NATIONAL RESEARCH
UNIVERSITY

National Research University
Higher School of Economics

Faculty of Informatics, Mathematics, and Computer Science
Department of Information Systems and Technology

VIDEO-BASED AGE AND GENDER CLASSIFICATION WITH CONVOLUTIONAL NEURAL NETWORKS

Kharchevnikova Angelina
14BI-1

Nizhny Novgorod
2018

AGENDA

- Problem statement
- Literature survey
- Methodology
- Experimental results and discussion
- Concluding comments and future plans

PROBLEM STATEMENT

Age and gender characteristics can be applied in retail for contextual advertising for particular group of customers

The reliability of the existing solutions remains insufficient for practical application

PROBLEM STATEMENT

The Purpose

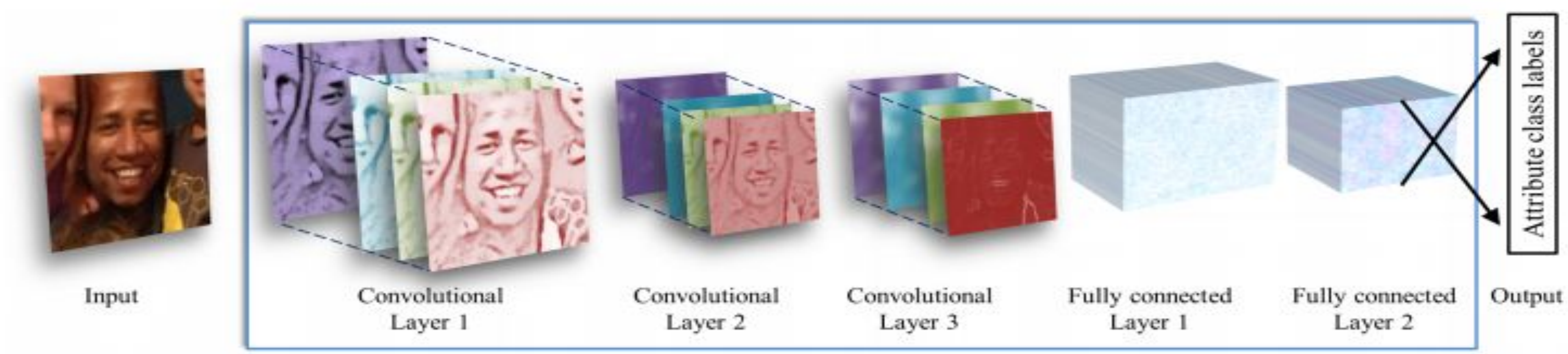
Developing a mobile off-line application for the video-based age and gender recognition using convolutional neural networks with classifier fusion methods

The Goal Set

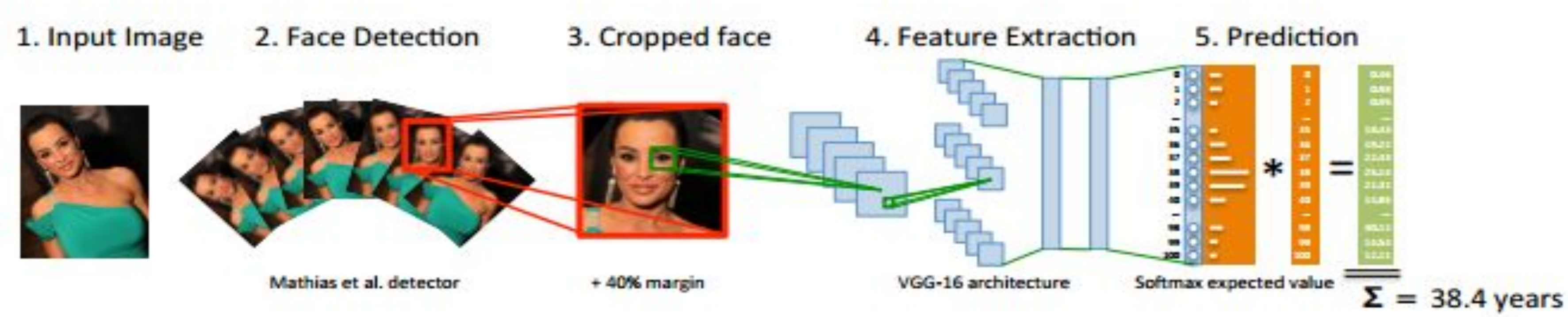
- To conduct a review of the existing solutions for image recognition by age and gender with an emphasis on convolutional neural networks (CNNs)
- To describe the algorithm scheme of the proposed system
- To develop the age/gender recognition architecture, taking into account the classifier fusion methods
- To analyze the accuracy of decision making on the basis of each aggregation method by conducting experiments
- To implement the algorithm scheme on mobile platform

LITERATURE SURVEY

CNNs architectures



Gender_net and Age_net (44MB) (Levi, G. Age and gender classification using convolutional neural networks 2015)



VGG-16 (500MB) (Rasmus Rothe, Radu Timofte, Luc Van Gool DEX: Deep EXpectation of apparent age from a single image 2015)

METHODOLOGY

CLASSIFIER FUSION METHODS

Simple voting

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

Arithmetical mean (sum rule)

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

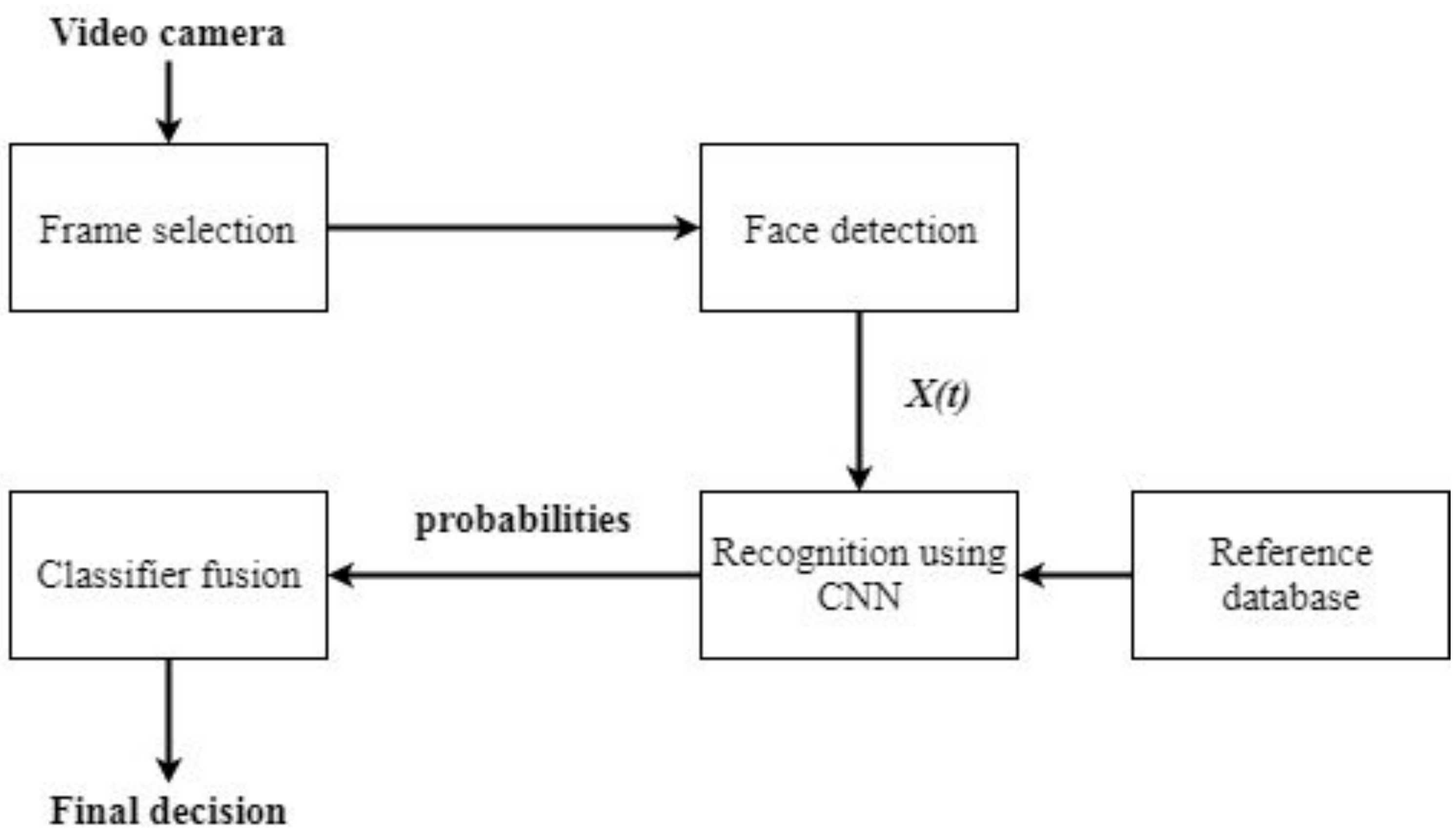
Geometric mean (product rule)

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

Expected value (mathematical expectation)

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

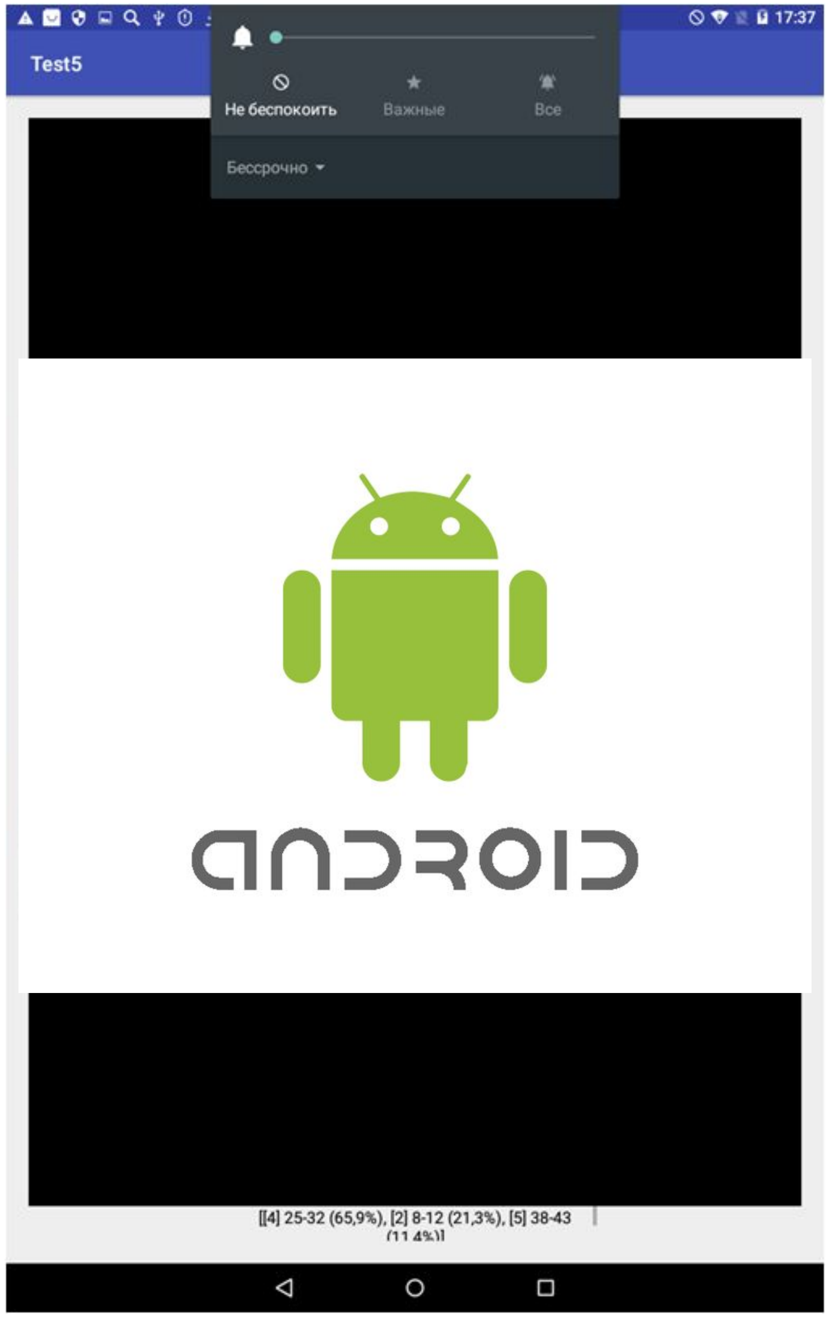
METHODOLOGY



The output of the CNN - the Softmax layer

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

EXPERIMENTAL RESULTS



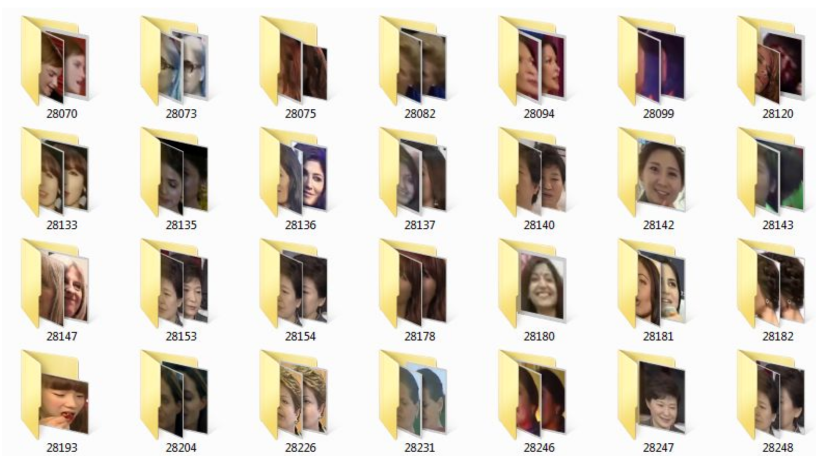
TensorFlow

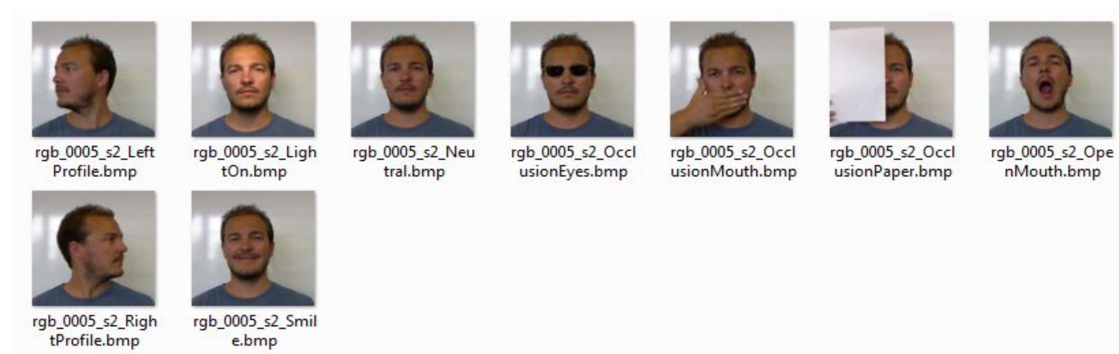
GoogleTensorFlow for recognition

EXPERIMENTAL RESULTS

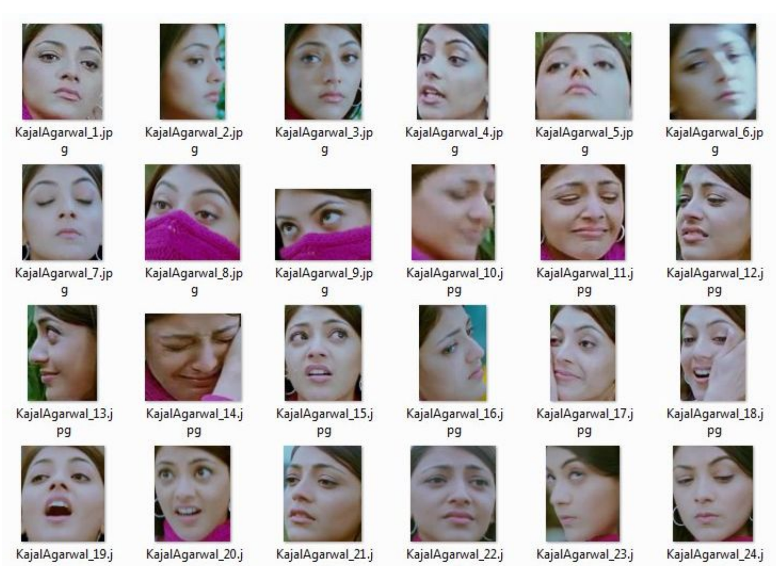
Datasets

IARPA Janus Benchmark A (IJB-A): 2043 video, 13900 frames + gender information





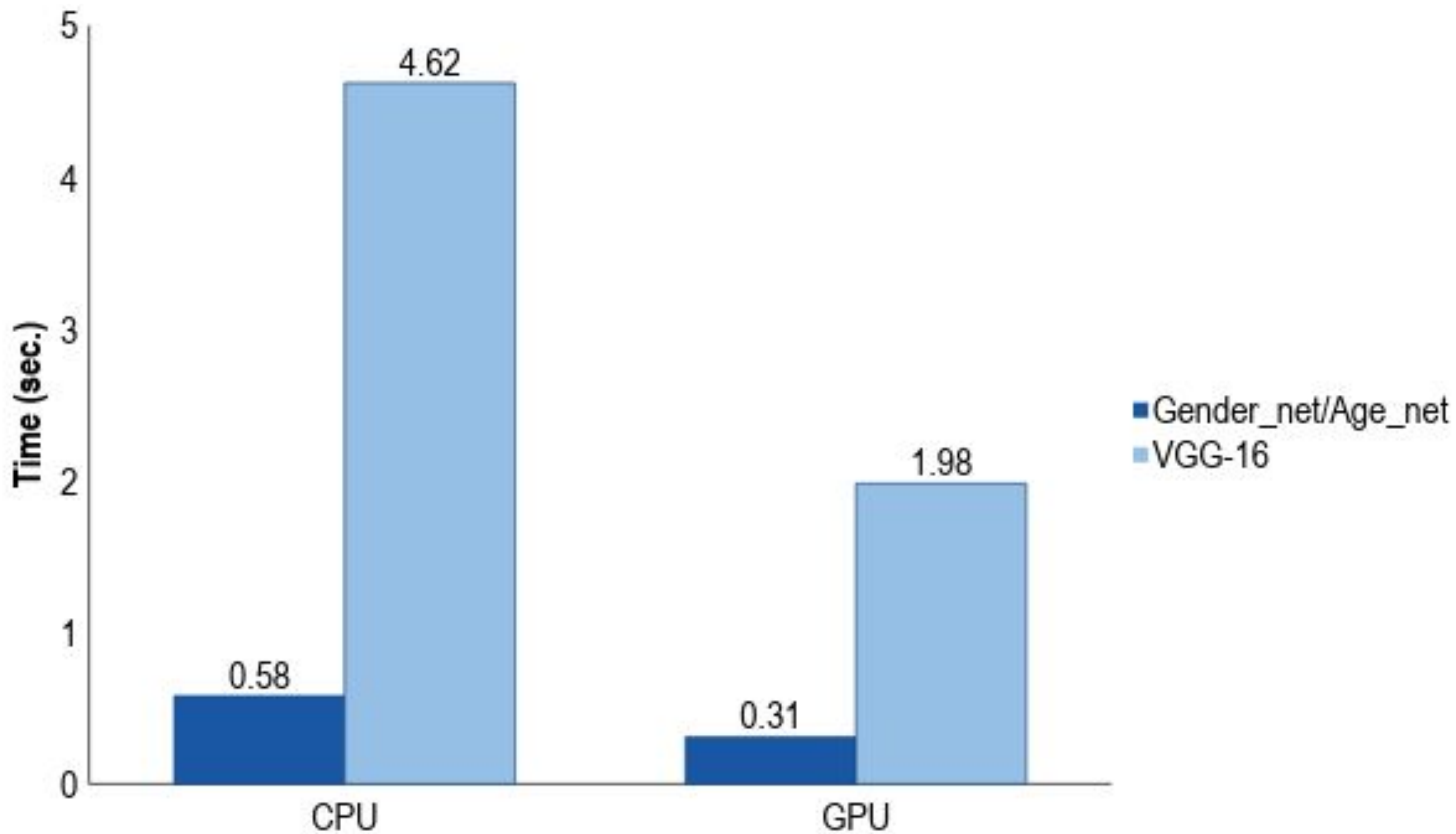
Kinect: 104 video, 936 frames + gender and age information



Indian Movie: 332 video, 28312 frames + gender and age information

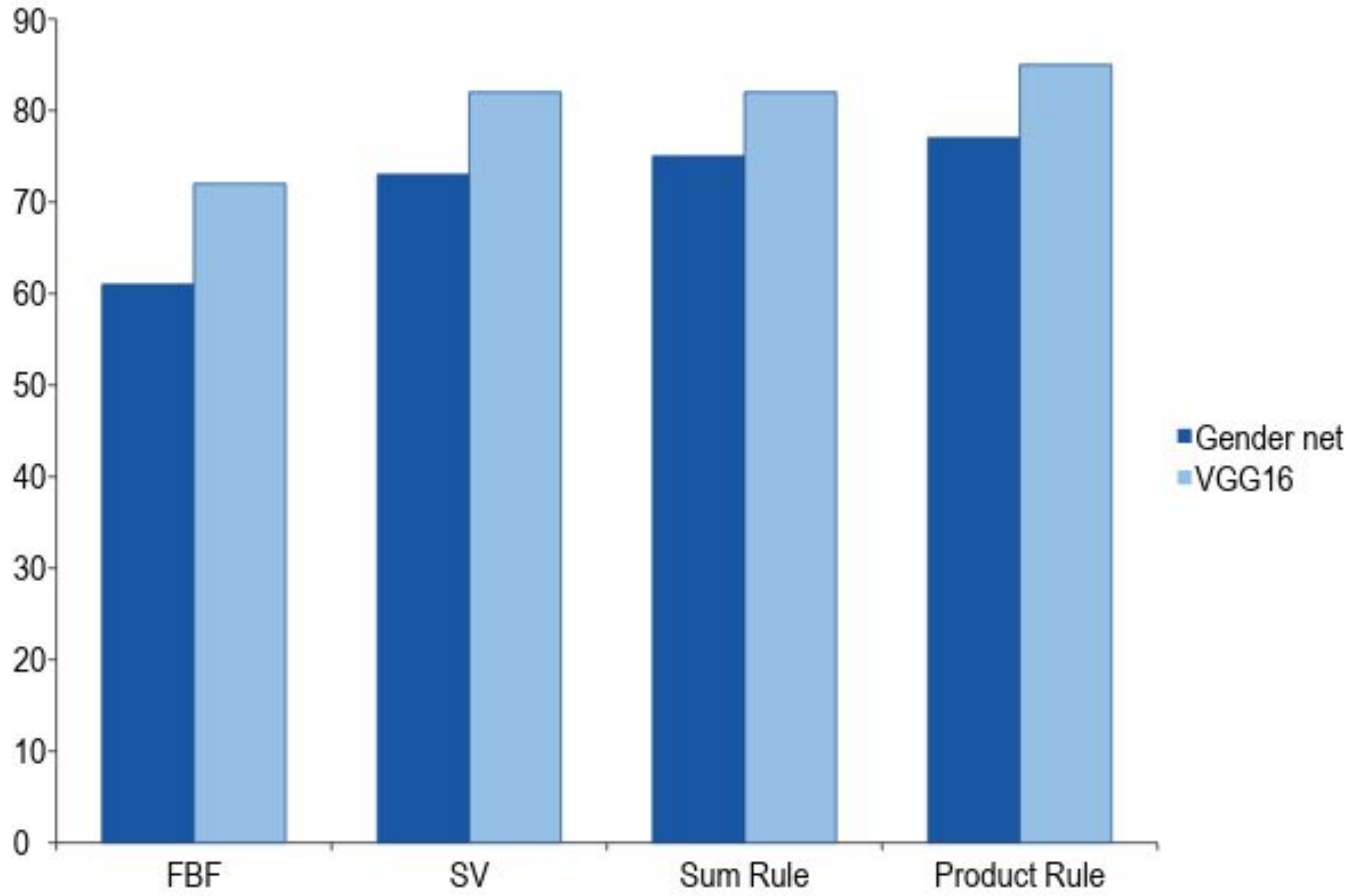
EXPERIMENTAL RESULTS

Inference time



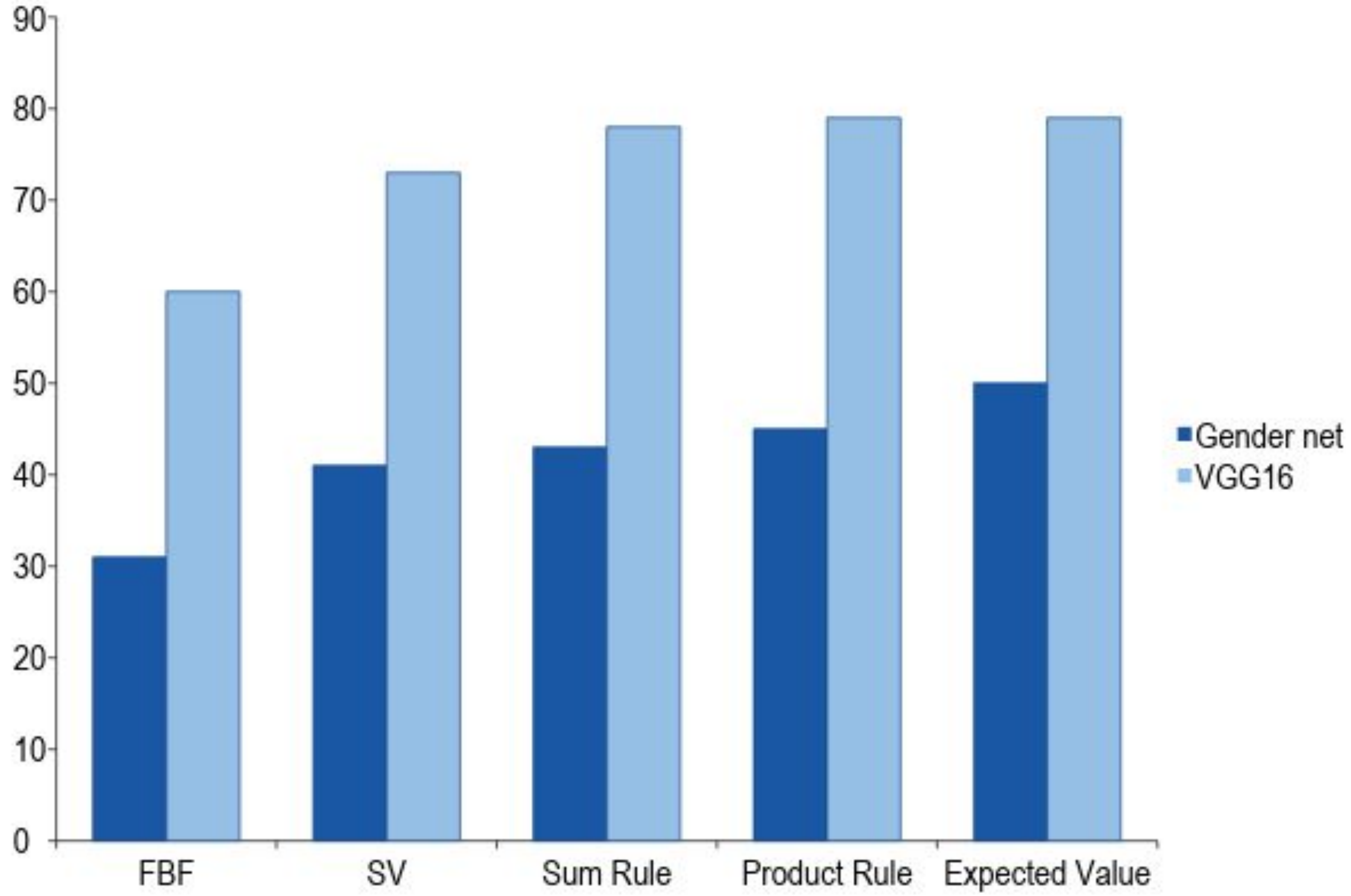
EXPERIMENTAL RESULTS

Kinect dataset
Gender recognition



EXPERIMENTAL RESULTS

Kinect dataset
Age recognition



Concluding comments and future plans

Conclusion

- The geometric mean (product rule) with normalization of the input video images is the most accurate in gender classification task
- The most accurate age prediction is achieved with the computation of the expected value
- The accuracy of the VGG-16 architecture is about 15% and 20% higher for the gender recognition and age prediction than Age and Gender net models
- The inference time of the VGG-16 is 4-9 times lower

Future work

- Completion of implementation of Android mobile off-line application
- Applying the modern techniques for fast classification and optimization of deep CNNs

References

1. Geng, X., Zhou, Z. H., Zhang, Y., Li, G., & Dai, H. (2006). Learning from facial aging patterns for automatic age estimation. In *Proceedings of the 14th ACM international conference on Multimedia* (pp. 307-316). ACM.
2. Golomb, B. A., Lawrence, D. T., & Sejnowski, T. J. (1990). A neural network identifies sex from human faces. In *NIPS* 572-579
3. Guo, G., Mu, G., Fu, Y., & Huang, T. S. (2009). Human age estimation using bio-inspired features. In *Computer Vision and Pattern Recognition, CVPR 2009. IEEE Conference on* (pp. 112-119). IEEE
4. Klare, B. F., Klein, B., Taborsky, E., Blanton, A., Cheney, J., Allen, K., & Jain, A. K. (2015). Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1931-1939.
5. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097-1105.
6. Kwon, Y. H., & Vitoria Lobo, N. (1999). Age classification from facial images. *Computer vision and image understanding*, *74*(1), 1-21.
7. Levi, G., Hassner T. (2015) Age and gender classification using convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 34-42
8. Makinen, E., & Raisamo, R. (2008). Evaluation of gender classification methods with automatically detected and aligned faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(3), 541-547.
9. Mariufic, M. P. (2016). Automatic determination of a person's gender in a photo using OpenCV. *Almanac of the world science,* (1-1), 75-77.
10. Min, R., Kose, N., & Dugelay, J. L. (2014). Kinectfacedb: A kinect database for face recognition. IEEE *Transactions on Systems, Man, and Cybernetics: Systems,*44(11), 1534-1548.
11. Nosov, G. Yu. (2014). Estimating age based on the video sequence. *Problems of modern science and education*, (12), 42-43.

THANK YOU FOR ATTENTION!

HIGHER · SCHOOL · OF · ECONOMICS ·

NATIONAL RESEARCH
UNIVERSITY