



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

19-я Всероссийская конференция «Математические методы распознавания образов»

Распознавание пола и возраста лица на видеоизображениях для мобильных платформ

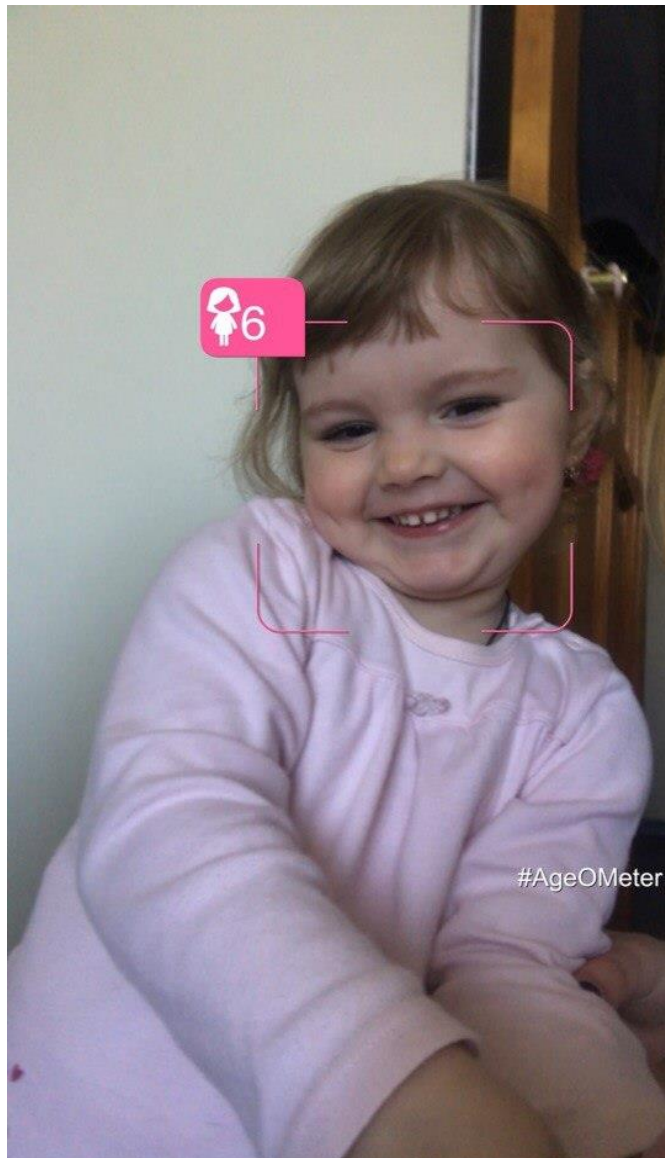
Харчевникова Ангелина Сергеевна,
д.т.н, Савченко Андрей Владимирович,
НИУ ВШЭ Нижний Новгород

Предметная область



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

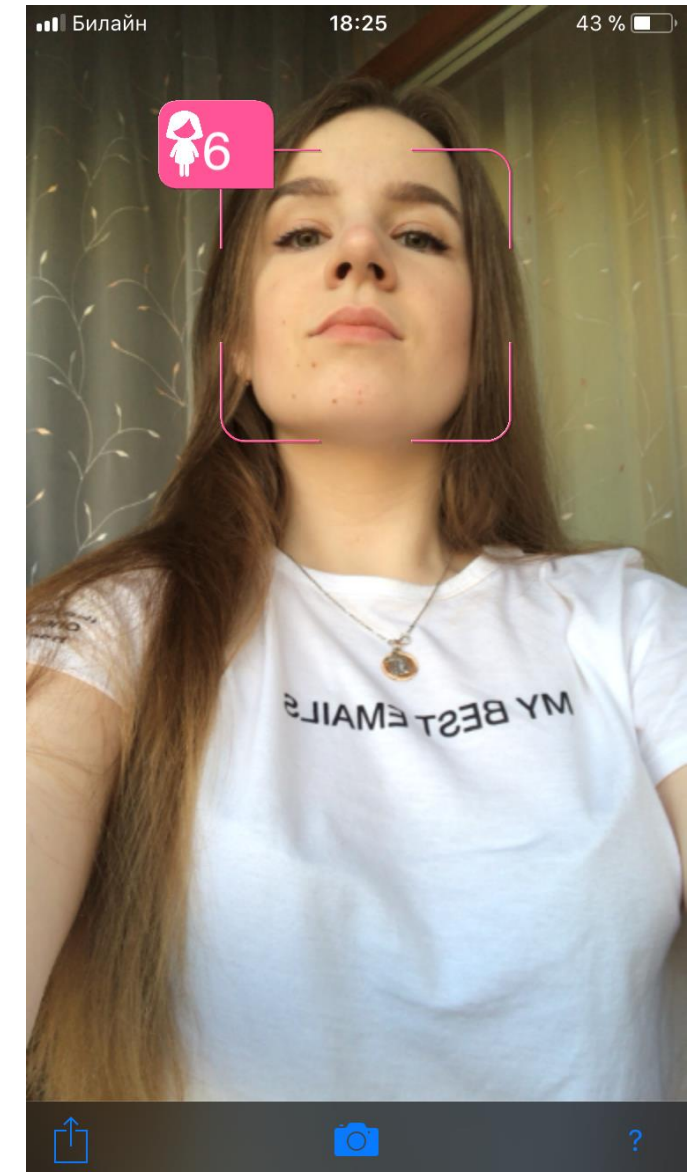
Определение пола и возраста



Luxand FaceSDK



Microsoft How-old.net

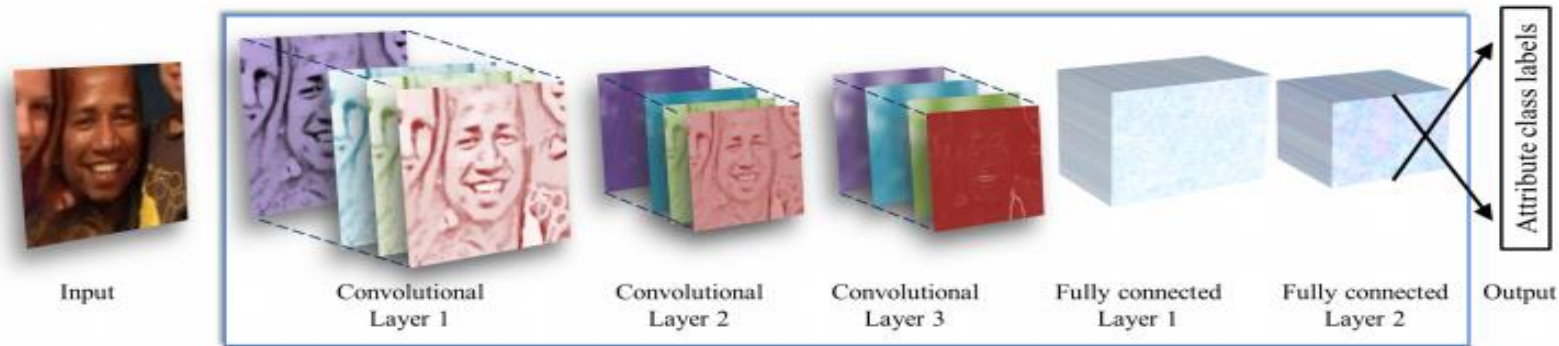


Luxand FaceSDK

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



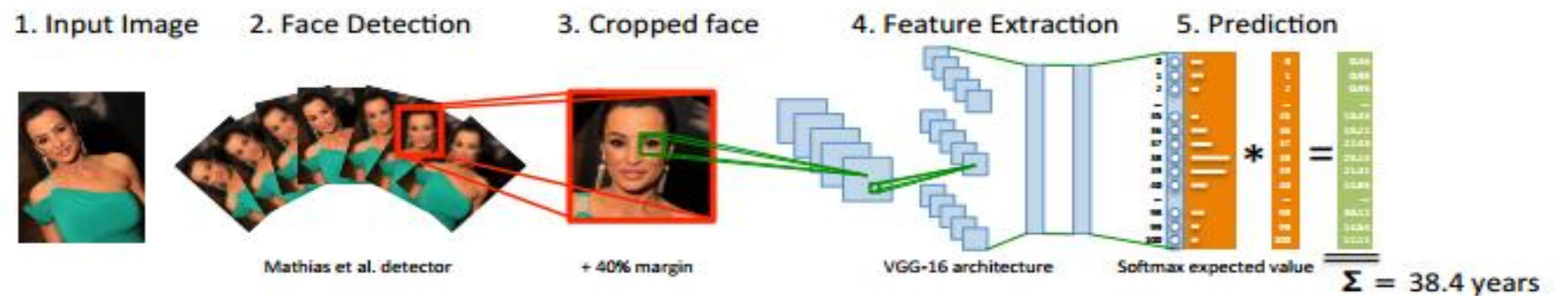
Сверточные нейронные сети



44 Mб

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

514 Mб



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



MobileNet

14 Mб

MobileNet-finetuned

13 Mб

MobileNet-quantized

4 Mб



Комитеты классификаторов

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

Выход нейросетевой модели
- слой Softmax() (1)

$$l^*(t) = \underset{l=1,2,\dots,L}{\operatorname{argmax}} P(l|X(t))$$

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

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

Простое голосование (3)

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

Правило суммы (4)

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

Правило произведения (5)



Комитеты классификаторов. Математическое ожидание

$$l^* = \sum_{l=1}^L P(l|X(t)) \cdot l \quad \text{Оценка математического ожидания} \quad (6)$$

Предложенный подход

$$l^* = \frac{\sum_{k=1}^K P^*(l_k|X(t)) \cdot l_k}{\sum_{i=1}^K P^*(l_i|X(t))} \quad \text{Оценка математического ожидания для } K \text{ наиболее вероятных классов} \quad (7)$$
$$P^*(l_1|X(t)) \geq P^*(l|X(t)) \geq P^*(l_L|X(t))$$



Комитеты классификаторов. Метод Демпстера-Шафера

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

Построение шаблонов решения с помощью базы данных изображений (8)

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

Вычисление близости между DT_l^i и решением для i -го фрейма (9)

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

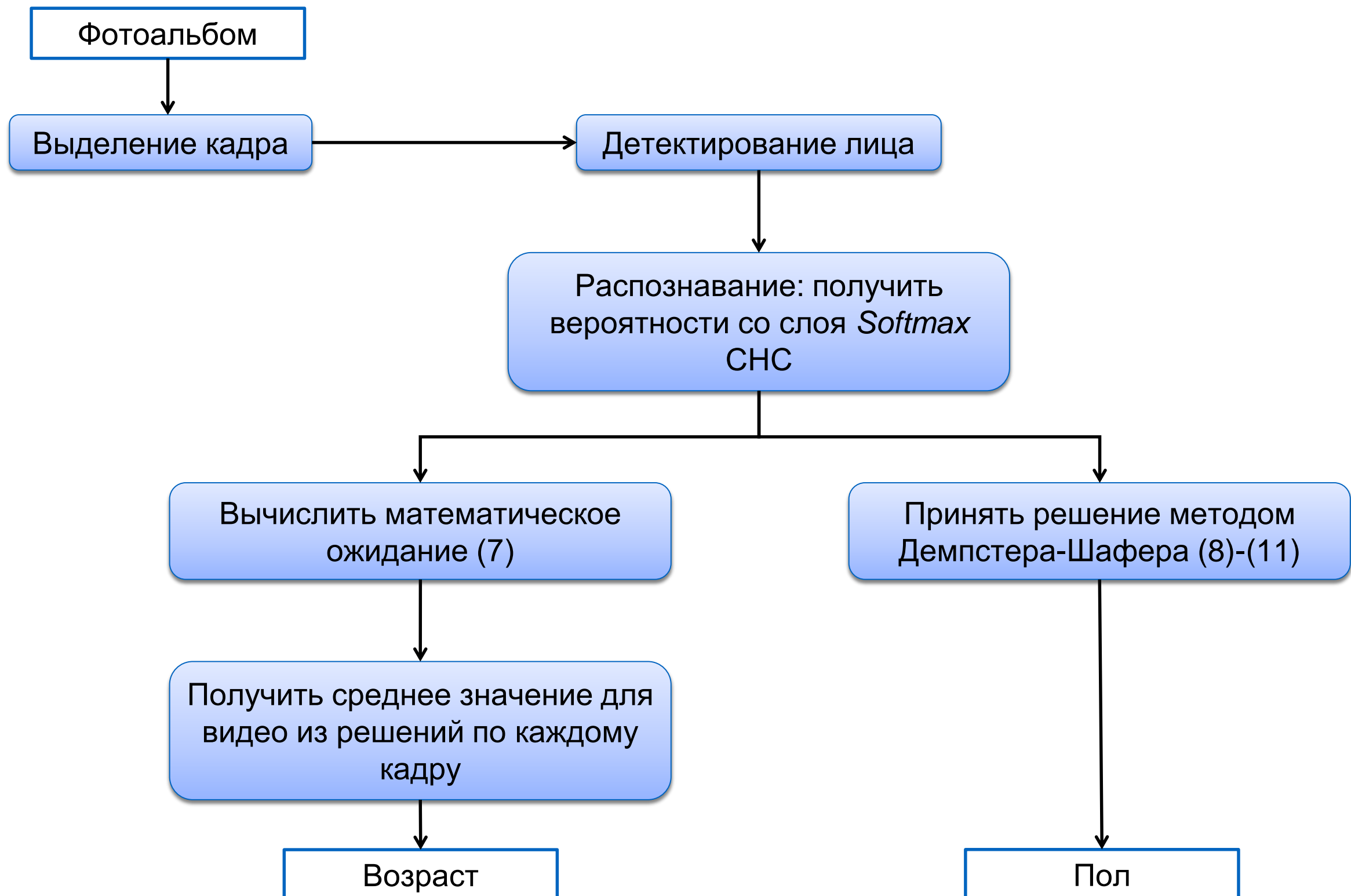
Вычисление степени доверия (10)

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

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

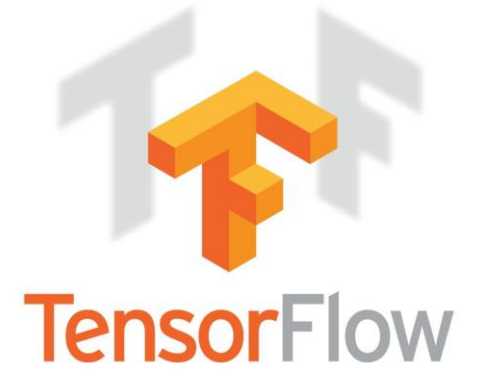
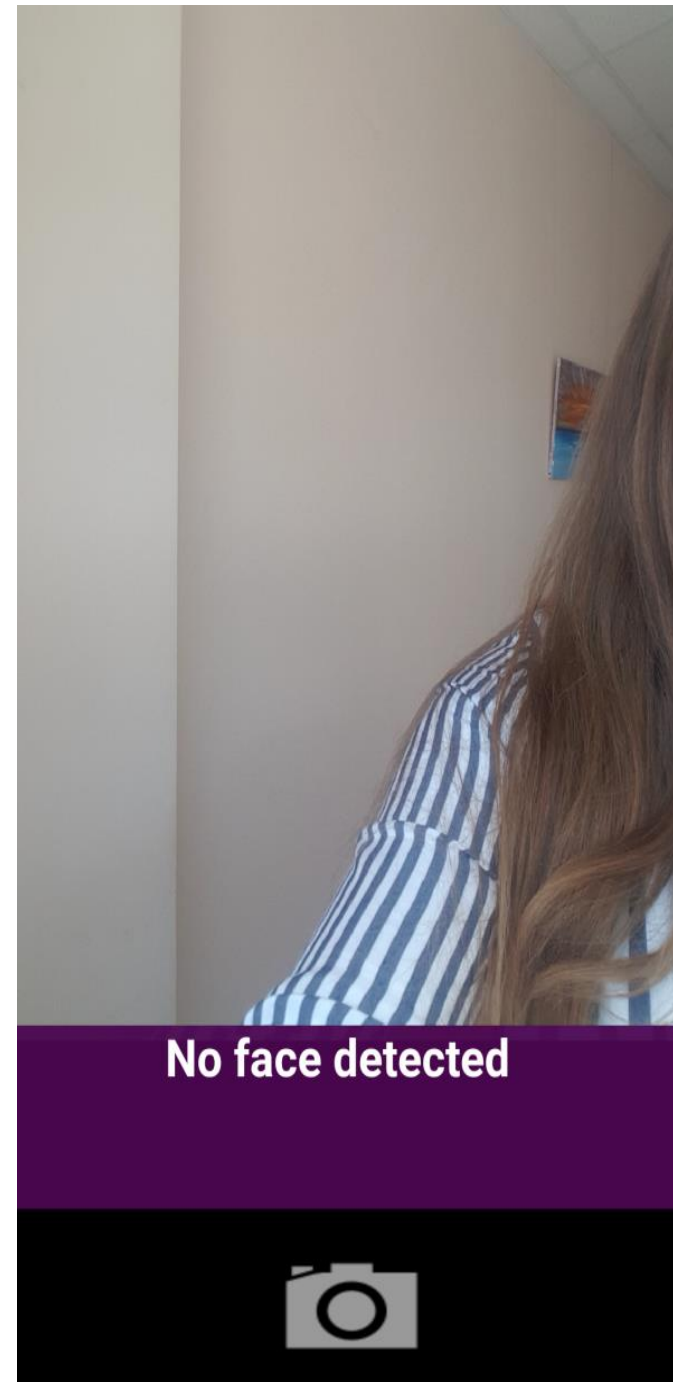


Структурная схема системы





Реализация системы на мобильной платформе



ANDROID

Результаты экспериментов. Время распознавания

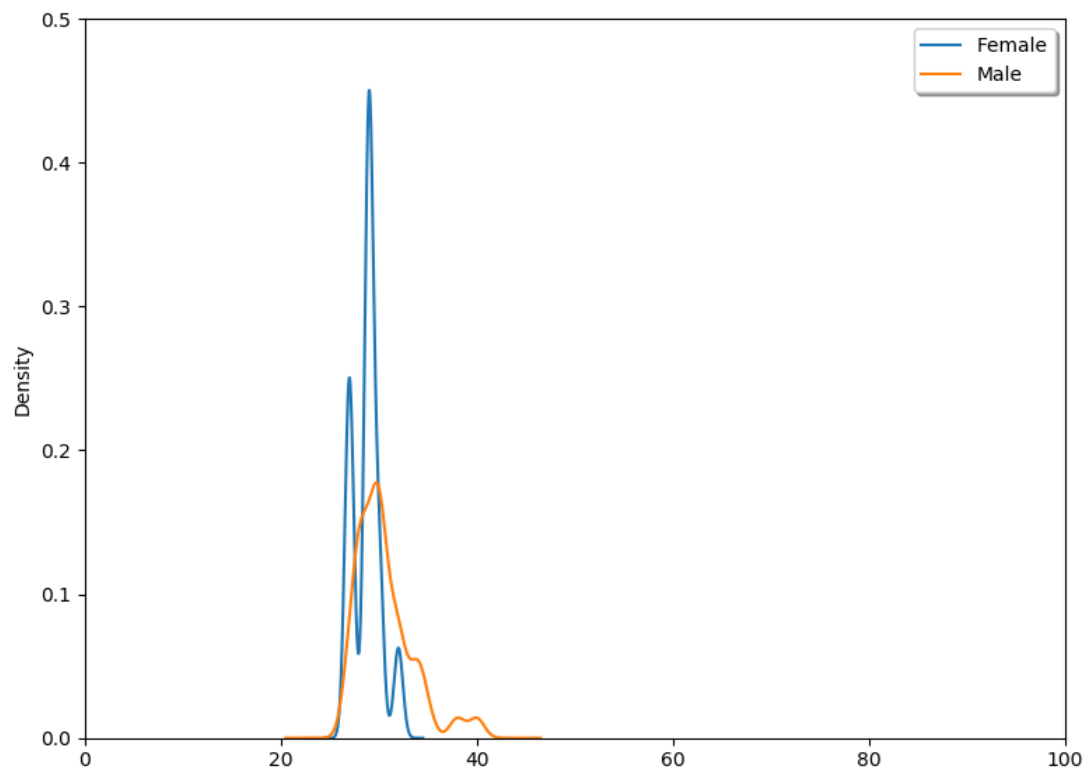
Модель СНС	Время, с
MobileNet	0.371
MobileNet_ft	0.368
MobileNet_quantized	0.373
Gender/Age net	1.076
VGG16	273



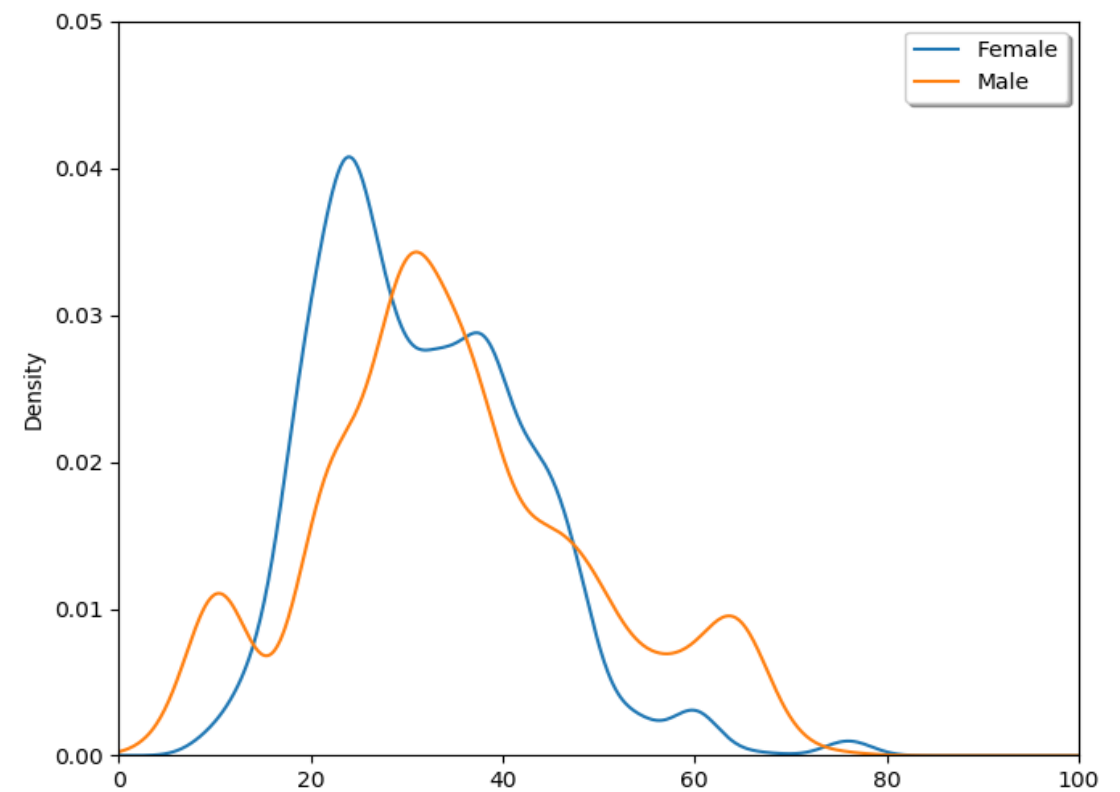


Результаты экспериментов. Наборы данных

- IJB-A
- Indian Movie
- Kinect
- EmotiW2018



Kinect

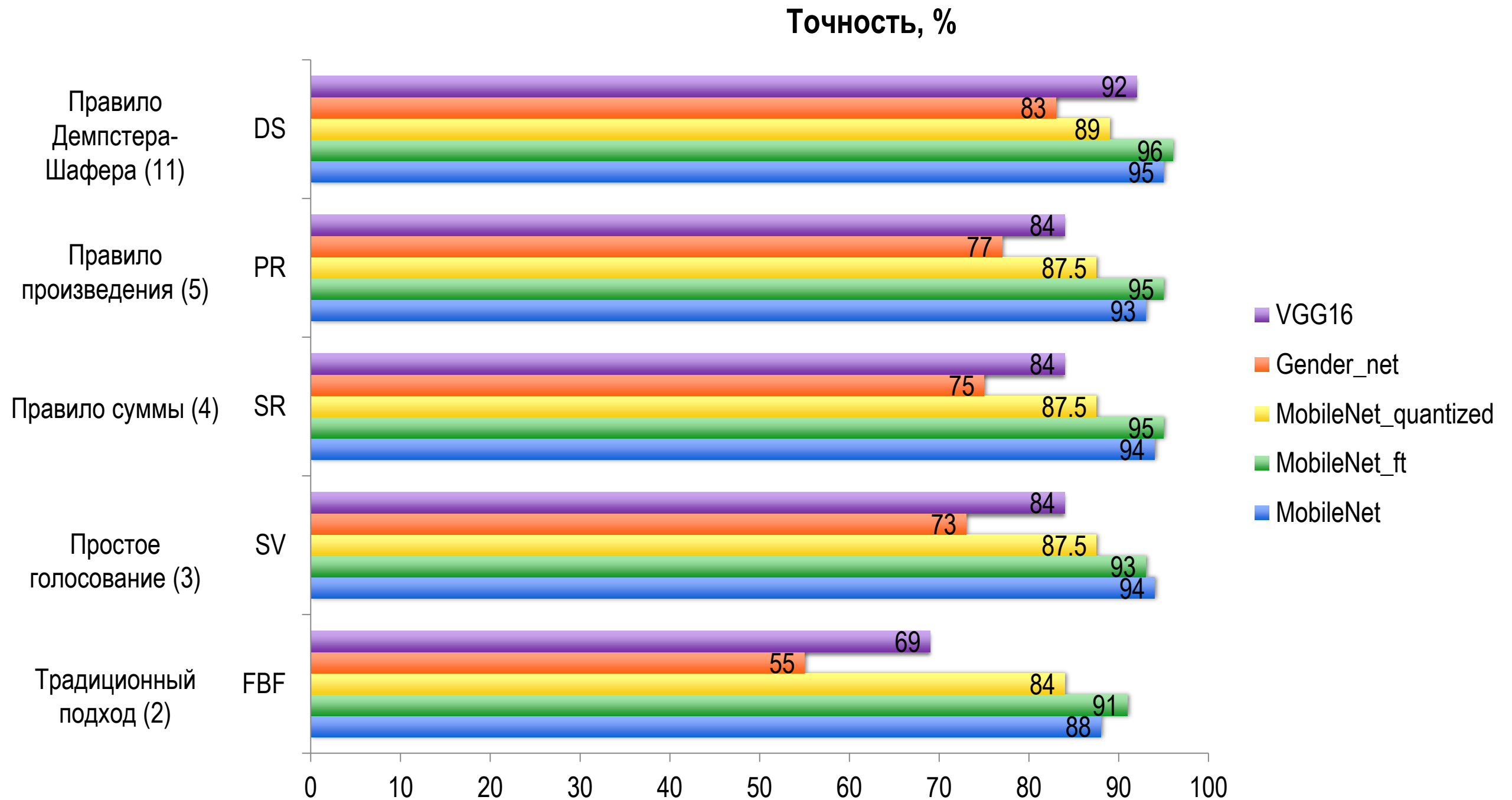


EmotiW2018



Результаты экспериментов. Распознавание пола

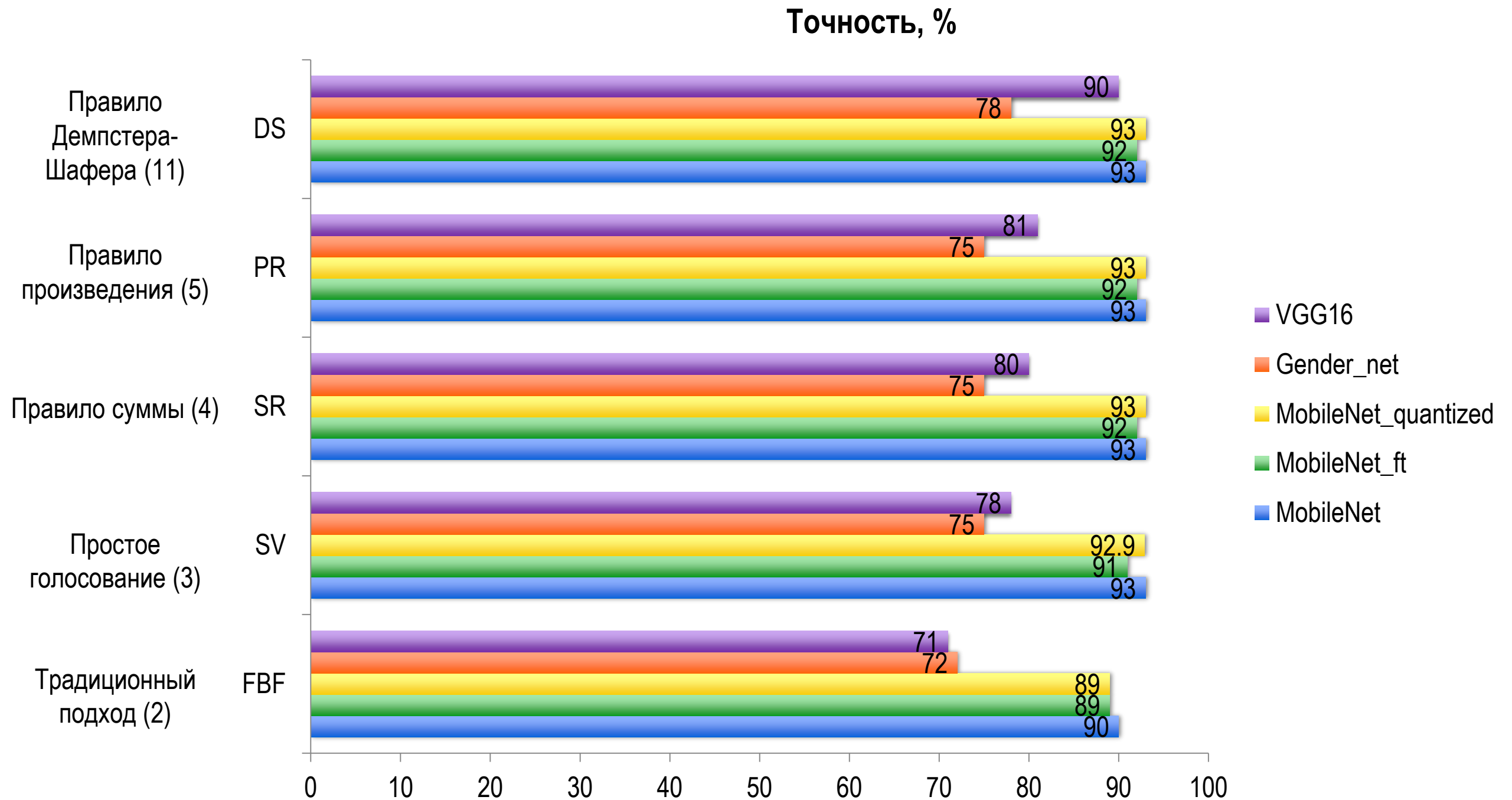
Датасет Kinect





Результаты экспериментов. Распознавание пола

Damascus EmotiW2018

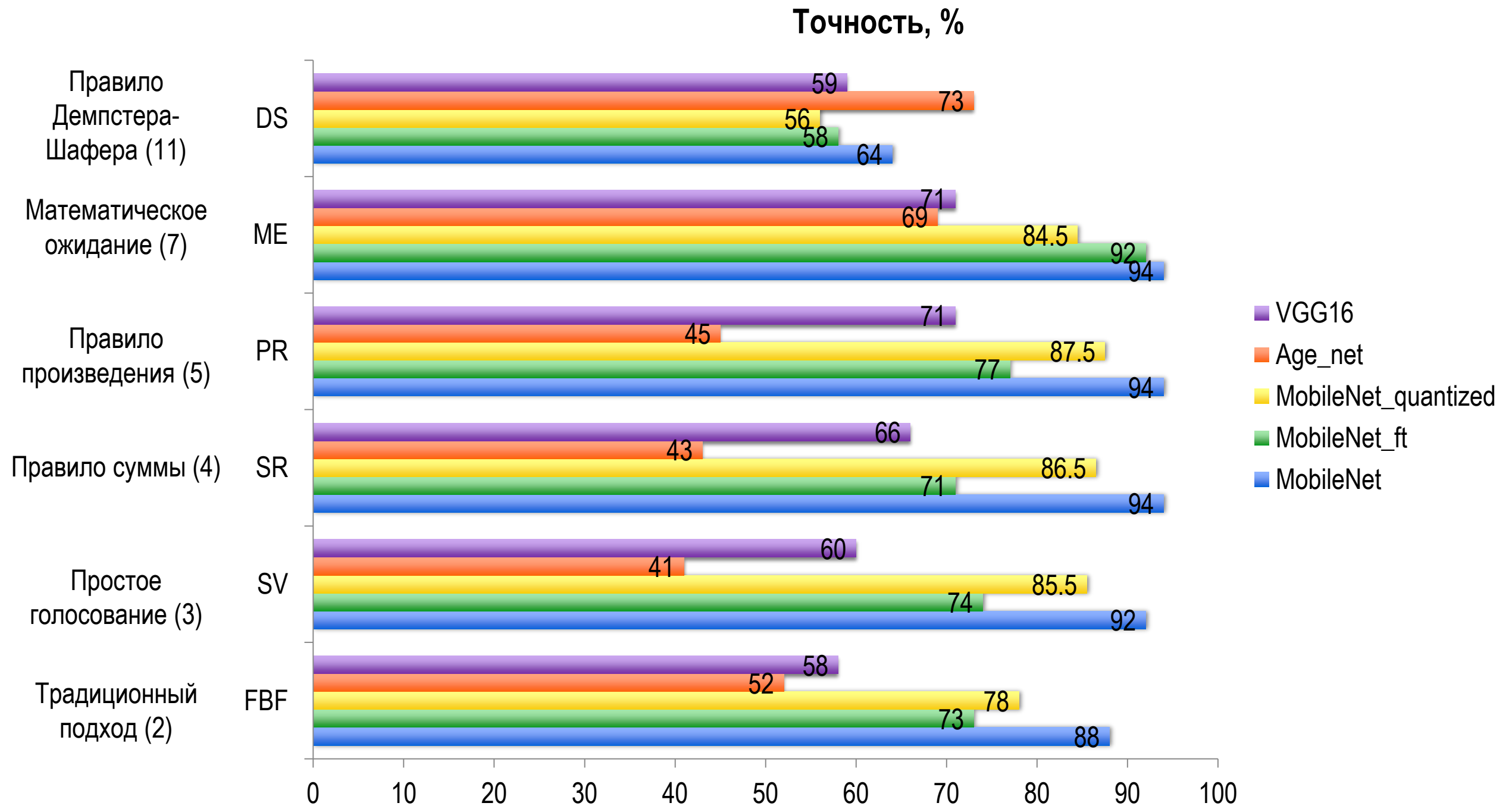




Результаты экспериментов. Распознавание возраста

Датасет Kinect

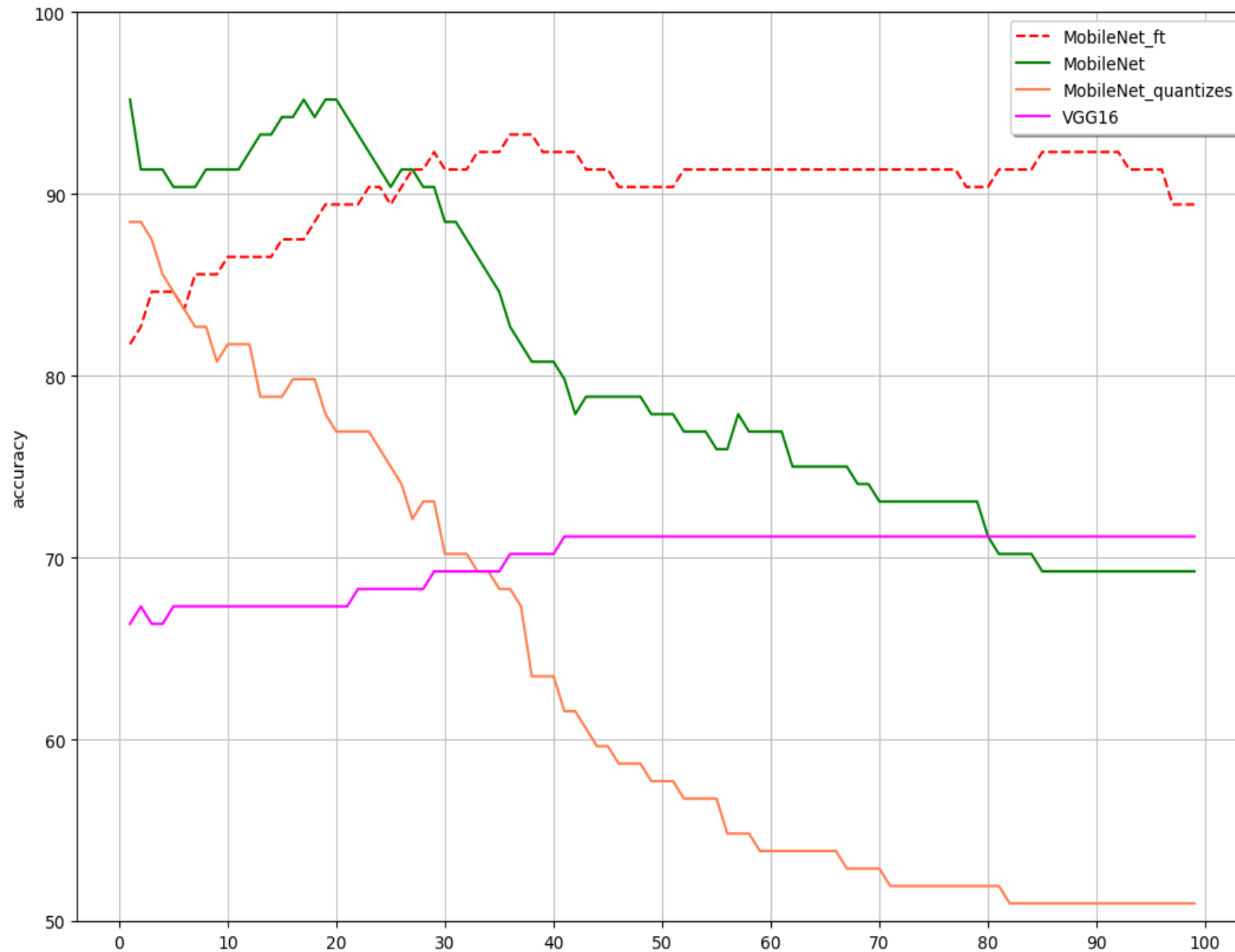
Возраст в базе данных рассматривается в рамках интервала ± 5 лет





Результаты экспериментов. Распознавание возраста

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

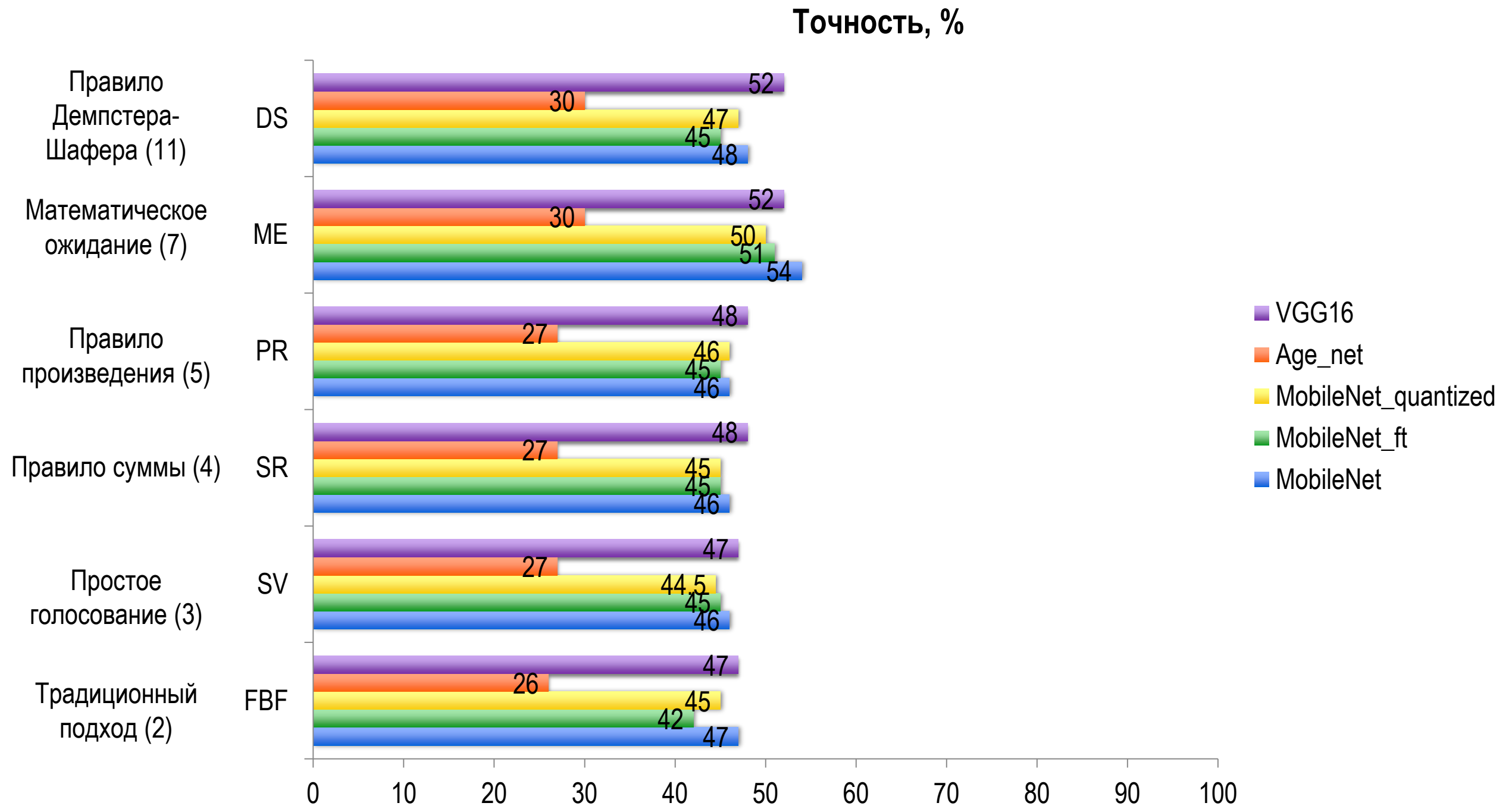




Результаты экспериментов. Распознавание возраста

Damascus EmotiW2018

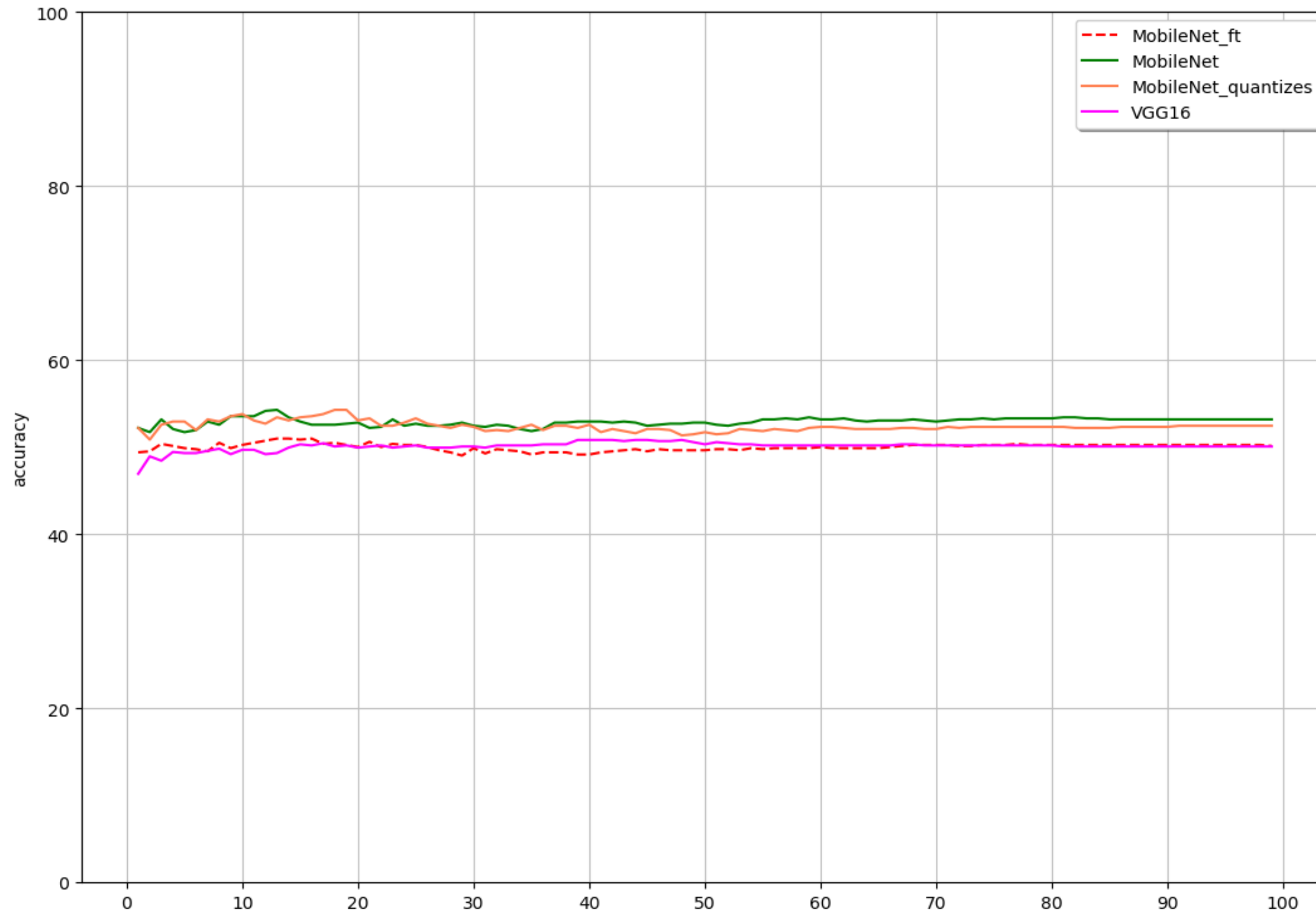
Возраст в базе данных рассматривается в рамках интервала ± 5 лет





Результаты экспериментов. Распознавание возраста

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



Результаты

- 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



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

СПАСИБО ЗА ВНИМАНИЕ!