НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

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

Харчевникова Ангелина Сергеевна

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Методы видеоаналитики для мобильных платформ в задаче определения пола и возраста покупателей

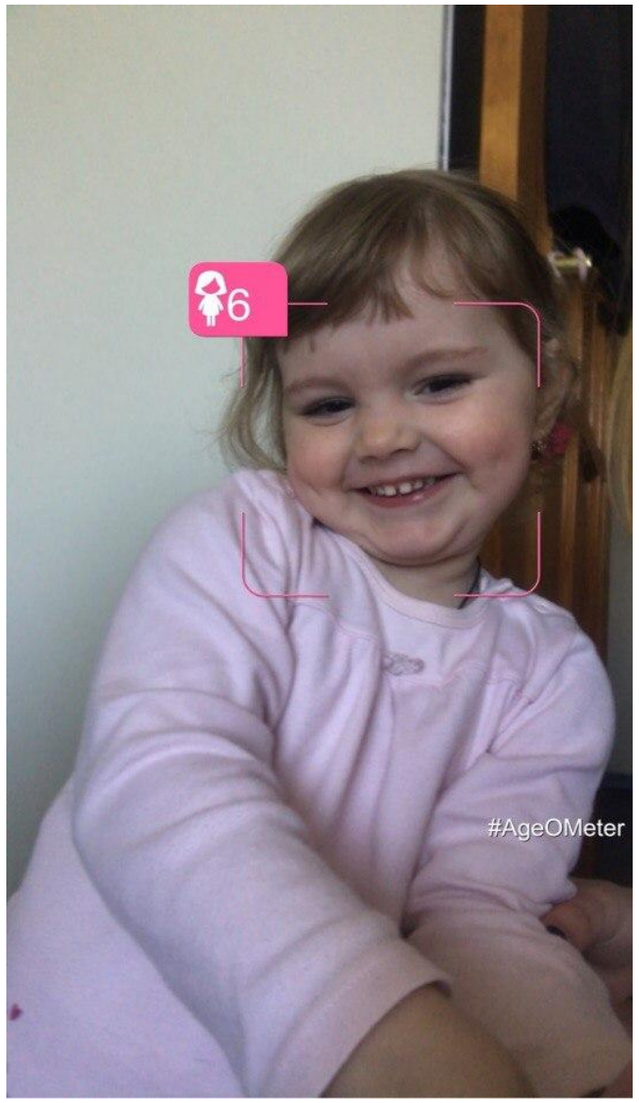
Научный руководитель:
Доктор технических наук, профессор
кафедры ИСиТ,
Савченко Андрей Владимирович

Нижний Новгород, 2018

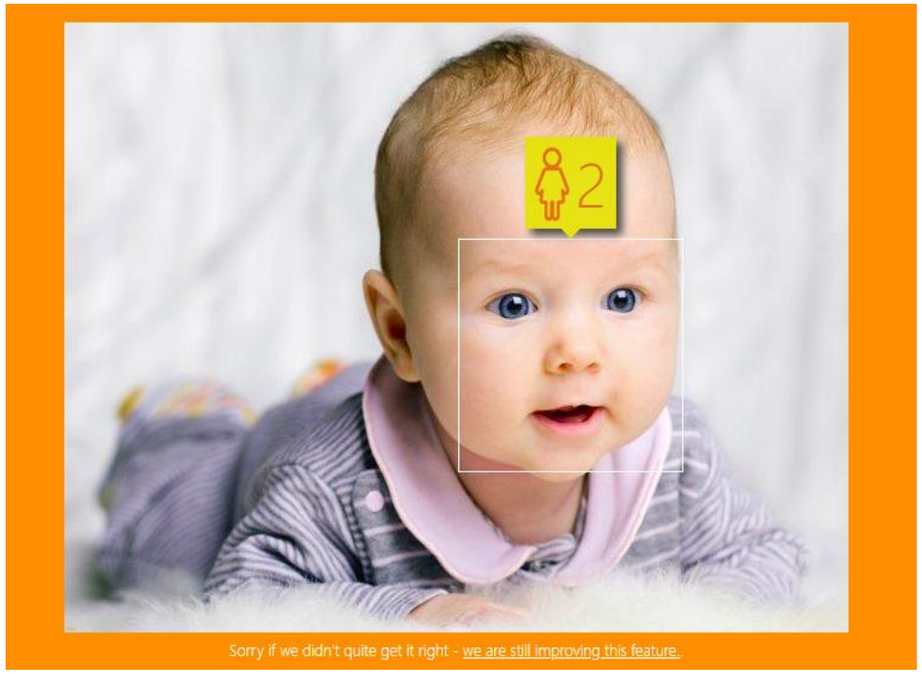
Предметная область

В зависимости от автоматически выявленных данных о клиенте, система выдает информацию, соответствующую конкретной целевой аудитории, что способствует увеличению продаж

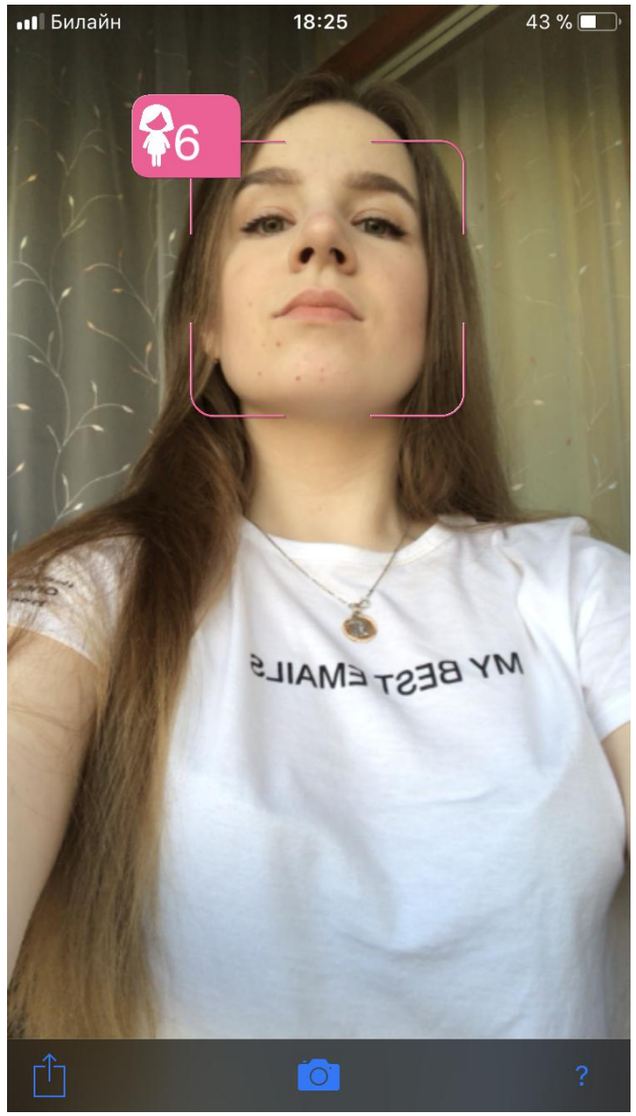
Актуальность



Luxand FaceSDK



Microsoft How-old.net



Luxand FaceSDK

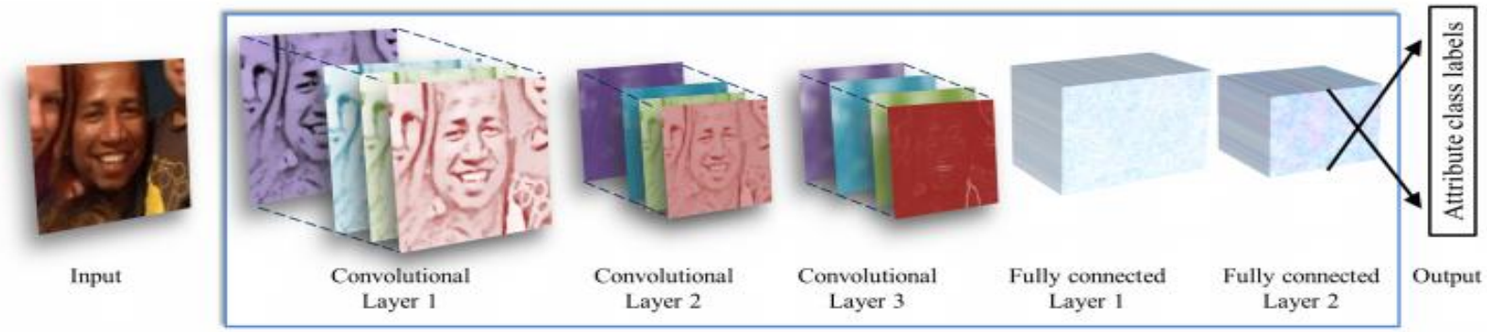
Внедрение подобных технологий в сфере розничной торговли все еще остается на стадии испытаний, так как надежность предложенных решений недостаточна для практического применения

Цель и задачи

Разработка прототипа мобильного офлайн-приложения для распознавания пола и возраста по видеоизображению лица на основании сверточных нейронных сетей, с применением комитетов классификаторов для повышения точности идентификации

- Провести обзор существующих решений распознавания пола и возраста по изображению лица с акцентированием внимания на сверточные нейронные сети
- Описать алгоритм предлагаемой системы с учетов методов агрегации решений
- Разработать прототип системы на платформе Android
- Провести экспериментальные исследования по анализу точности и производительности работы каждого из алгоритмов

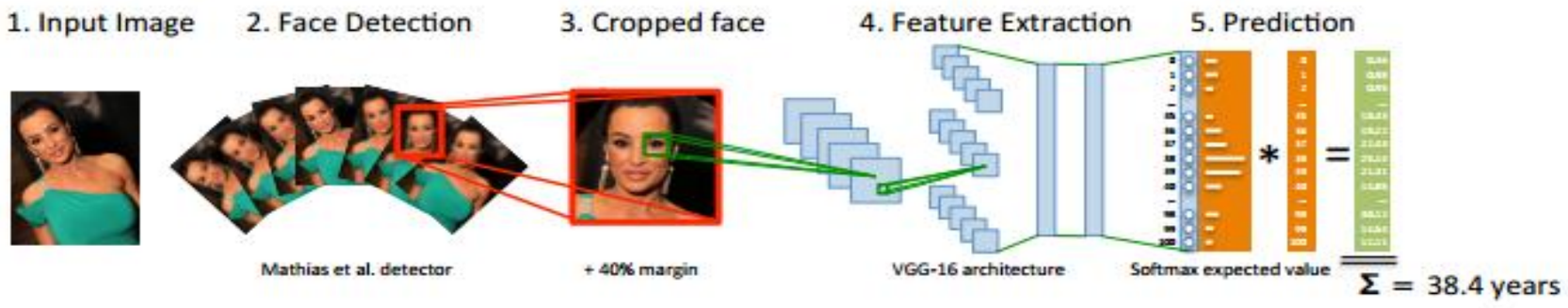
Сверточные нейронные сети



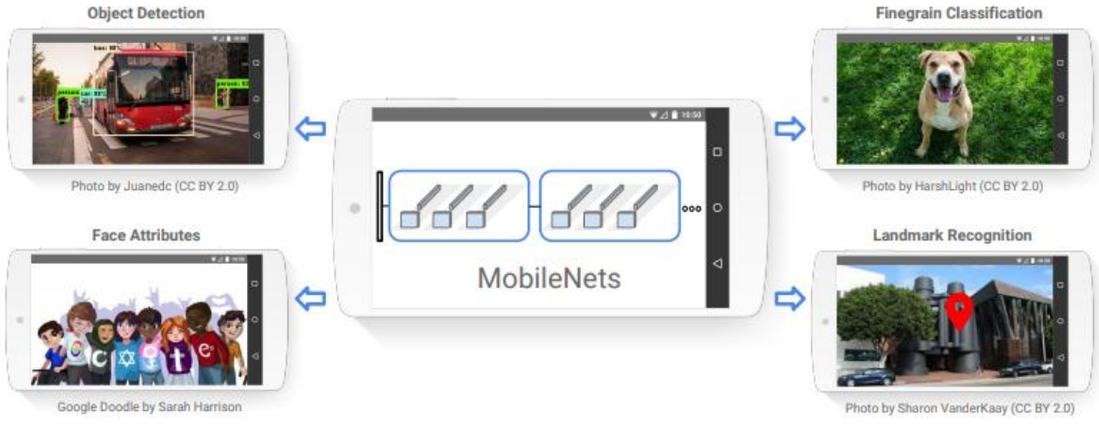
44 Мб

Gender net and Age net Levi, G. Age and gender classification using convolutional neural networks 2015

514 Мб



VGG-16 Rasmus Rothe DEX: Deep EXpectation of apparent age from a single image 2015



MobileNet **14 Мб**

MobileNet-finetuned **13 Мб**

MobileNet-quantized **4 Мб**

Комитеты классификаторов

$$P\left(l\middle|X(t)\right)=softmax\, z_l(t)=\frac{exp(z_l(t))}{\sum_{j=1}^{L} exp(z_j(t))},l=1,2,\ldots,L$$

Выход нейросетевой модели - слой Softmax() **(1)**

$$l^*(t)=\underset{l=1,2,\ldots,L}{argmax}\, P\left(l\middle|X(t)\right)$$

Принятие решения в пользу класса с максимальной апостериорной вероятностью (традиционный подход) **(2)**

$$l^*=\underset{l=\overline{1,L}}{argmax}\sum_{t=1}^{T}\delta(l^*(t)-l)$$

Простое голосование **(3)**

$$l^*=\underset{l=\overline{1,L}}{argmax}\frac{1}{T}\sum_{t=1}^{T}P\left(l\middle|X(t)\right)$$

Правило суммы **(4)**

$$l^*=\underset{l=\overline{1,L}}{argmax}\left(\prod_{t=1}^{T}P\left(l\middle|X(t)\right)\right)^{1/T}=\underset{l=\overline{1,L}}{argmax}\,exp\left(\frac{1}{T}\sum_{t=1}^{T}log\, P\left(l\middle|X(t)\right)\right)$$

Правило произведения **(5)**

Комитеты классификаторов. Математическое ожидание

$$l^* = \sum_{l=1}^{L} P\left(l \middle| X(t)\right) \cdot l$$

Оценка математического ожидания **(6)**

$$l^* = \frac{\sum_{k=1}^{K} P^*\left(l_(k) \middle| X(t)\right) \cdot l_(k)}{\sum_{i=1}^{K} P^*\left(l_(i) \middle| X(t)\right)}, \mathrm{K} \in \{(1)\ldots(L)\}$$

Оценка математического ожидания для К наиболее вероятных классов **(7)**

$$P^*\left(l_(1) \middle| X(t)\right) \geq P^*\left(l_(i) \middle| X(t)\right) \geq P^*\left(l_(L) \middle| X(t)\right)$$

Комитеты классификаторов. Метод Демпстера-Шафера

$$DT_l = \frac{1}{N_l} \sum_{z_k \in l} P(l \mid X(z_k))$$

Построение шаблонов решения с помощью базы данных изображений **(8)**

$$\Phi = \frac{\left(1 + \left\| DT_l^i - D_i(x) \right\|\right)^{-1}}{\sum_{k=1}^{L} \left(1 + \left\| DT_l^{i\,k} - D_i(x) \right\|\right)^{-1}}$$

Вычисление близости между DT^i_l и решением для *i-го* фрейма **(9)**

$$b_l(D_i(x)) = \frac{\Phi_{l,i}(x) \prod_{k \neq l} (1 - \Phi_{k,i}(x))}{1 - \Phi_{l,i}(x) \left[1 - \prod_{k \neq l} (1 - \Phi_{k,i}(x))\right]}$$

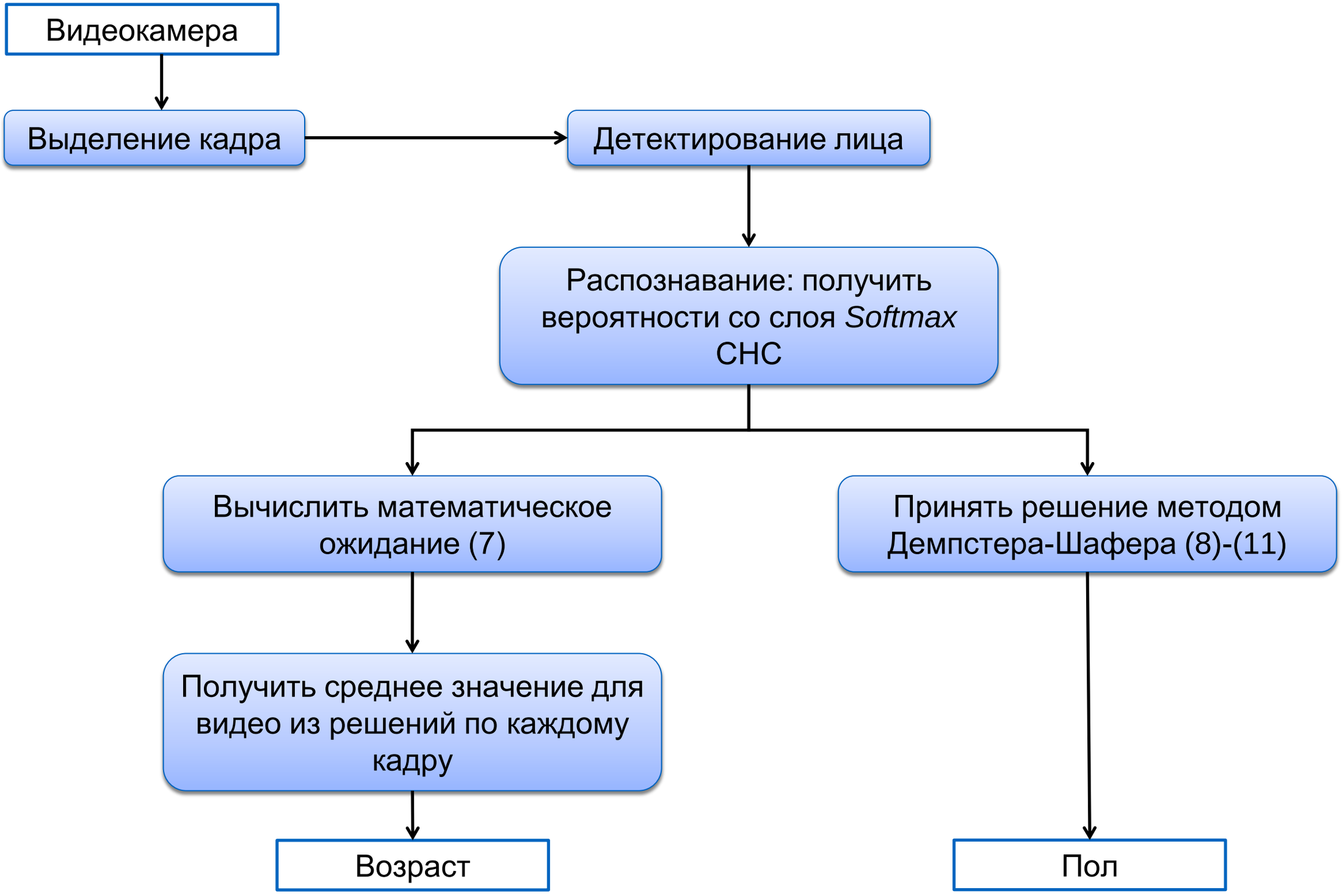
Вычисление степени доверия **(10)**

$$\mu_l(x) = K \prod_{i=1} b_l(D_i(x))$$

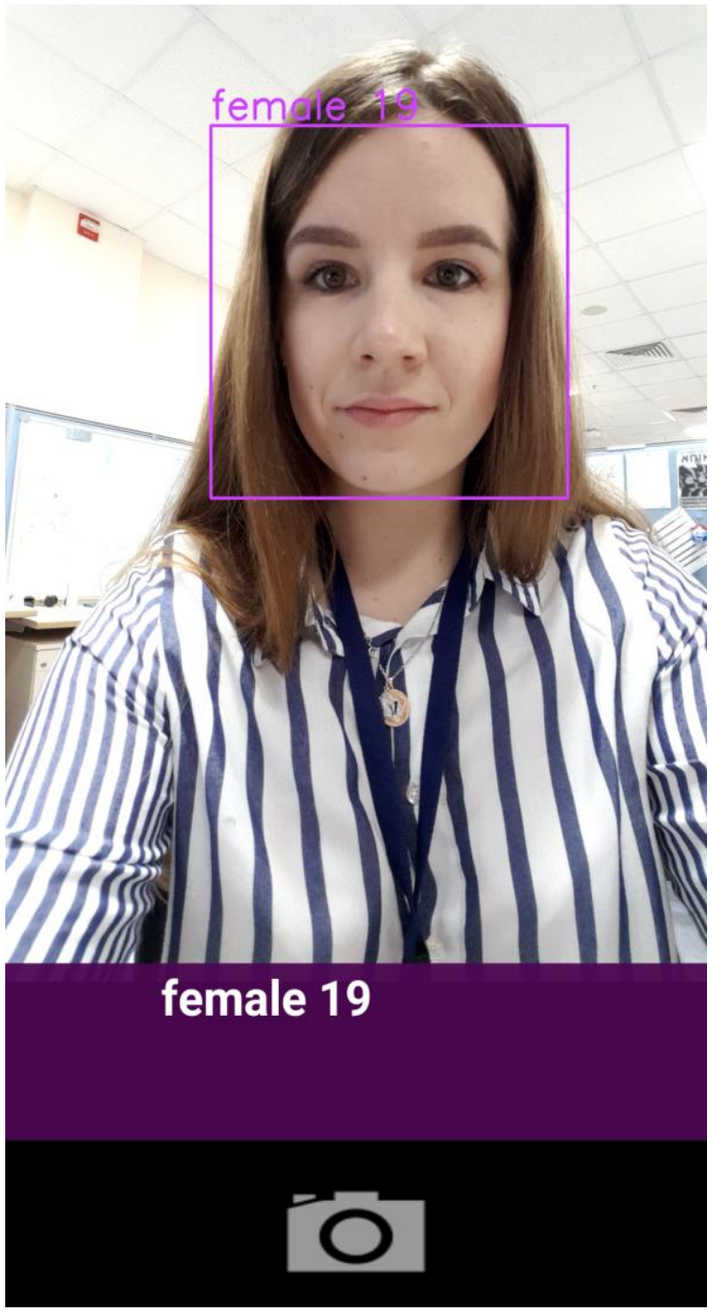
Финальное решение в пользу максимального значения $\mu_l(x)$ **(11)**

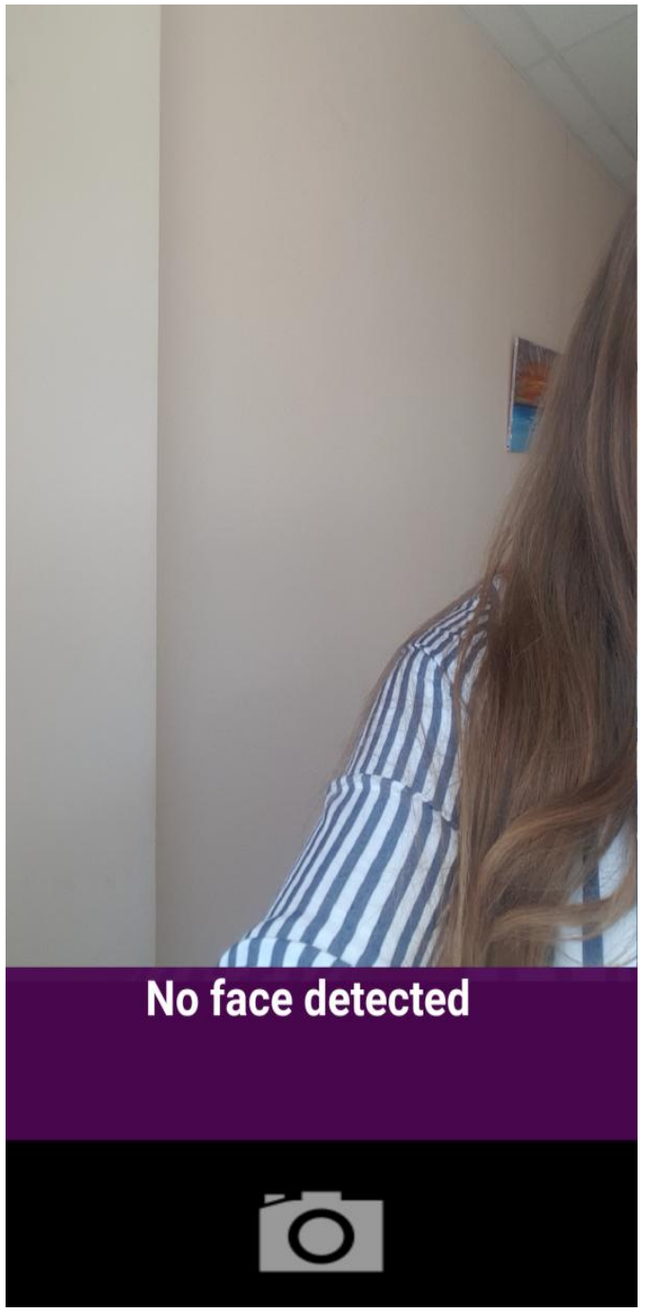
Bagheri, M. A. Logo recognition based on the dempster-shafer fusion of multiple classifiers In *CanadianConference on Artificial Intelligence*. 1-12p. Springer, Berlin, Heidelberg. 2013

Структурная схема системы



Реализация системы на мобильной платформе

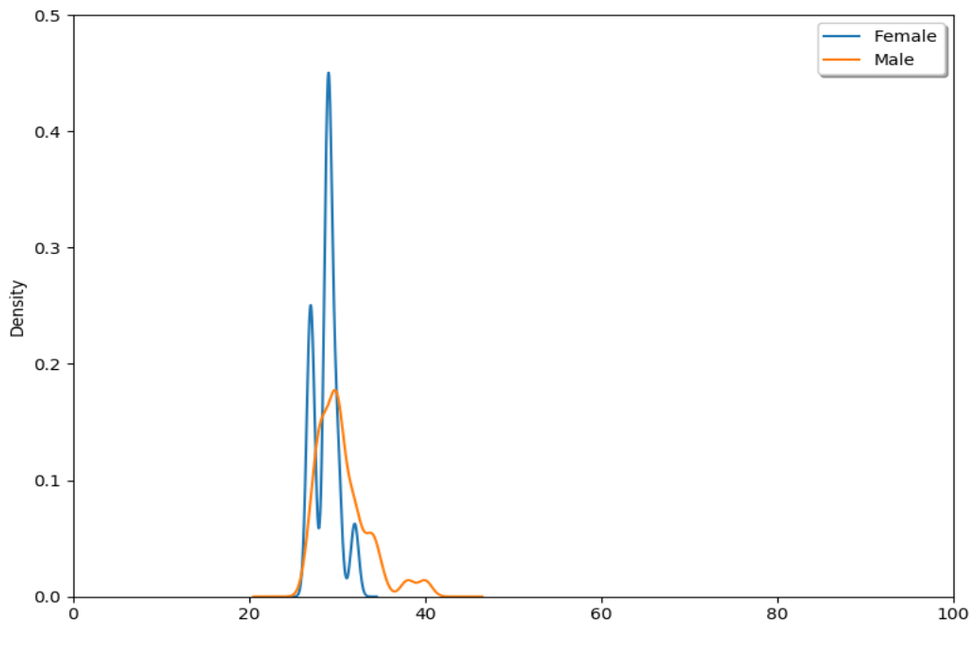




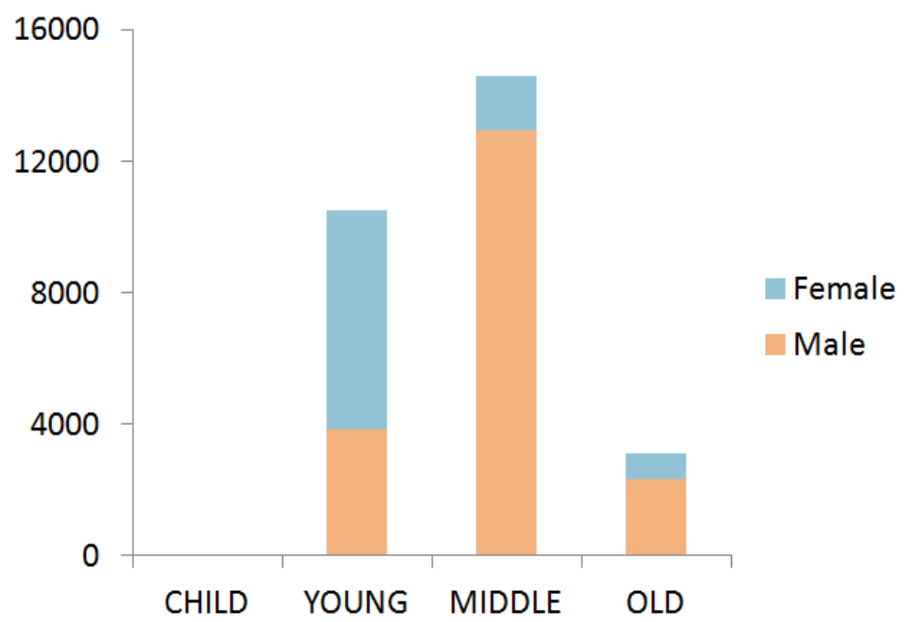
Результаты экспериментов. Время распознавания

Модель СНС	Время, с
MobileNet	0.371
MobileNet_ft	0.368
MobileNet_quantized	0.373
Gender/Age net	1.076
VGG16	273

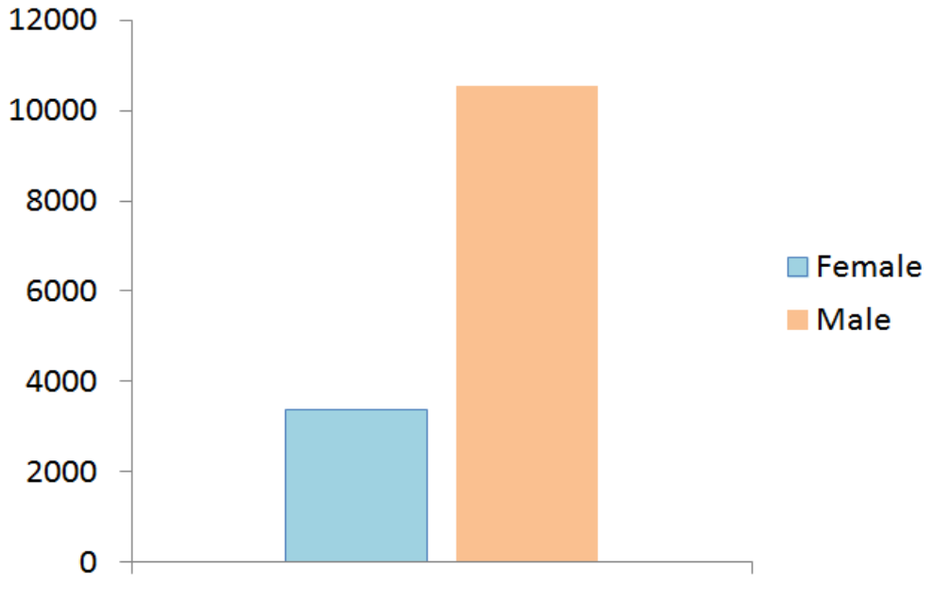
Результаты экспериментов. Наборы данных



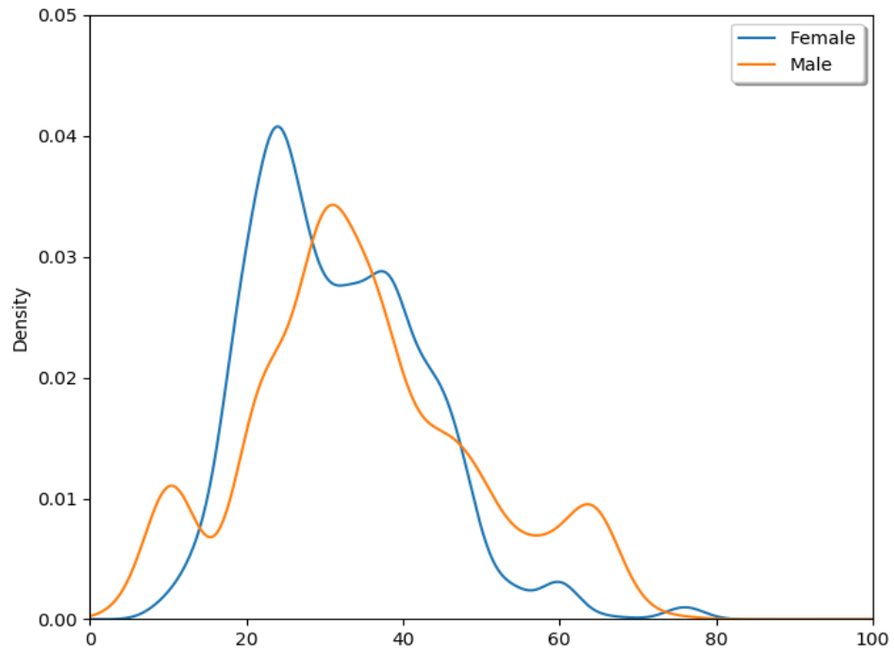
Kinect



Indian Movie



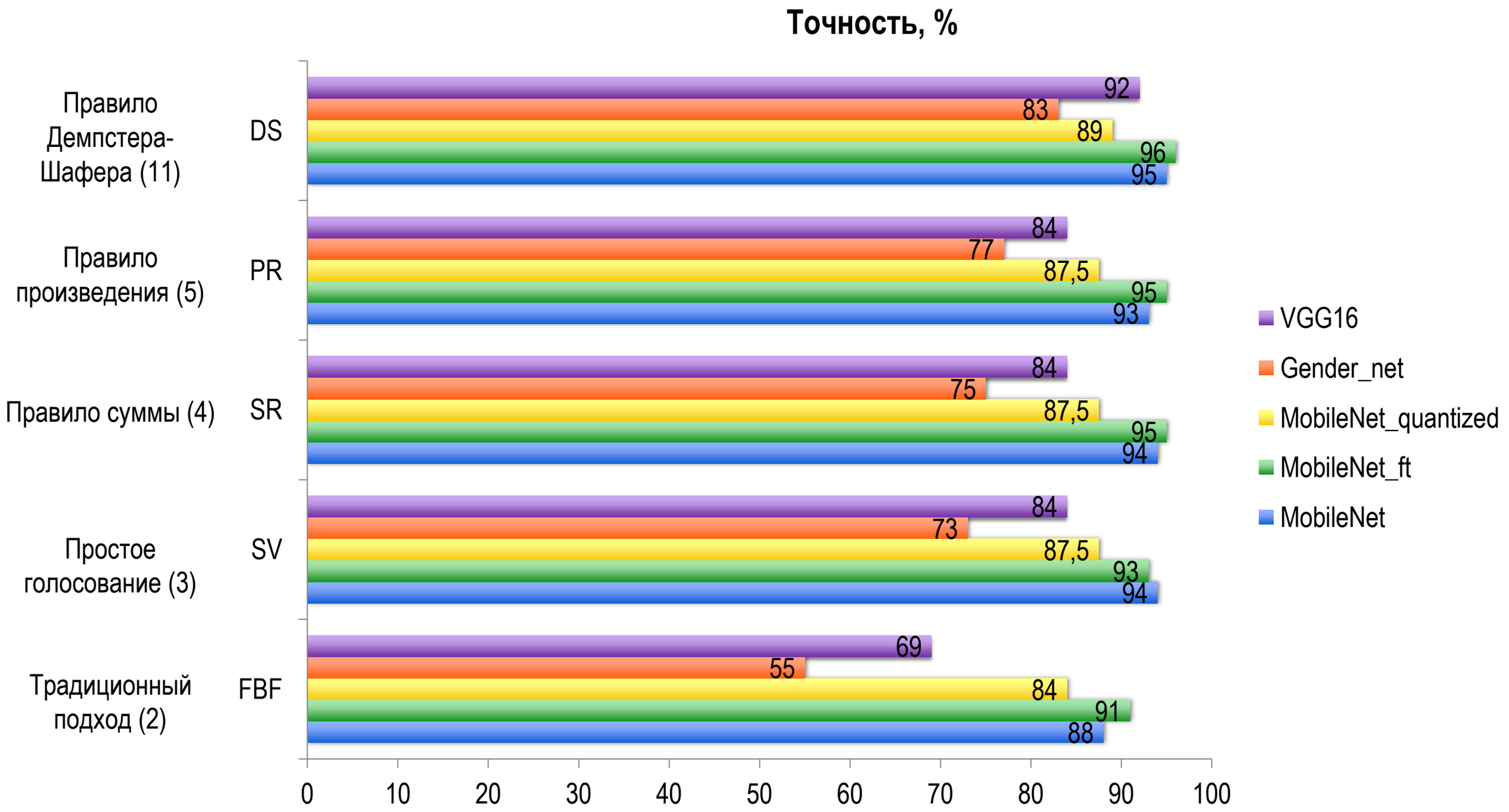
IJB-A



EmotiW2018

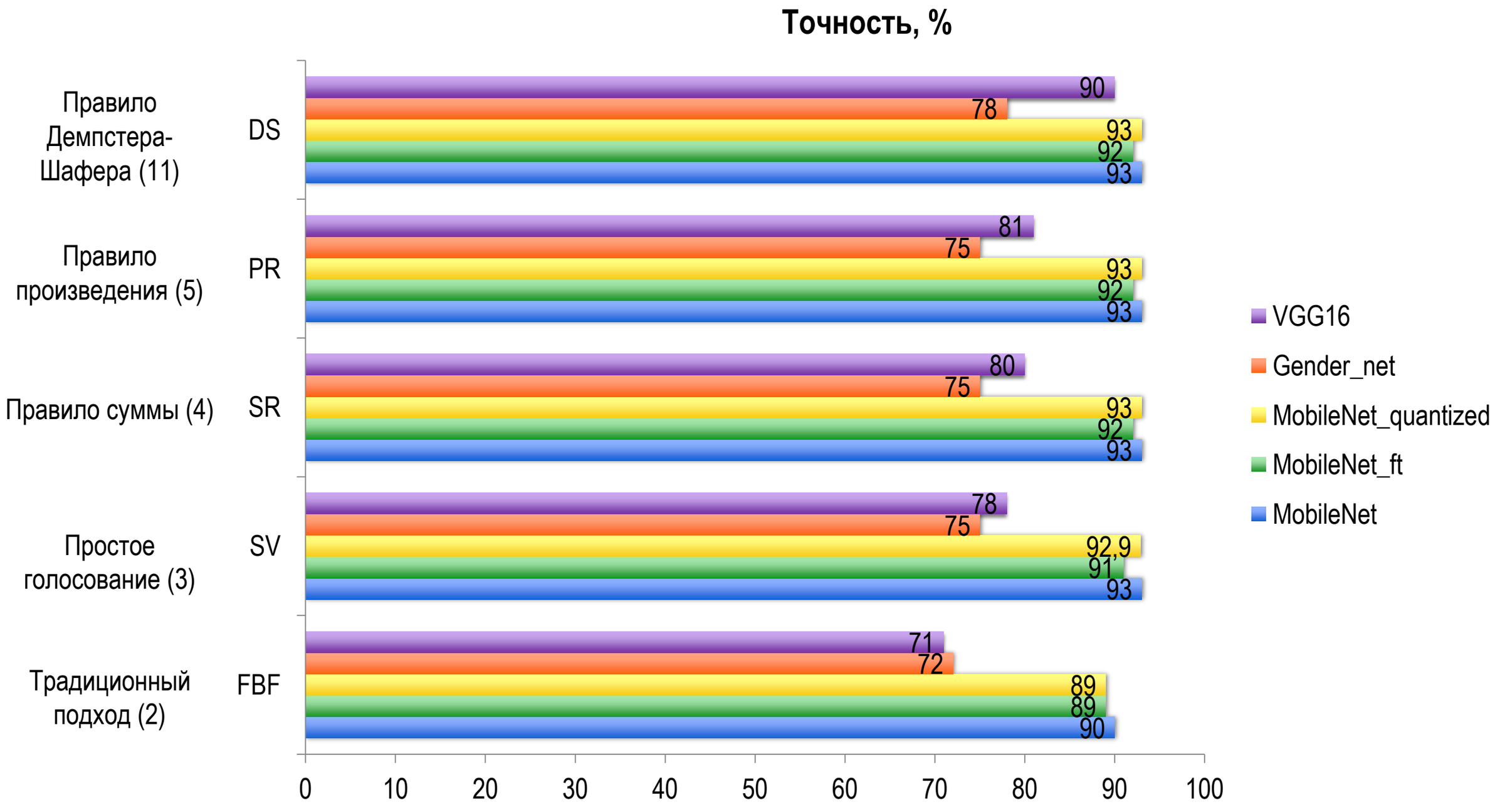
Результаты экспериментов. Распознавание пола

Datacem Kinect



Результаты экспериментов. Распознавание пола

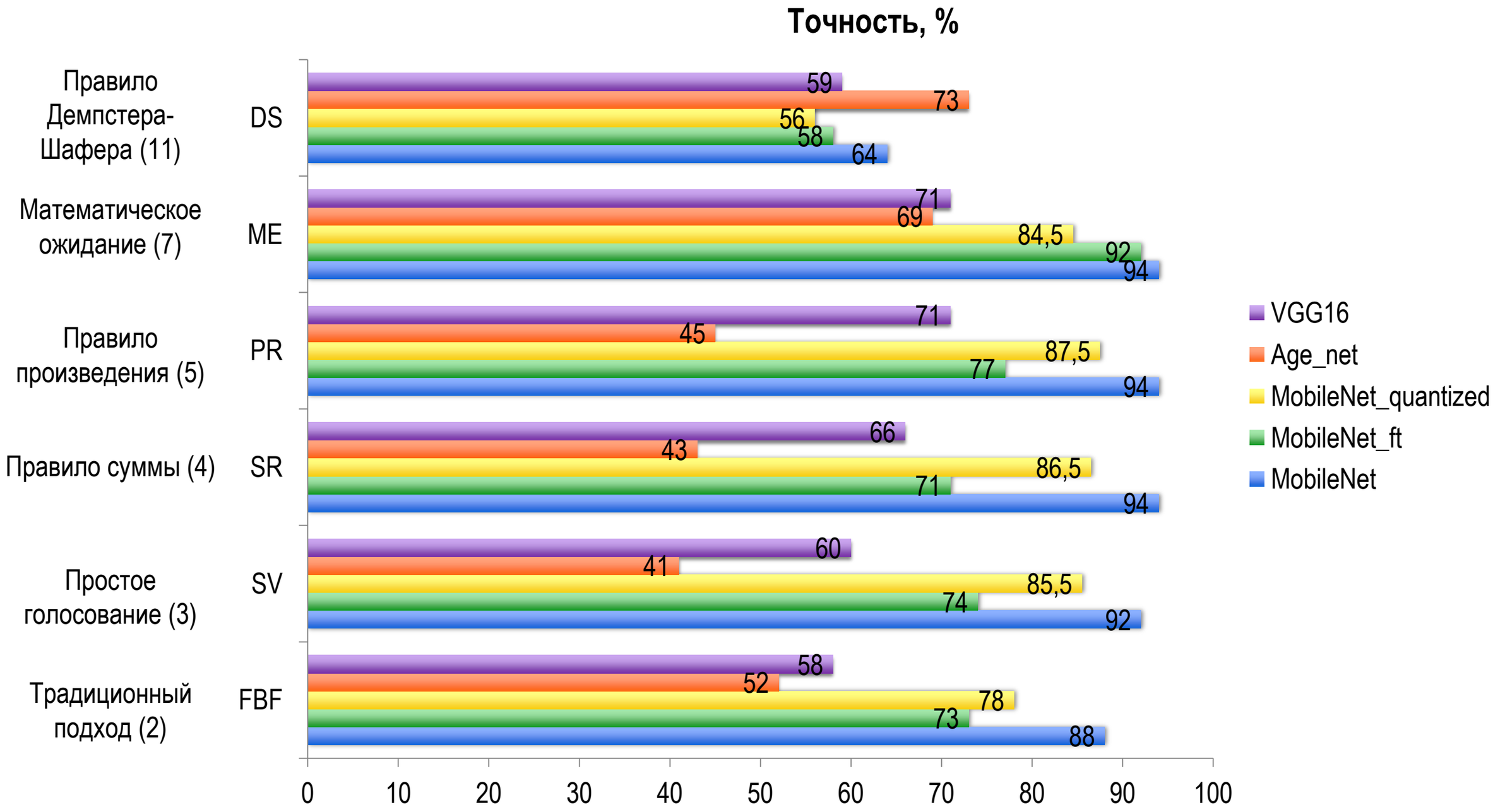
Датасет EmotiW2018



Результаты экспериментов. Распознавание возраста

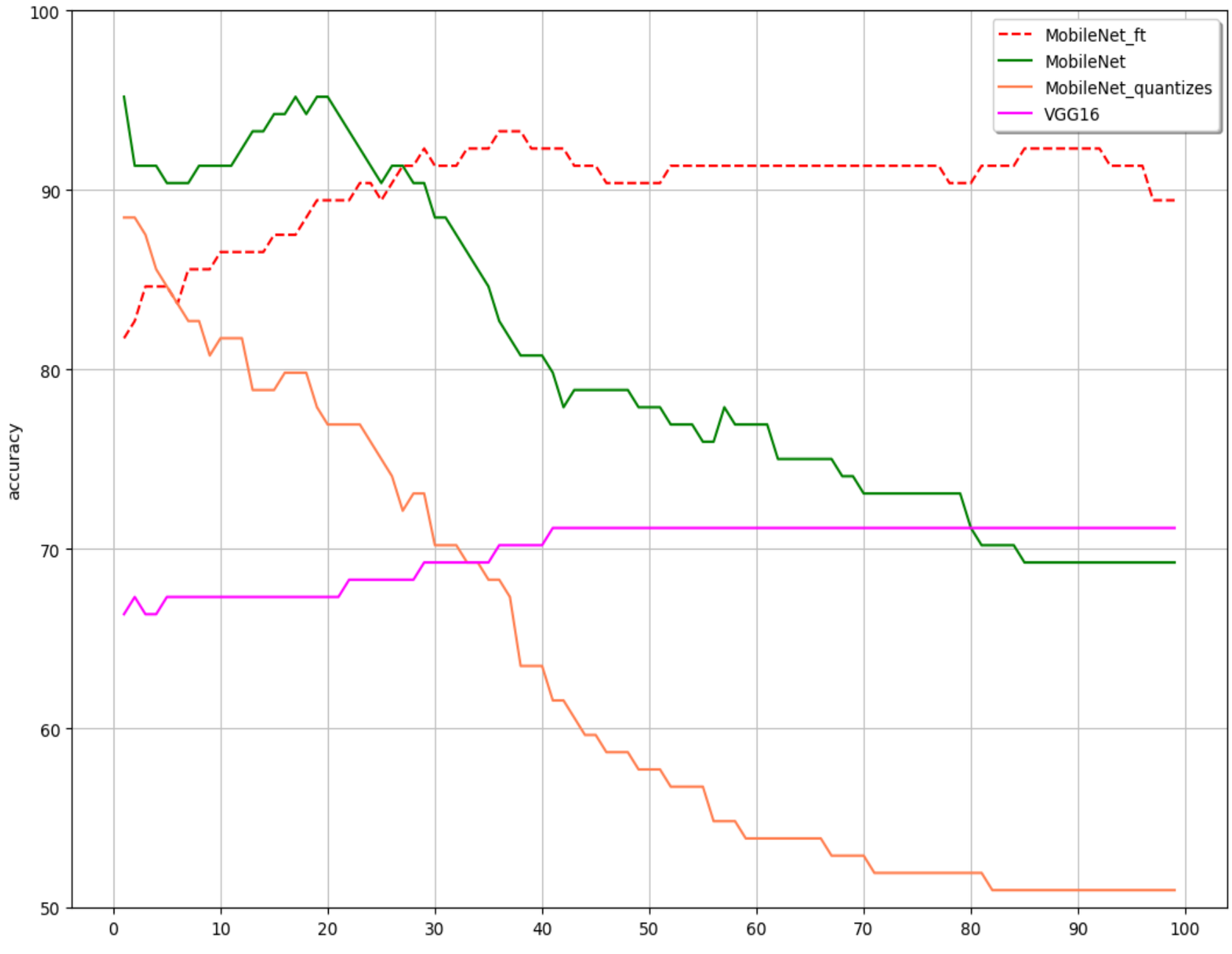
Datacem Kinect

Возраст в базе данных рассматривается в рамках интервала +- 5 лет



Результаты экспериментов. Распознавание возраста

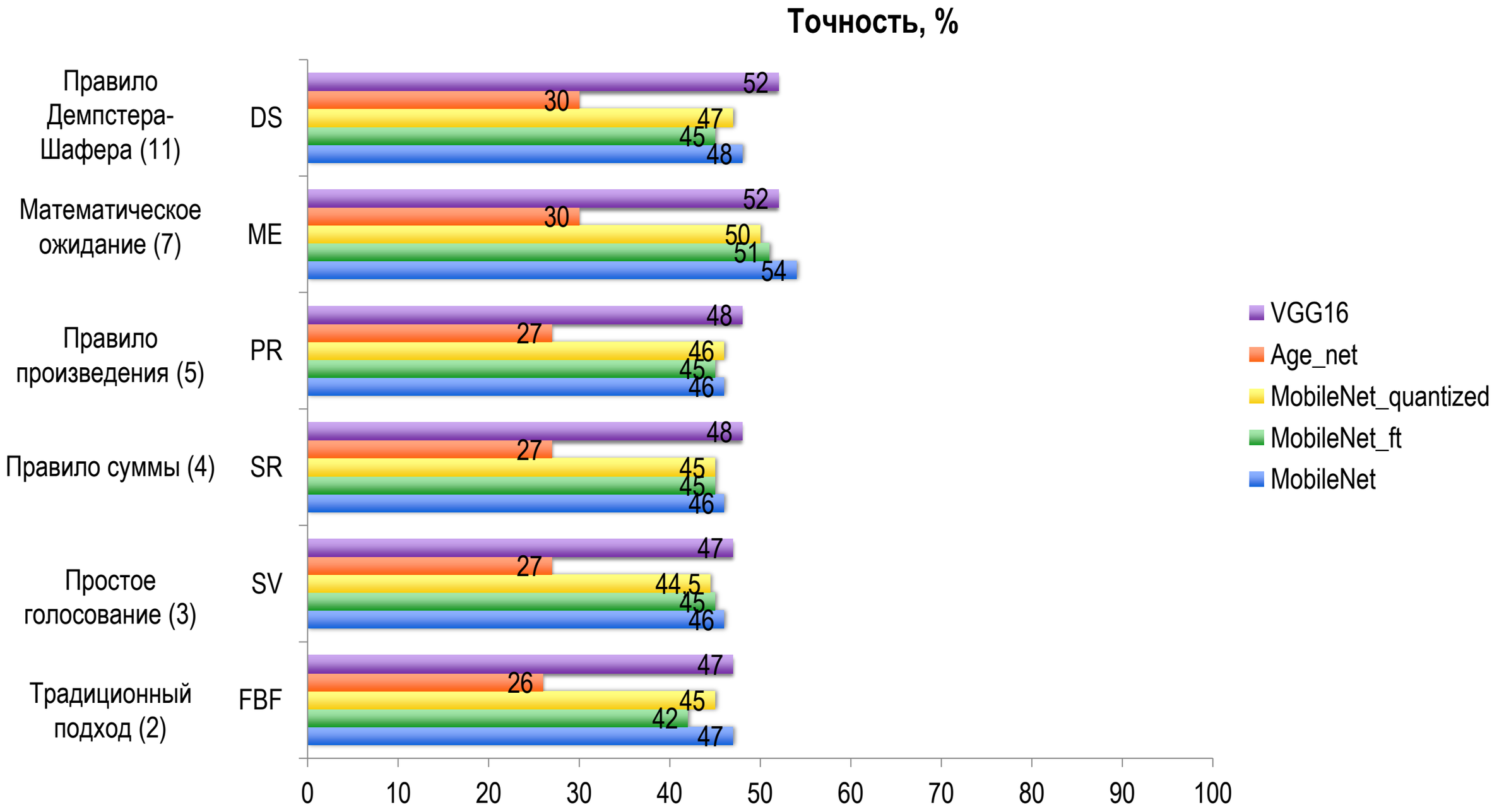
Зависимость точности математического ожидания от количества наиболее вероятных классов K (7)



Результаты экспериментов. Распознавание возраста

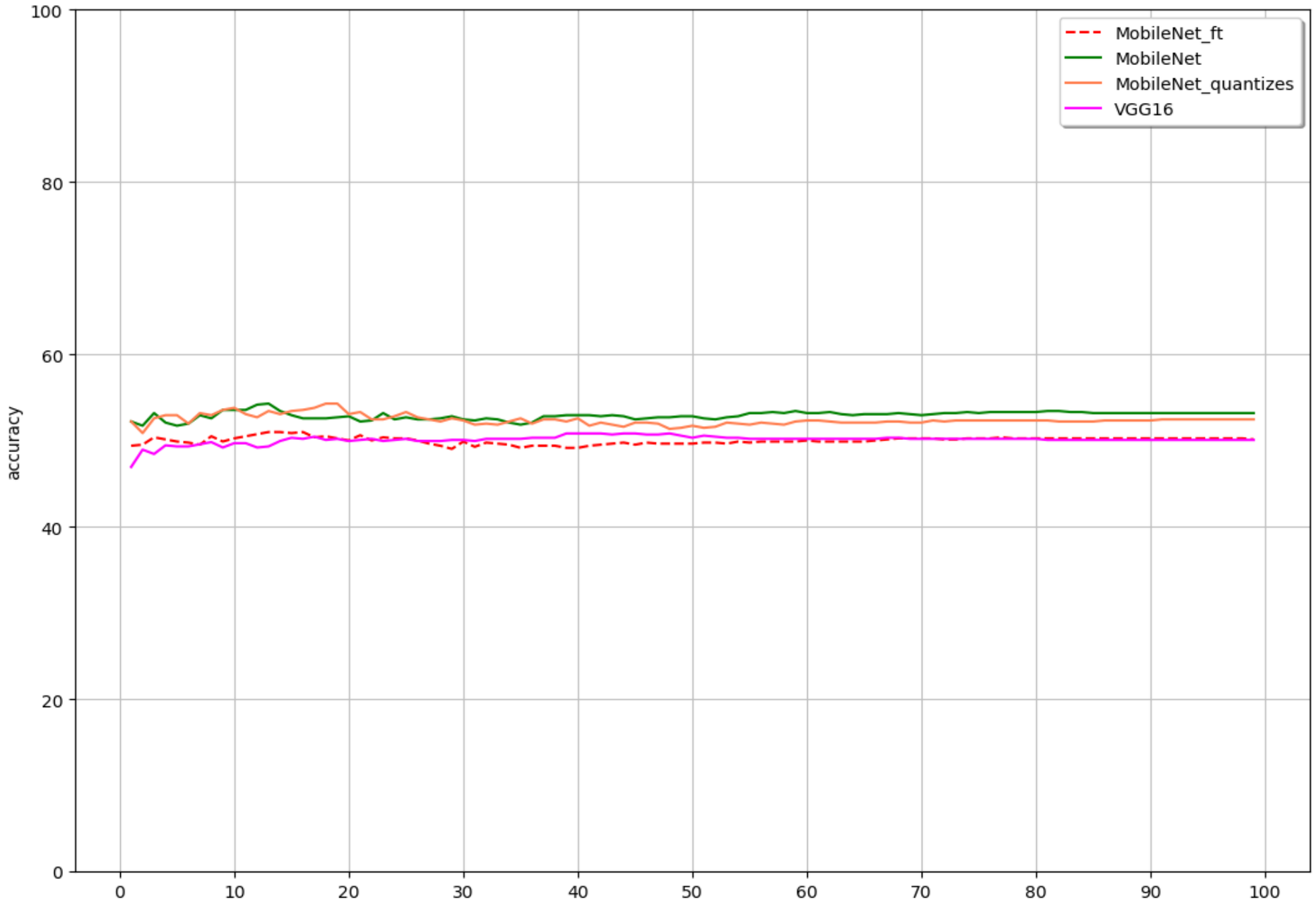
Datacem EmotiW2018

Возраст в базе данных рассматривается в рамках интервала +- 5 лет



Результаты экспериментов. Распознавание возраста

Зависимость точности математического ожидания от количества наиболее вероятных классов K (7)



Результаты

- Проведен аналитический обзор литературы по методам распознавания видеоизображений, с акцентированием внимания на архитектуры сверточных нейронных сетей

- Разработана схема алгоритма распознавания пола и возраста по видео. Предложенный алгоритм реализован в виде прототипа программной системы на мобильной платформе Android

- Проведены экспериментальные исследования для определения пола и возраста из наборов данных *Kinect, IJB-A, Indian Movie, EmotiW2018.* Показано, что наиболее эффективным методом агрегации решений для определения пола является алгоритм Демпстера-Шафера, который оказывается на 2-5% точнее по сравнению с известными аналогами. В то же время вычисление математического ожидания показало свою эффективность на 5-10% для проблемы распознавания возраста

Результаты

- Международная конференция «Информационные системы и технологии 2017»
- IV Международная конференции «Информационные технологии и нанотехнологии» (ИТНТ-2018)
- 2 место в конкурсе научно-исследовательских работ студентов (НИРС) НИУ ВШЭ в номинации «Лучшая научно-исследовательская работа по компьютерным наукам для студентов бакалавриата»
- 3 место на IV Междисциплинарной студенческой конференции "Интеллектуальный город: наука как искусство"

В будущем планируется произвести более подробное исследование возможности реализации сверточных нейронных сетей на мобильных устройствах в задачах определения пола и возраста, а также продолжить рассматривать методы агрегации, требующие обучение алгоритма

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

СПАСИБО ЗА ВНИМАНИЕ!