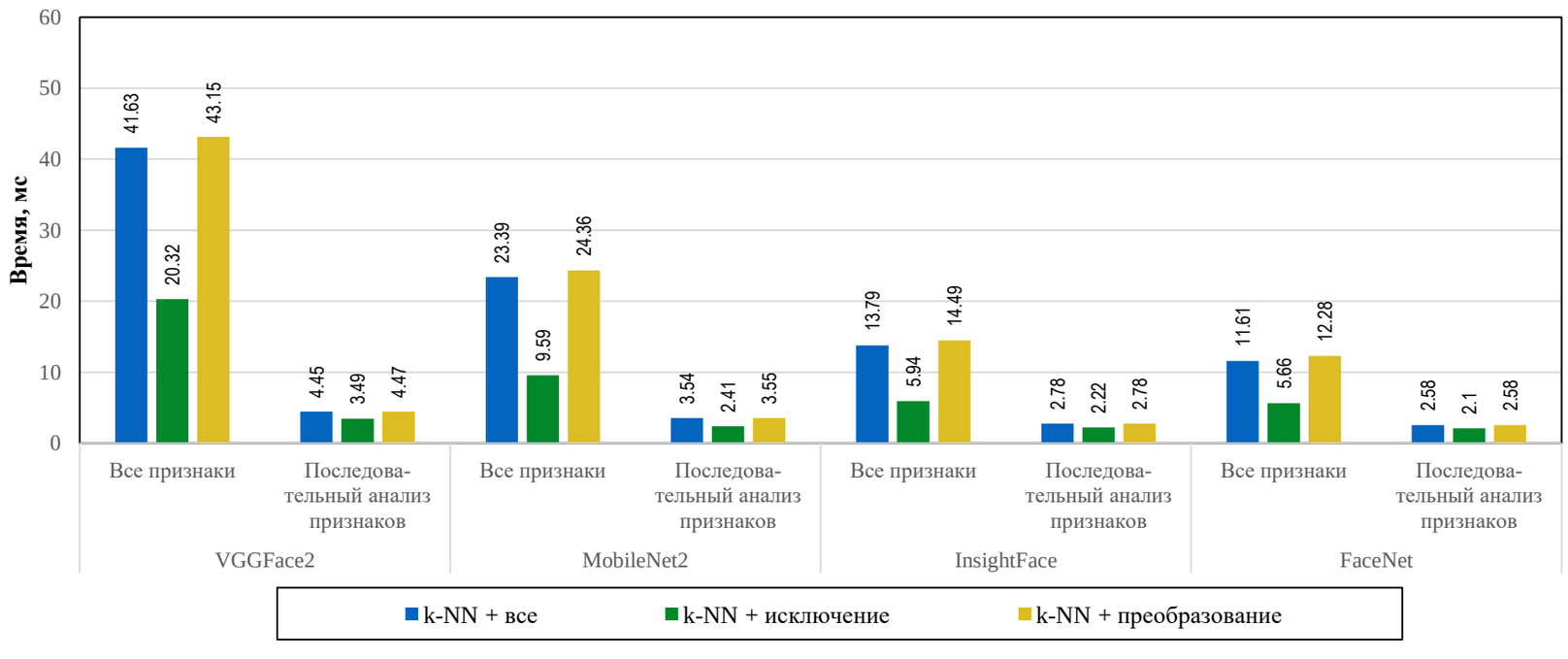
СНС	TPR	TNR	FPR	FNR
VGGFace2	97.29%	93.21%	6.79%	2.71%
MobileNet2	**98.44%**	**98.14%**	**1.86%**	**1.56%**
InsightFace	96.55%	89.57%	10.43%	3.45%
FaceNet	94.73%	90.52%	9.48%	5.27%

- Матрица ошибок определения «out-of-distribution» изображений (MS1M набор данных)

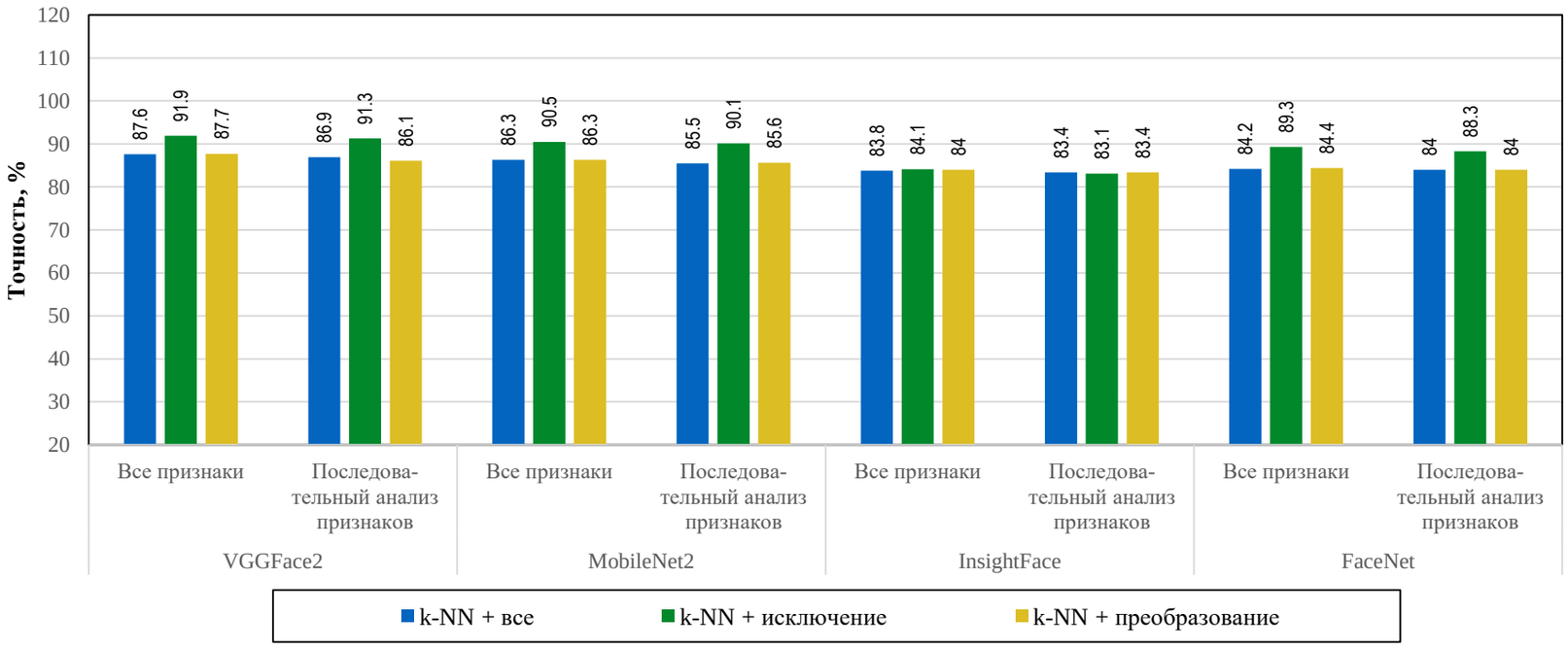
СНС	TPR	TNR	FPR	FNR
VGGFace2	96.54%	**92.46%**	**7.54%**	3.46%
MobileNet2	**97.28%**	91.57%	8.43%	**2.72%**
InsightFace	93.16%	88.95%	11.05%	6.84%
FaceNet	94.55%	90.69%	9.31%	5.45%

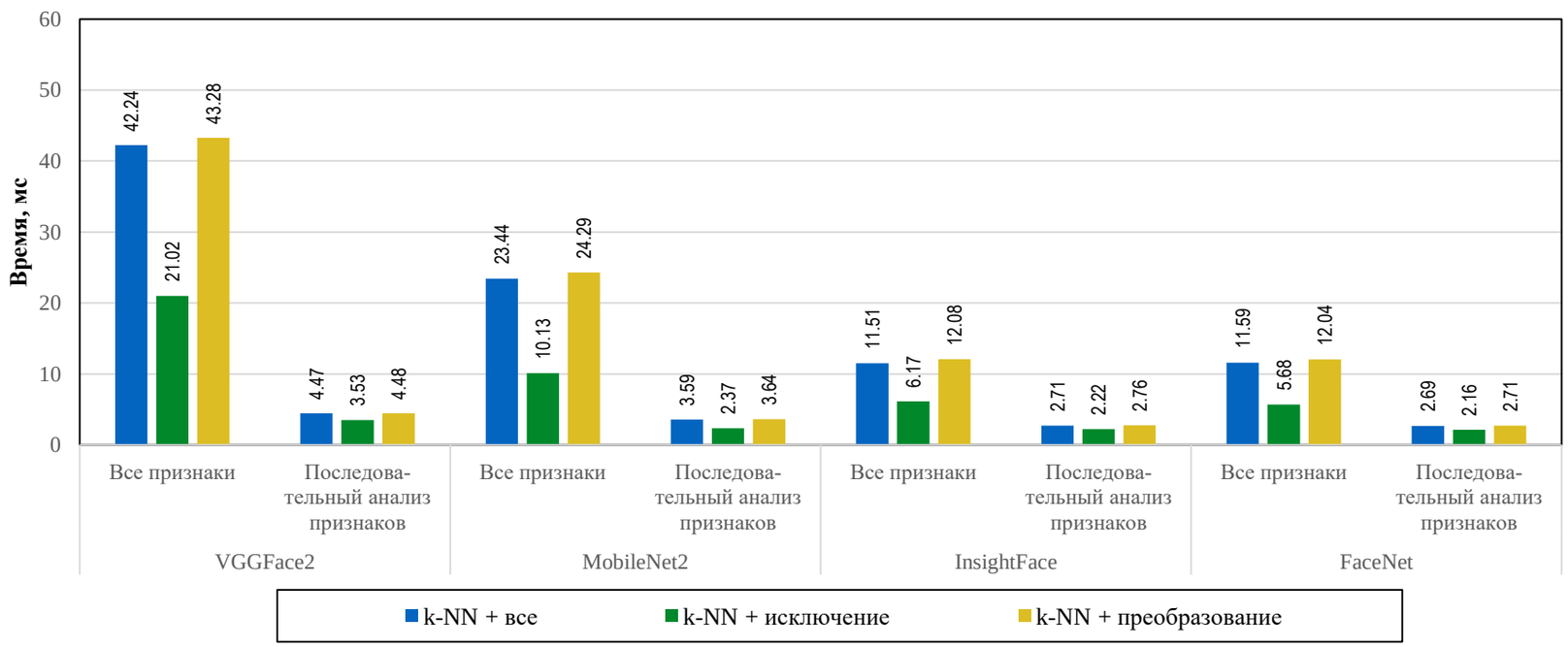
РЕЗУЛЬТАТ ПОСЛЕДОВАТЕЛЬНОЙ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ (VGGFACE2)



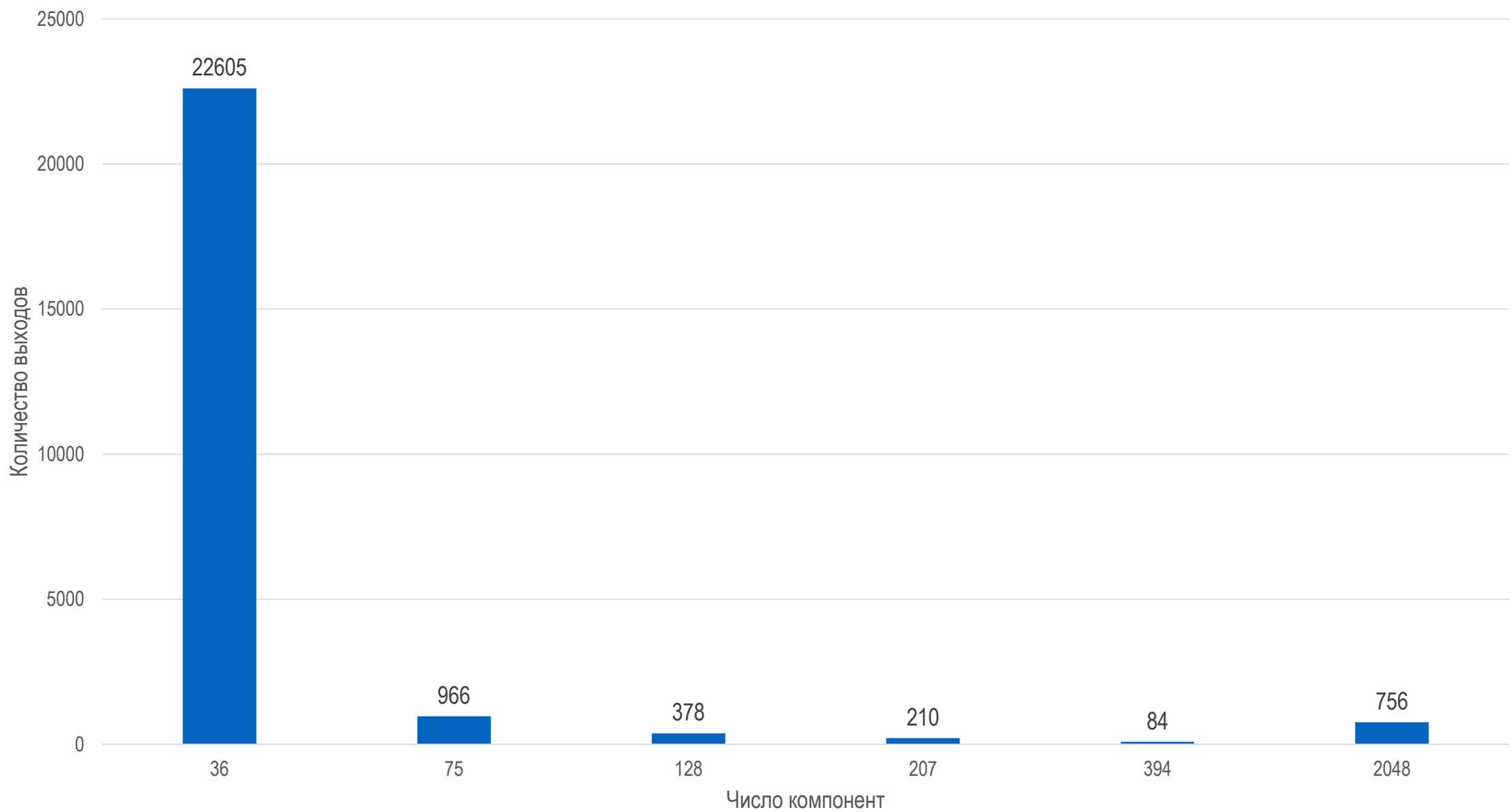


РЕЗУЛЬТАТ ПОСЛЕДОВАТЕЛЬНОЙ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ (MS1M)





РЕЗУЛЬТАТ ВЫХОДОВ НА ОПРЕДЕЛЕННОМ ЧИСЛЕ КОМПОНЕНТ (VGGFACE2)



РЕЗУЛЬТАТ ПОСЛЕДОВАТЕЛЬНОЙ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ (LFW)

Метрика	Классификатор	Отбор изображений	VGGFace2	MobileNet2	InsightFace	FaceNet
Точность (%)	k-NN, все признаки	все	**98.66**	**96.58**	**92.43**	**98.15**
		исключение	**98.71**	**96.61**	**92.51**	**98.21**
	Последовательный k-NN, фиксированное число компонент	все	98.32	96.24	91.78	98.15
		исключение	98.37	96.25	91.79	98.19
	Последовательный k-NN, число компонент по объясненной дисперсии	все	97.98	96.28	91.88	98.15
		исключение	98.02	96.32	91.90	98.18
Время (мс)	k-NN, все признаки	все	34.78	15.32	9.42	9.37
		исключение	34.72	15.25	9.21	9.19
	Последовательный k-NN, фиксированное число компонент	все	2.84	2.73	2.65	2.69
		исключение	2.79	2.69	2.41	2.42
	Последовательный k-NN, число компонент по объясненной дисперсии	все	**2.81**	**2.16**	**1.89**	**1.87**
		исключение	**2.78**	**2.14**	**1.74**	**1.73**

ПОКАЗАТЕЛИ ТОЧНОСТИ РАБОТЫ КЛАССИФИКАТОРОВ ИЗ БИБЛИОТЕКИ SCIKIT-LEARN (LFW)

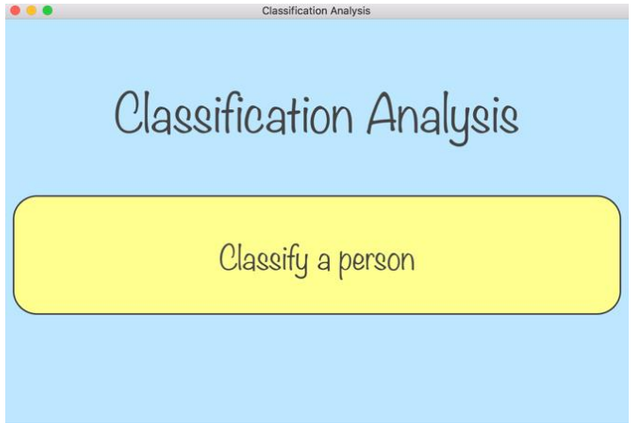
Точность (%)

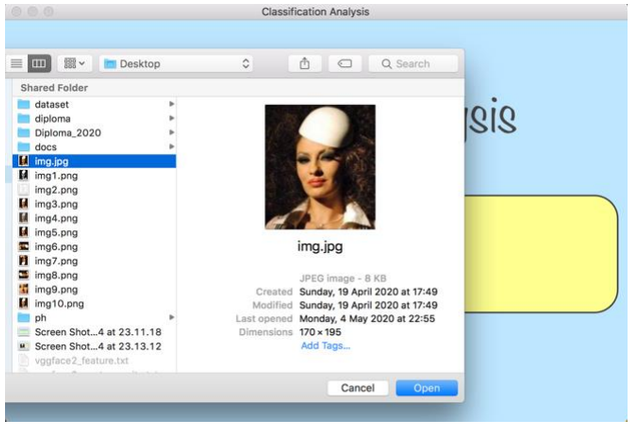
СНС	Число компонент	SVC	Random Forest	Extremely Randomized Trees	AdaBoost	Gradient Boosting
VGGFace2	Все признаки	14.00	81.50	81.50	19.30	22.41
	Фиксированное число компонент	0	79.50	82.10	17.40	21.01
MobileNet2	Все признаки	19.32	82.64	81.95	16.72	22.79
	Фиксированное число компонент	7.46	81.05	80.49	15.39	20.36
InsightFace	Все признаки	10.12	69.50	69.65	19.24	19.41
	Фиксированное число компонент	19.34	64.23	69.98	14.40	19.01
FaceNet	Все признаки	23.88	74.51	75.91	23.84	24.93
	Фиксированное число компонент	44.07	77.87	78.99	24.11	18.80

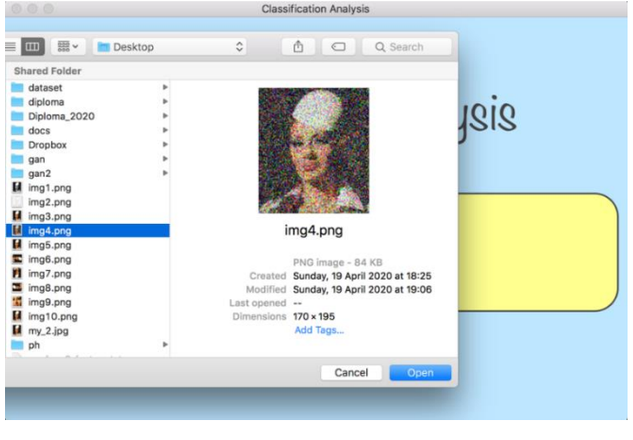
Время (мс)

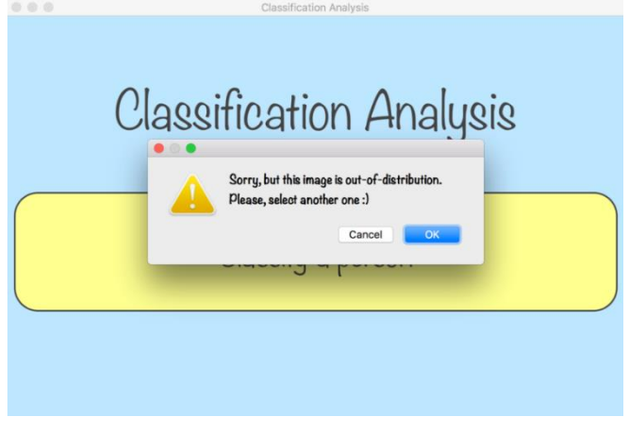
СНС	Число компонент	SVC	Random Forest	Extremely Randomized Trees	AdaBoost	Gradient Boosting
VGGFace2	Все признаки	7.17	16.55	1.00	24.98	0.23
	Фиксированное число компонент	3.20	16.14	0.69	21.15	0.16
MobileNet2	Все признаки	4.42	6.55	0.78	18.21	0.35
	Фиксированное число компонент	1.76	6.03	0.51	16.94	0.31
InsightFace	Все признаки	4.12	5.99	0.52	7.99	0.22
	Фиксированное число компонент	2.15	5.34	0.41	7.74	0.19
FaceNet	Все признаки	6.35	6.03	0.42	9.34	0.43
	Фиксированное число компонент	4.63	5.7	0.33	9.01	0.17

ПРОТОТИП СИСТЕМЫ









РЕЗУЛЬТАТЫ

- Проведен обзор современной литературы по темам классификации изображений, СНС, трансформации изображений;
- Собран набор данных, демонстрирующий различные виды данных, которые снижают точность работы классификаторов – такие данные определены как «редкие»;
- Описана блок схема системы классификации изображений людей, которая на вход принимает изображение, далее детектирует на нем область лица, извлекает из данной области вектор признаков, определяет какого распределения изображение, после этого только выполняется последовательная классификация;
- Реализован алгоритм классификации лиц, который позволяет улучшить точность с использованием предложенного метода;
- Разработан прототип системы с помощью кроссплатформенного фреймворка Qt, использующий предложенный подход

АПРОБАЦИЯ РЕЗУЛЬТАТОВ

- Исследование проходило в рамках научно-учебной группы «Анализ мультимедийных данных пользователей мобильных устройств», которая реализует проект № 19-04-004 «Эффективные методы распознавания мультимедийных данных для задач анализа предпочтений пользователей мобильных устройств»
- Планируется выступление с докладом на 10й Международной конференции по сетевому анализу NET2020 (в формате Zoom)
- Представлен постерный доклад на международном саммите «Computer Vision and Deep Learning summit» (Москва)
- Представлен доклад на международной конференции по анализу изображений, социальных сетей и текстов АИСТ19 (Казань)
- Представлен доклад на конференции с международным участием «Математические методы распознавания образов» (РАН, Москва)
- 2 статьи в Springer:

  «Fast Nearest-Neighbor Classifier Based on Sequential Analysis of Principal Components»
  «Computation-Efficient Face Recognition Algorithm Using a Sequential Analysis of High Dimensional Neural-Net Features»

СПАСИБО ЗА ВНИМАНИЕ!

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ