НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

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

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Упорядочивание данных в системах видеонаблюдения на основе технологий глубокого обучения

Студент группы 14БИ1:
Соколова Анастасия Дмитриевна

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

Нижний Новгород, 2018

АКТУАЛЬНОСТЬ РАБОТЫ

- Автоматическое распознавание объектов в сфере общественной безопасности
- Группировка видеоданных для ведения статистики
- Сокращение времени на поиски определенных объектов
- Облегчение работы пользователей

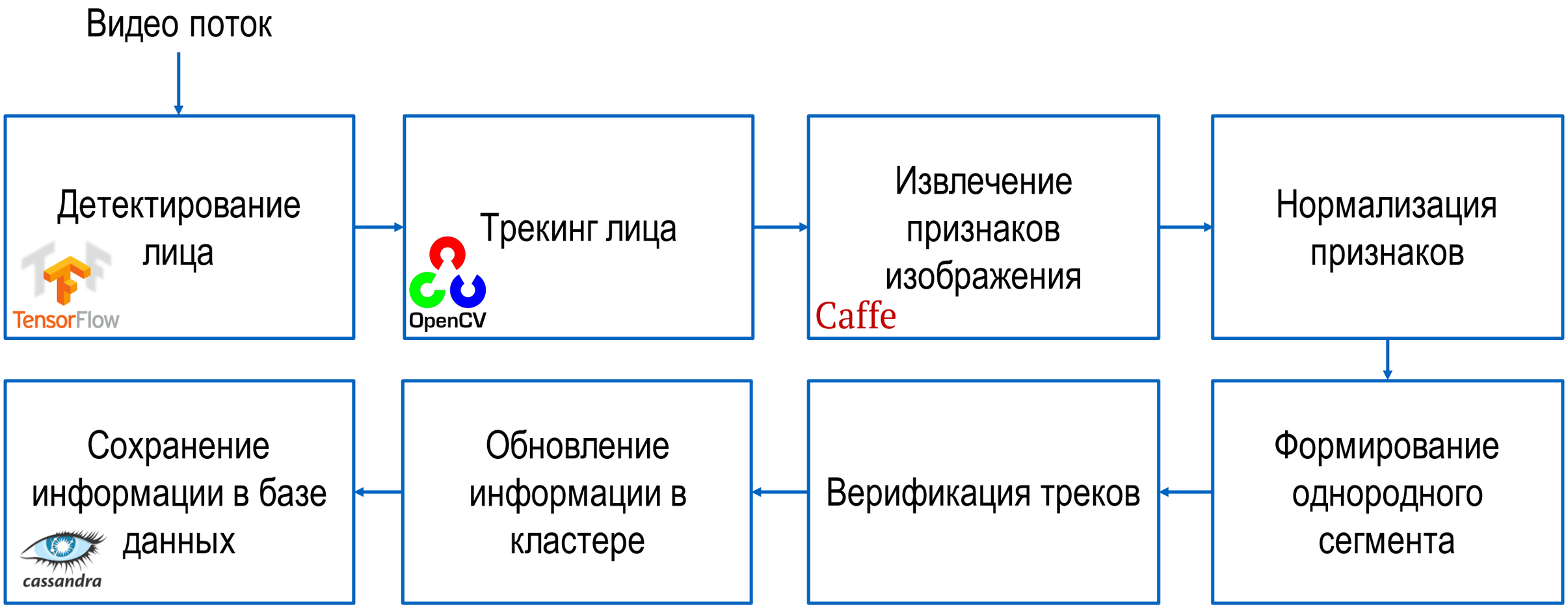
ЦЕЛЬ И ЗАДАЧИ РАБОТЫ

Цель: разработка системы видеонаблюдения с автоматической кластеризацией изображений на основе сверточных нейронных сетей, которая будет выявлять число различных людей на видео с меньшей вероятностью ошибки.

Задачи:

- Провести аналитический обзор современной литературы по кластеризации видеопоследовательностей лиц
- Описать схему работы алгоритма системы
- Разработать прототип системы кластеризации видеопоследовательностей
- Провести эксперименты реализованных алгоритмов

СХЕМА СИСТЕМЫ



БАЗЫ ДАННЫХ

- LFW (Labeled Faces in the Wild)
  - 1680 людей
  - 13000 изображений
  - 1-10 изображений в одном треке
- YTF (YouTube Faces)
  - 1595 людей
  - 3425 видео
  - 48-6070 изображений в одном треке
- CASIA WebFaces
  - 10500 людей
  - 494000 изображений
  - 1-42 изображений в одном треке
- IJB-B (IARPA Janus Benchmark-B)
  - 1895 людей
  - 7000 видео
  - 67000 изображений

http://vis-www.cs.umass.edu/lfw/
https://www.cs.tau.ac.il/~wolf/ytfaces/
https://www.nist.gov/itl/iad/image-group/ijb-dataset-request-form
https://blog.csdn.net/alpine_climbing/article/details/50502492

МЕТОДЫ ДЕТЕКТИРОВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ (CASIA WEBFACES)

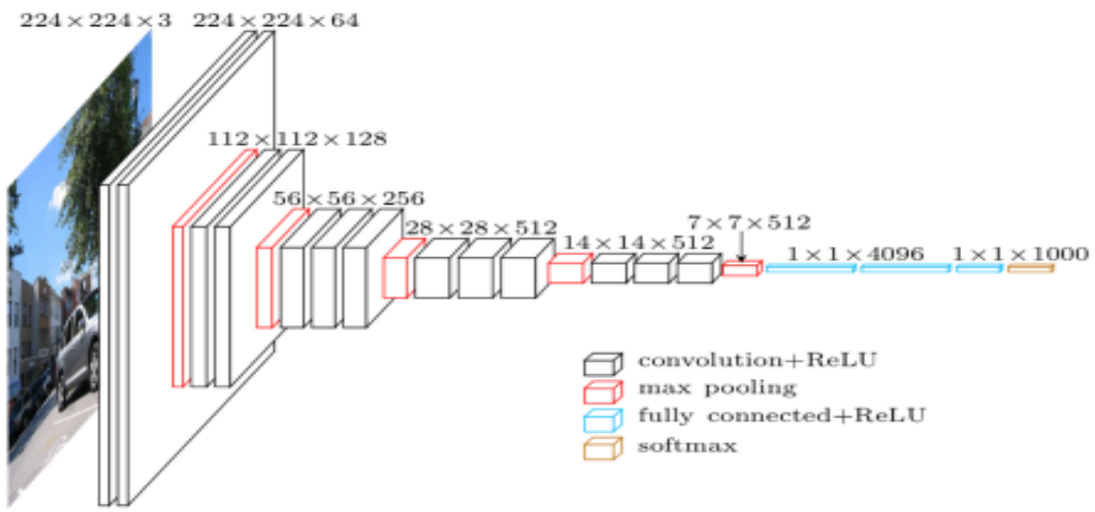
	Каскады Хаара	MobileNet SSD	ResNet-10
Precision	85%	94.4%	92%

Видеокарта GeForce GTX 970

Chen D. et al. Joint cascade face detection and alignment //European Conference on Computer Vision – 2014
https://github.com/yeephycho/tensorflow-face-detection
https://github.com/opencv/opencv/tree/master/samples/dnn/face_detector

ИЗВЛЕЧЕНИЕ ПРИЗНАКОВ ИЗОБРАЖЕНИЯ

- OpenFace
  1x128 вектор
- Lightened CNN (version C)
  1x256 вектор
- VggFace2
  1x2048 вектор
- VGGNet
  1x4096 вектор



Caffe

ВЫДЕЛЕНИЕ ОДНОРОДНОГО СЕГМЕНТА (LFW)

0,4285

0,521

0,8934

АЛГОРИТМЫ ВЕРИФИКАЦИИ

- Медоид

$$\rho(X(m_1), X(m_2)) = \rho(x^*(m_1),\ x^*(m_2)),$$

$$x^*(m_1) = \underset{x(t),t\in[t_1(m_i),t_2(m_i)],t`=t_1(m_i)}{\operatorname{argmin}} \sum_{t`=t_1(m_i)}^{t_2(m_i)} \rho(x(t), x(t`))$$

- Среднее расстояние между всеми кадрами двух треков

$$\rho(X(m_1), X(m_2)) = \frac{1}{\Delta t(m_1),\Delta t(m_2)} \sum_{t=t_1(m_1)}^{t_2(m_1)} \sum_{t`=t_1(m_2)}^{t_2(m_2)} \rho(x(t), x(t`))$$

- Средний вектор расстояний одного трека

$$\rho(X(m_1), X(m_2)) = \rho\left(\tilde{x}(m_1), \tilde{x}(m_2)\right),$$

$$\tilde{x}(m_i) = \frac{1}{\Delta t(m_i)} \sum_{t`=t_1(m_i)}^{t_2(m_i)} x(t`)$$

- Агрегированный вектор

$$a(t) = \frac{\exp(q^T x(t))}{\sum_{t`=t_1(m)}^{t_2(m)} \exp(q^T x(t`))}, \qquad r(m) = \sum_{t=t_1(m)}^{t_2(m)} a(t)x(t)$$

Sokolova A., Savchenko A. Cluster analysis of facial video data in video surveillance systems using deep learning // Large Scale Networks - Computational Aspects and Applications. Springer Proceedings in Mathematics and Statistics – 2018

АЛГОРИТМЫ ВЕРИФИКАЦИИ

- Медиана

$$\rho(X(m_1), X(m_2)) = \rho\big(x`(m_1), x`(m_2)\big),$$
$$x`(m_i) = median$$

- Среднее расстояние между элементами и дисперсия расстояния внутри трека, а также между элементами двух разных треков

$$\bar{\rho}_{in} = \frac{2}{N_{1q}(N_{1q}-1)} \sum_{i=1}^{N_{1q}} \sum_{j=i+1}^{N_{1q}} \rho_{ij},$$

$$D(\rho_{in}) = \frac{2}{N_{1q}(N_{1q}-1)} \sum_{i=1}^{N_{1q}} \sum_{j=i+1}^{N_{1q}} \rho_{ij}^2 - \bar{\rho}_{in}^2,$$

$$\bar{\rho}_{ex} = \frac{2}{N_{1q}(N_{2q}-1)} \sum_{i=1}^{N_{1q}} \sum_{j=i+1}^{N_{2q}} \rho_{ij},$$

$$D(\rho_{ex}) = \frac{2}{N_{1q}(N_{2q}-1)} \sum_{i=1}^{N_{1q}} \sum_{j=i+1}^{N_{2q}} \rho_{ij}^2 - \bar{\rho}_{ex}^2,$$

$$\begin{cases} \bar{\rho}_{ex} - \bar{\rho}_{in} < K\sqrt{D(\rho_{in})} \\ \bar{\rho}_{ex} - \bar{\rho}_{in} < K\sqrt{D(\rho_{ex})} \end{cases}$$

Шемагина О.В. et al. Системы обнаружения, сопровождения и кластеризации объектов на основе нейроноподобного кодирования //Информационно-Измерительные И Управляющие Системы - 2010

ВСПОМОГАТЕЛЬНЫЕ ВЫЧИСЛЕНИЯ

- Нормализация вектора

$$\|x\| = \sqrt{\sum_i |x_i|^2},$$

$$x_i = \frac{x_i}{\|x\|}, \forall i = 0, \dots, N-1$$

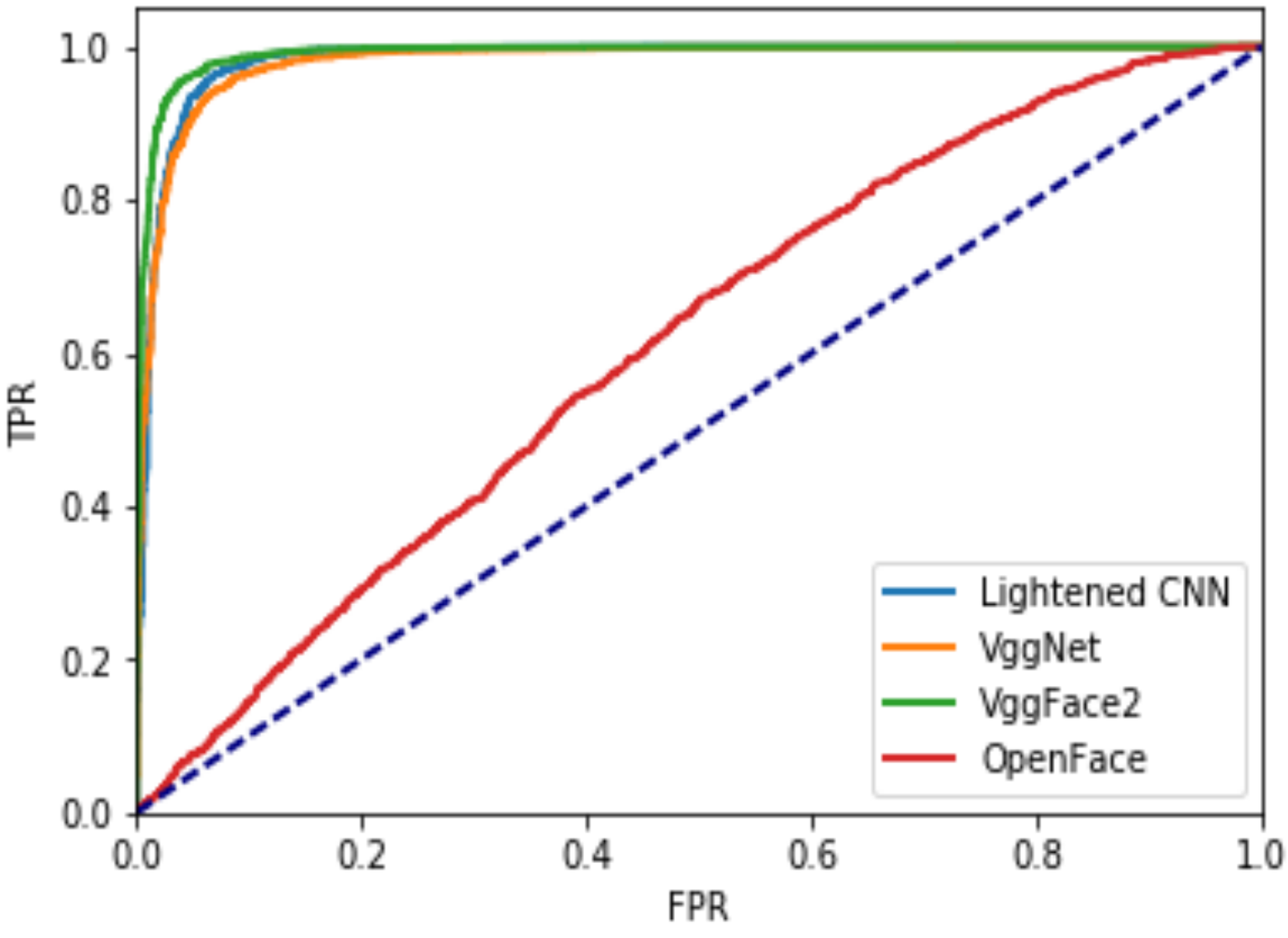
- Вычисление меры близости
  - Евклидово расстояние

$$\rho(x,y) = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1}(x_i - y_i)^2}$$

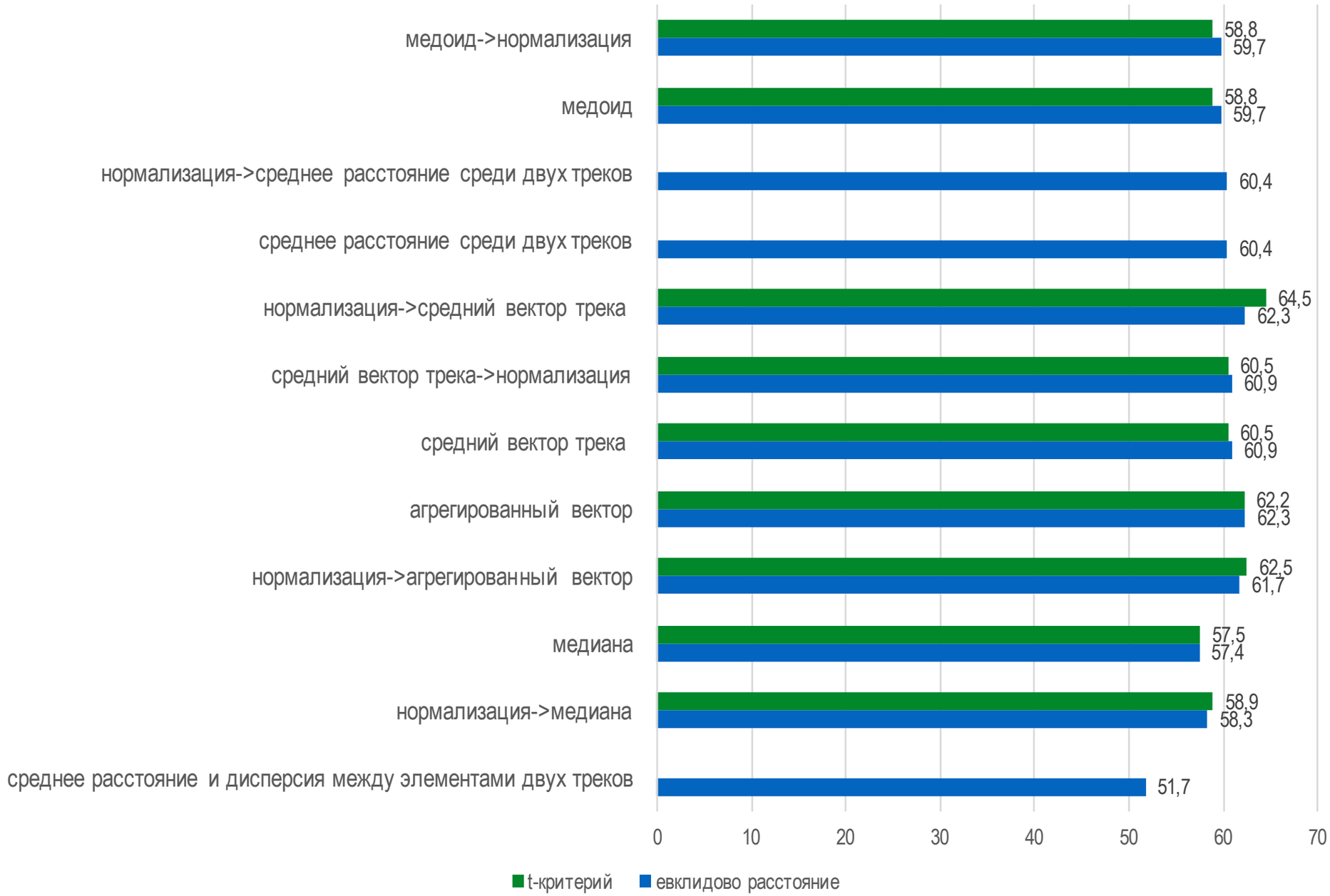
  - t - критерий

$$t = \frac{\rho(x,y)}{\sqrt{\frac{D_x^2}{N_x} + \frac{D_y^2}{N_y}}}$$

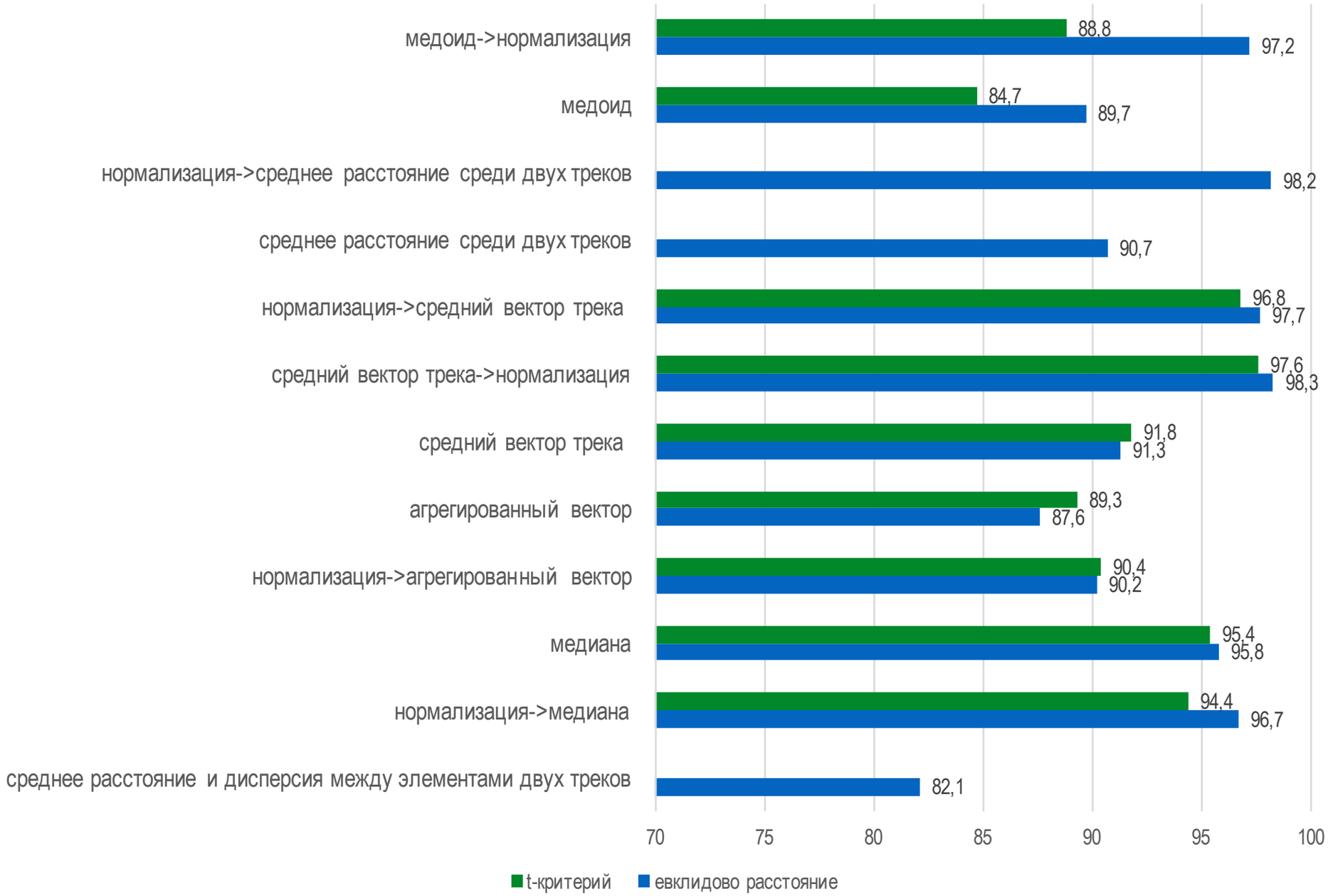
ROC-КРИВЫЕ СРАВНЕНИЯ СРЕДНЕГО ВЕКТОРА НОРМИРОВАННЫХ ПРИЗНАКОВ



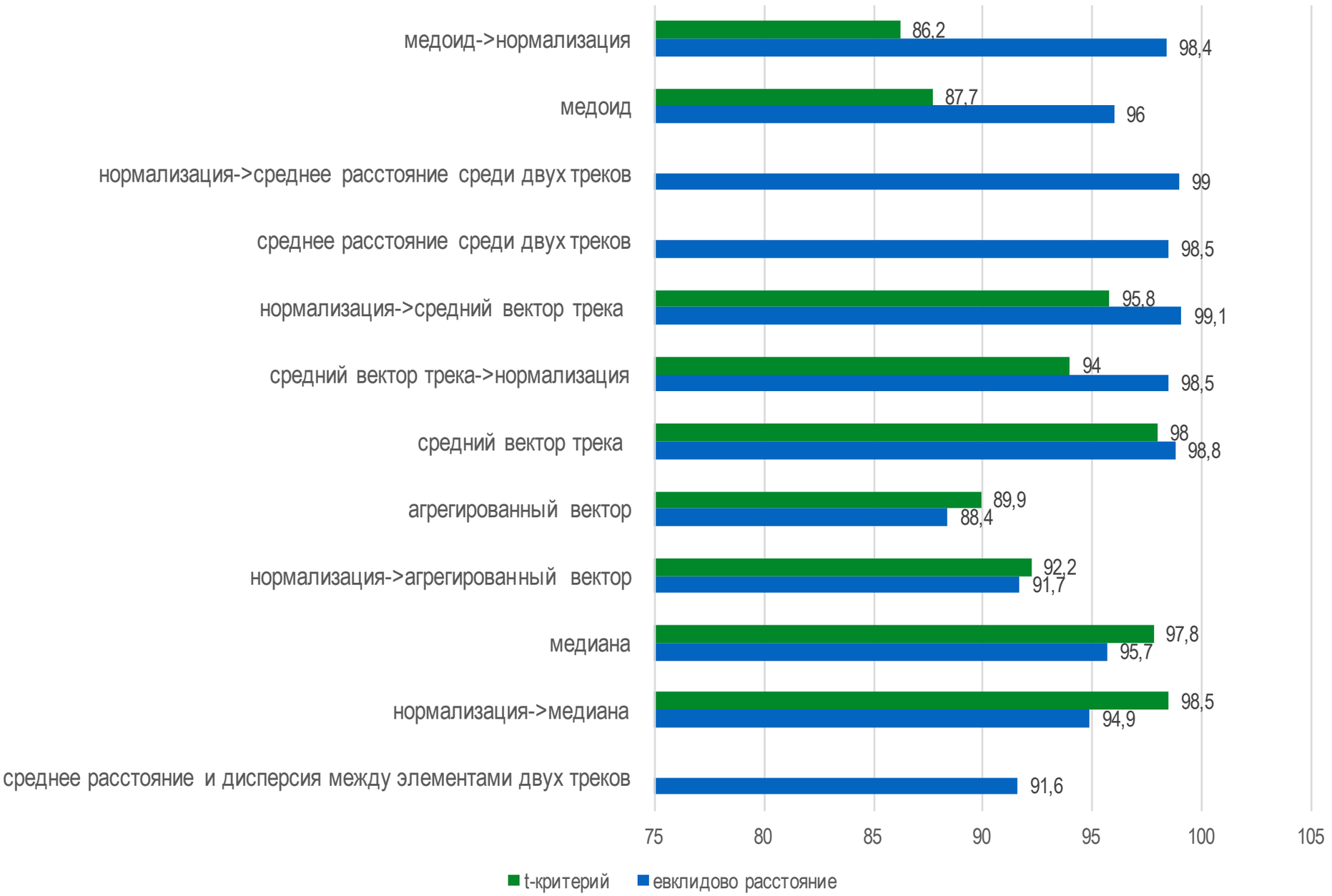
РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ. AREA UNDER CURVE. OPENFACE



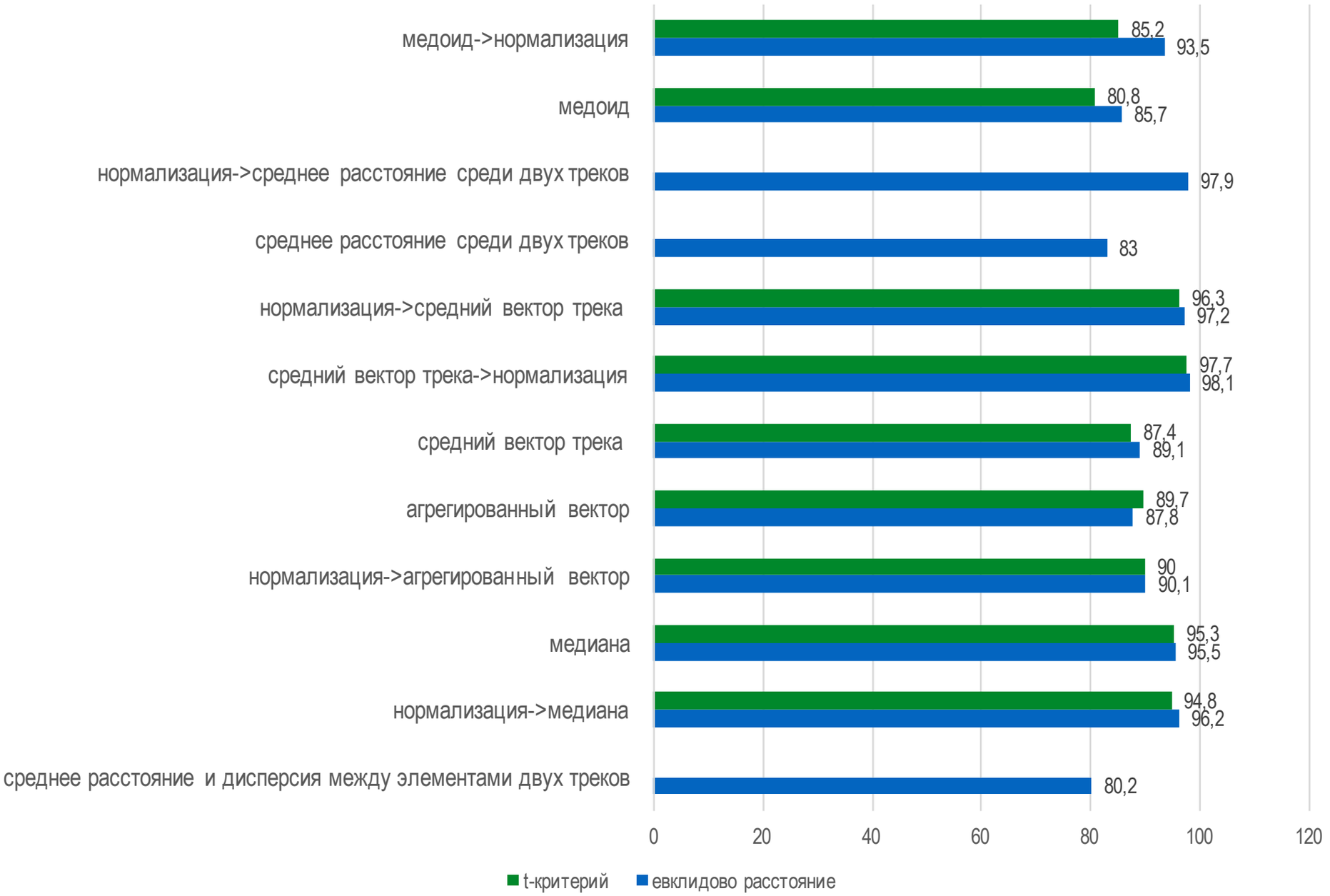
РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ. AREA UNDER CURVE. LIGHTENED CNN



РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ. AREA UNDER CURVE. VGGFACE2

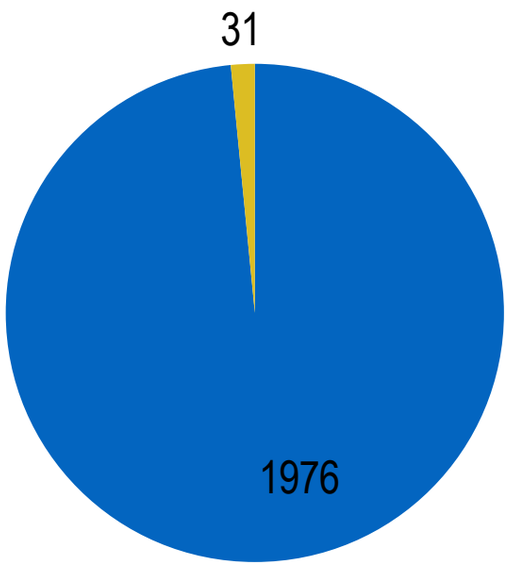


РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ. AREA UNDER CURVE. VGGNET

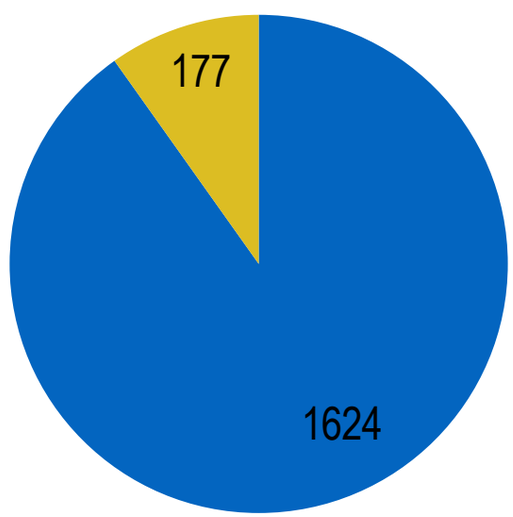


КЛАСТЕРИЗАЦИЯ (YTF)

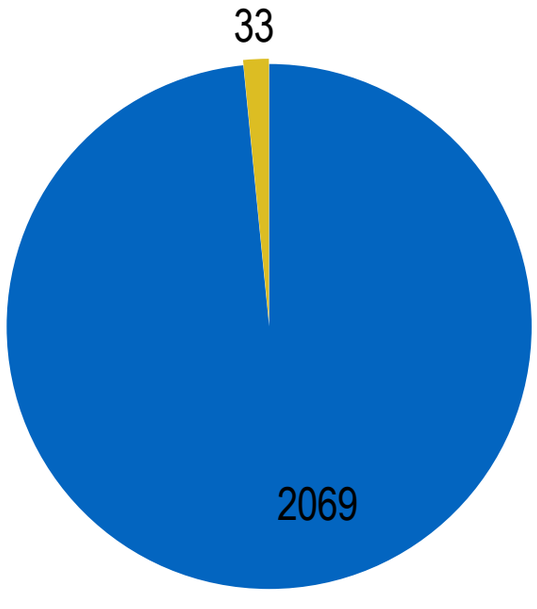
Rank-Order



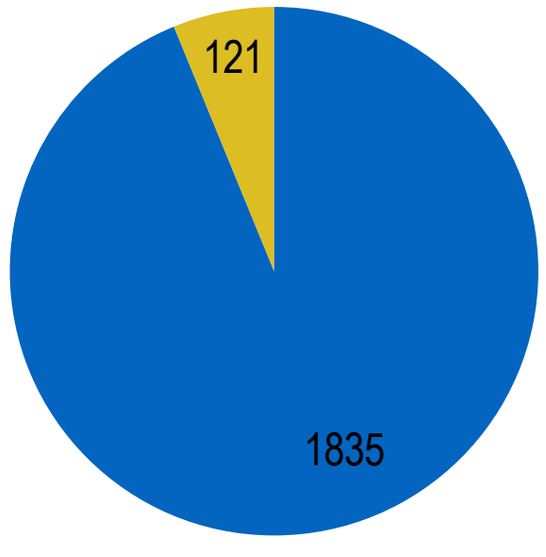
DominantSet



Последовательная кластеризация (FAR = 1%)

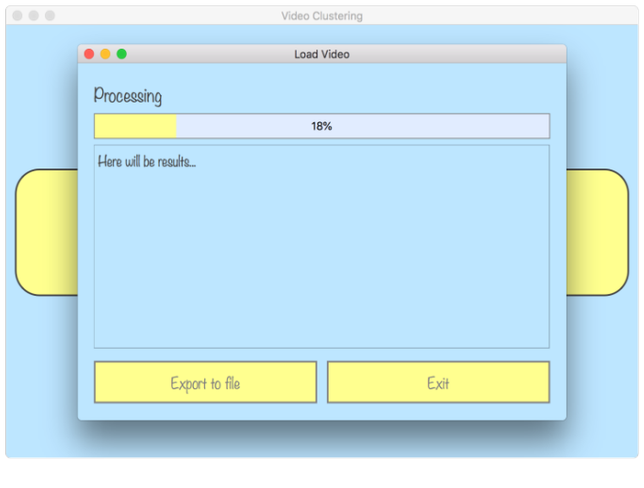


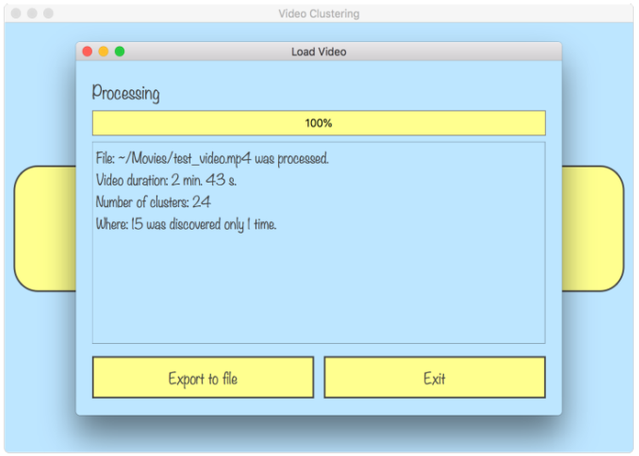
Последовательная кластеризация (FAR = 10%)

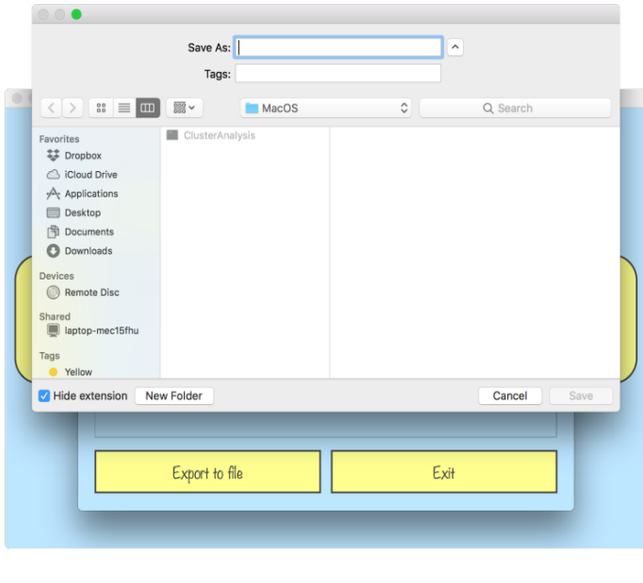


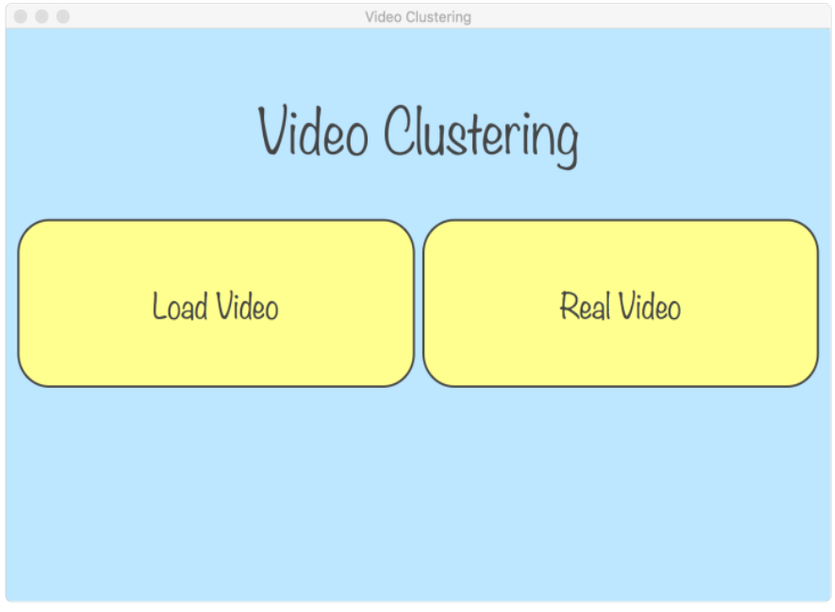
Zhu C. et al. A rank-order distance based clustering algorithm for face tagging //Computer Vision and Pattern Recognition – 2011
M. Pavan and M. Pelillo. Dominant sets and pairwise clustering //PAMI – 2007

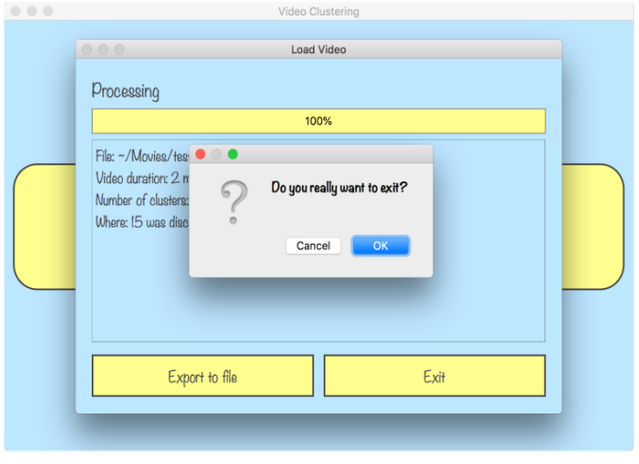
ПРОТОТИП СИСТЕМЫ

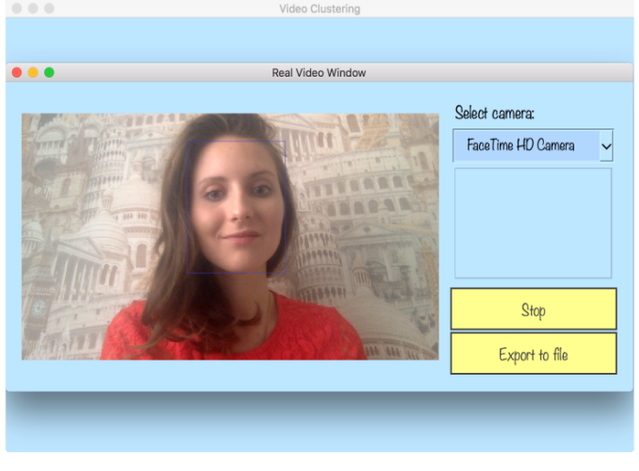


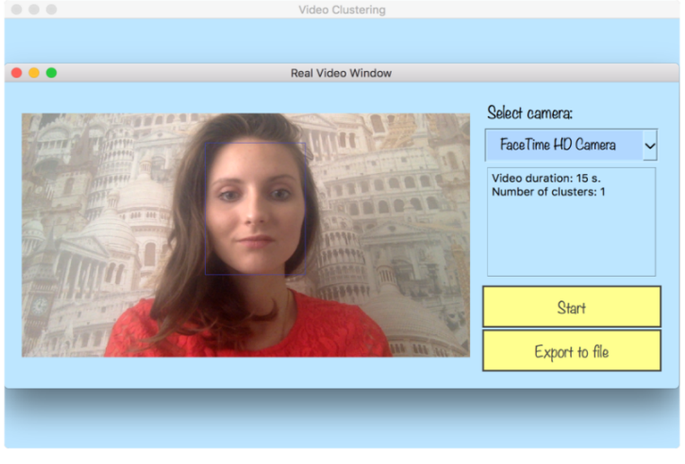




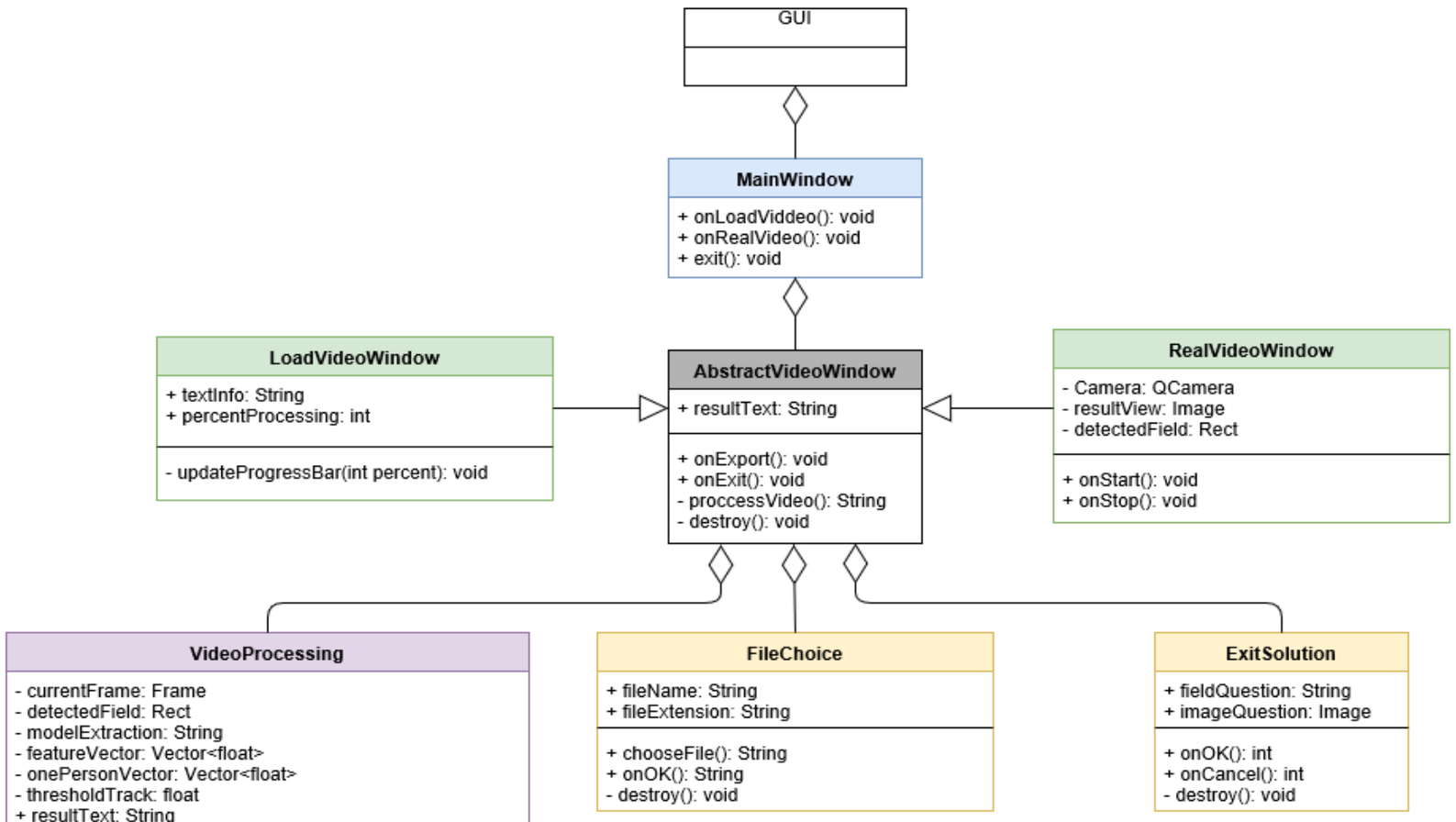


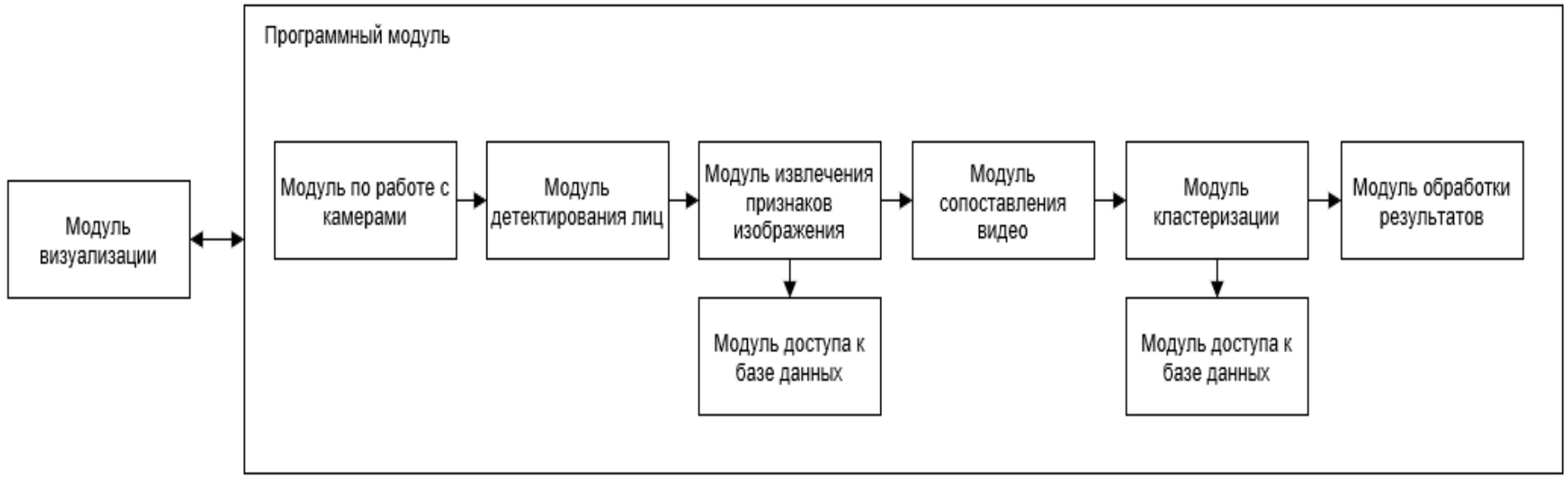






АРХИТЕКТУРА СИСТЕМЫ





РЕЗУЛЬТАТЫ

- Проведен обзор современной литературы на тему сверточных нейронных сетей, детектирования и распознавания объектов на изображении, их кластеризации
- Разработана схема реализации системы видеонаблюдения с автоматической кластеризацией
- Реализованы алгоритмы агрегации векторов признаков в задаче верификации лиц на видео. Экспериментально показано, что наилучшей точностью характеризуется алгоритм поиска среднего вектора одного трека нормализованных признаков
- Разработан прототип системы видеонаблюдения с автоматической кластеризацией с учетом алгоритмов, показывающих лучшие результаты по экспериментам

АПРОБАЦИЯ РЕЗУЛЬТАТОВ

2017	Опубликована статья в сборнике трудов конференции AIST (Analysis of Images, Social networks and Texts) 2017 в престижной серии Lecture Notes in Computer Science издательства Springer
2017	Результаты представлены на зимней школе по аналитике данных (НИУ ВШЭ - НН, ЛАТАС)
2018	Презентация исследования организации видеоданных на Международной конференции «Информационные технологии и нанотехнологии 2018» на базе Самарского национального исследовательского университета имени академика С.П. Королева и Института систем обработки изображений РАН
2018	Награждение вторым местом на студенческой конференции «Интеллектуальный город: наука как искусство» в секции «Прикладная математика и информатика: социальные и бизнес-приложения» (НИУ ВШЭ - НН)
2018	Получение диплома лауреата в конкурсе научно – исследовательских работ студентов бакалавриата НИУ ВШЭ по направлению «Бизнес – информатика»
Планы	Участие в международном саммите «Computer Vision and Deep Learning summit» (г. Москва)

СПАСИБО ЗА ВНИМАНИЕ!

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ