Robust Face Analysis Employing Machine Learning Techniques for Remote Heart Rate Estimation and towards Unbiased Attribute Analysis

By

Abhijit Das

STARS, Inria Sophia Antipolis – Méditerranée

30th January 2019

Contents

- Brief overview of my research
- Heart rate estimation from face videos
- Bias in face analysis

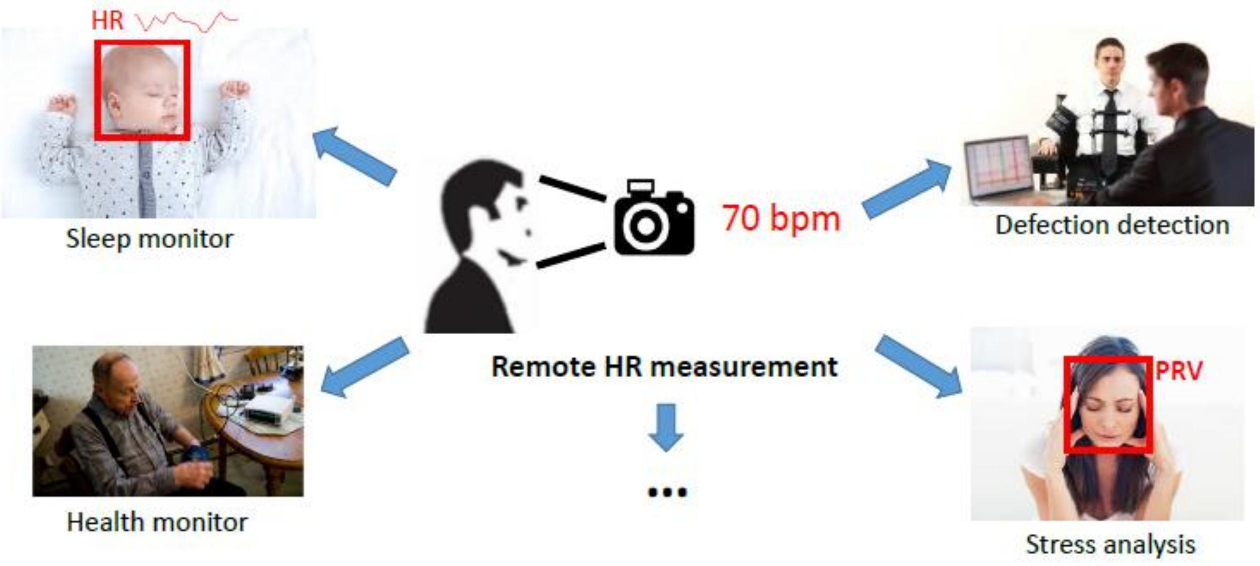
Brief overview of my research

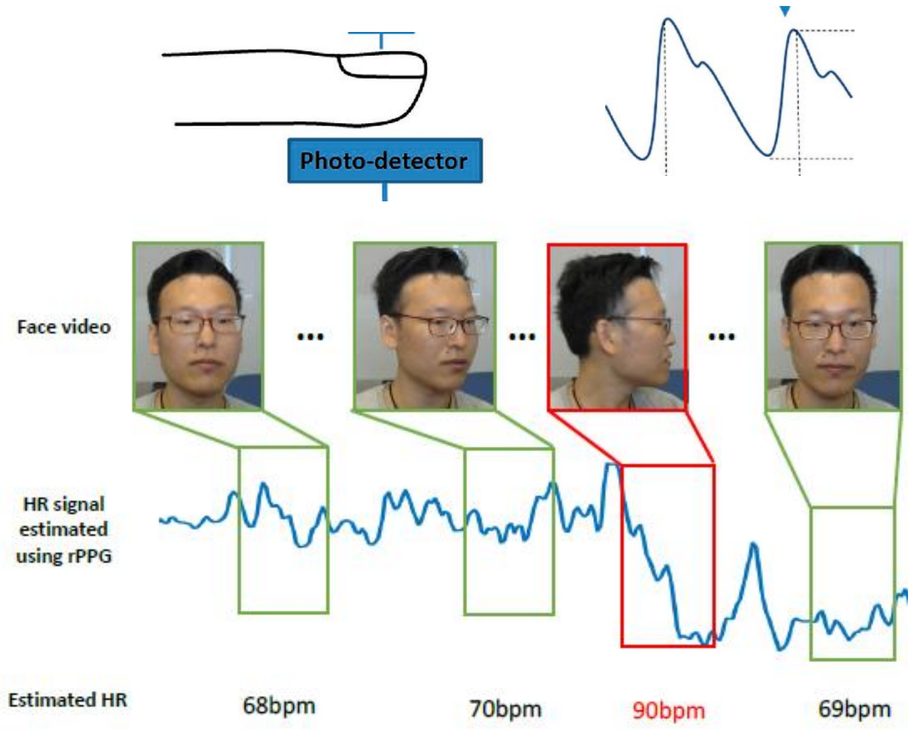
- Face analysis for health monitoring and security
- Multimodal iris and sclera biometrics
- Multiscript signature verification and recognition
- Lip biometrics
- Tattoo biometrics
- Script recognition
- Bird call recognition

Heart rate estimation from face videos

Introduction: HR estimation

- Remote photoplethysmography (rPPG) signals can be used for Heart Rate (HR) estimation
- rPPG based HR measurement has shown promising results under controlled conditions





Literature: HR estimation

HR estimation

- *Blind signal separation*: independent component analysis (ICA) to temporally filter red, green, and blue (RGB) color channel [1].
- *Optical model based methods*: Prior knowledge skin optical model RGB color channel analysis [2].
- *Data-driven methods*: aim at leveraging big training data to perform remote HR estimation for example employing spatial and temporal cues [3]

Representation Learning Utilizing Attention

Channel and spatial level attention is proposed in [4].

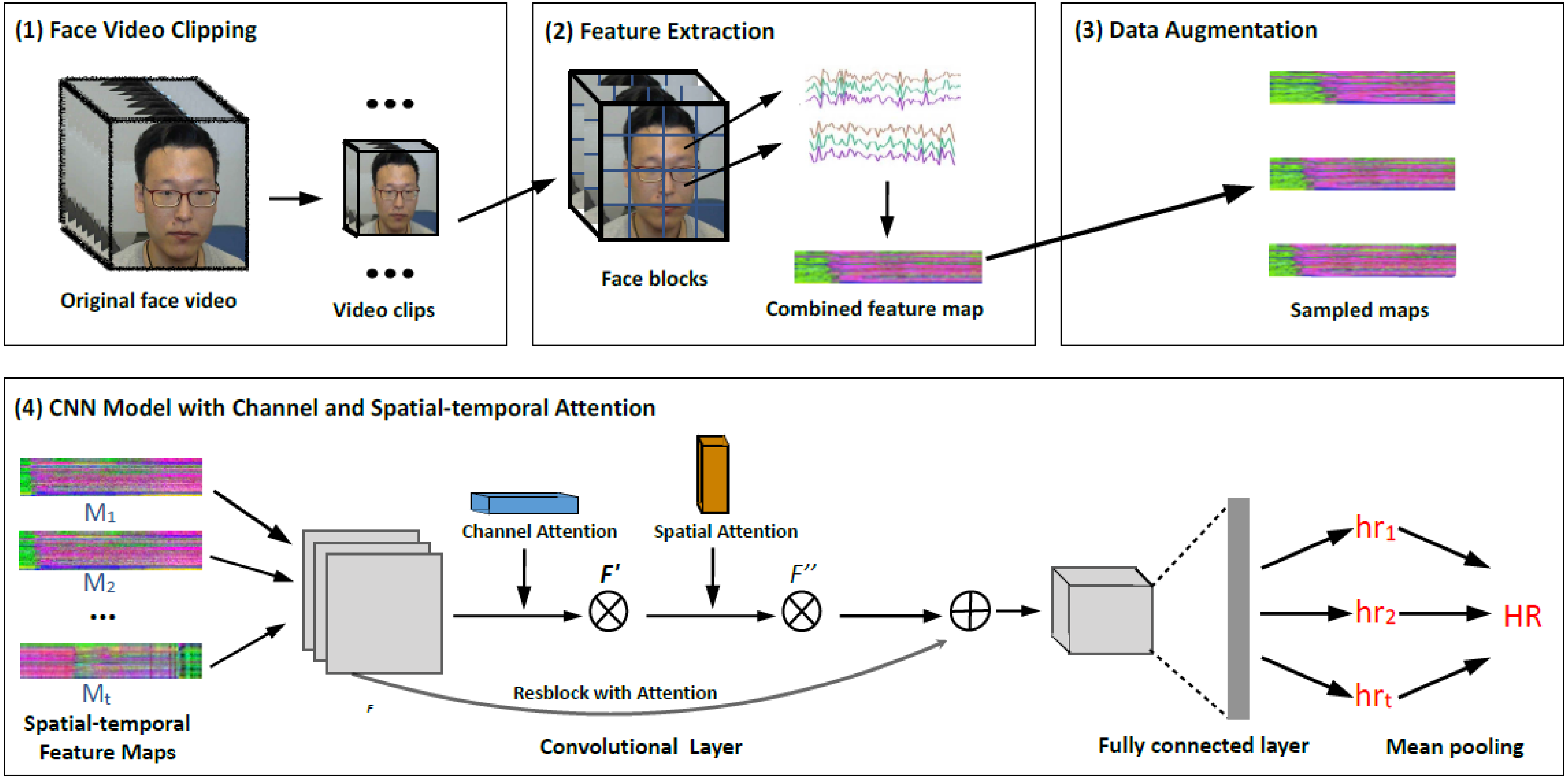
[1] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation." Opt. Express, vol. 18, no. 10, pp. 10 762–10 774, 2010.

[2] G. De Haan and V. Jeanne, "Robust pulse rate from chrominance based rppg," IEEE Trans. Biomed. Eng., vol. 60, no. 10, pp. 2878–2886, 2013.

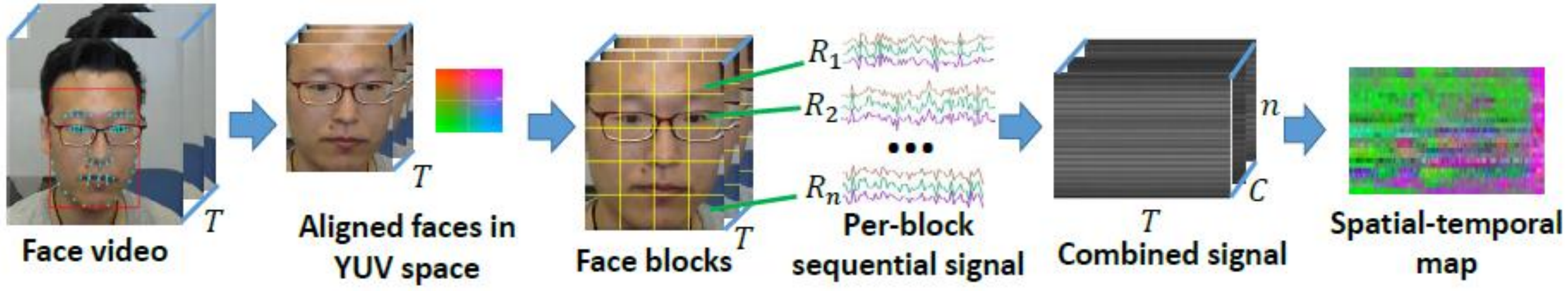
[3] X. Niu, H. Han, S. Shan, and X. Chen, "Synrhythm: Learning a deep heart rate estimator from general to specific," in Proc. IAPR ICPR, 2018.

[4] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in Proc. ECCV, 2018.

Proposed methodology: HR estimation



Proposed methodology: spatial-temporal map

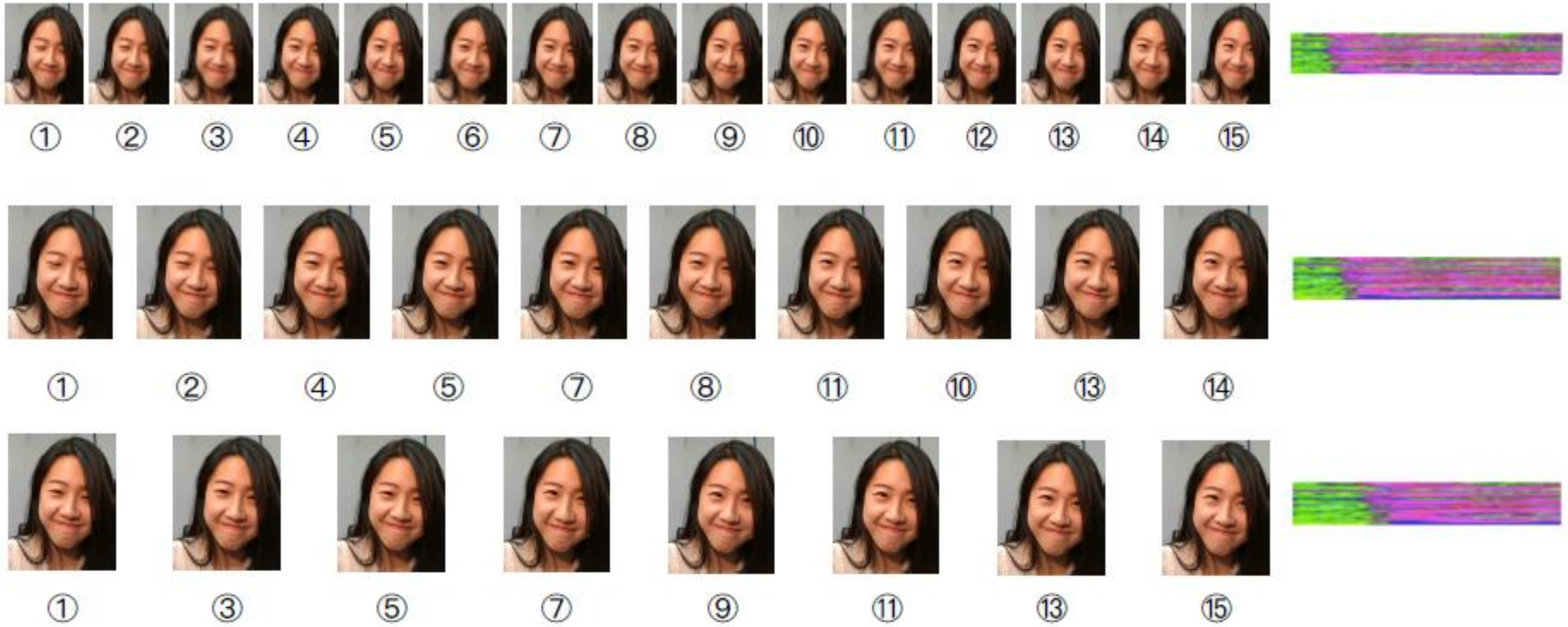


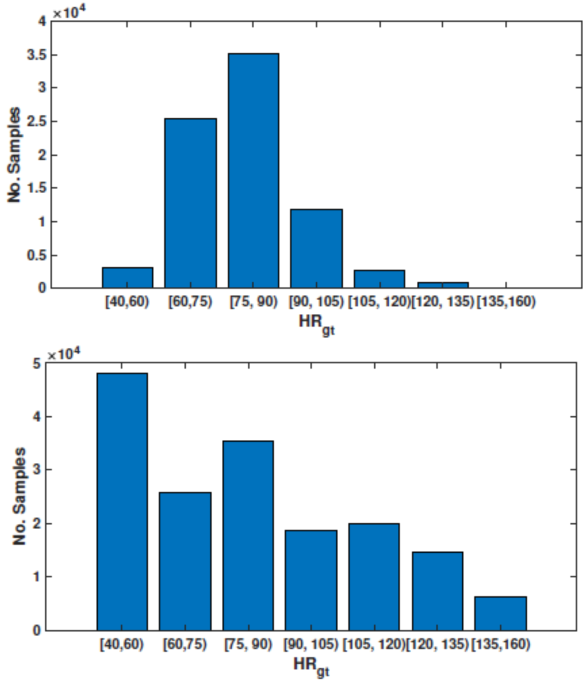
- Face alignment is performed based the two eye centers.
- Bounding box with the size of w×1.5h, where w= horizontal distance between the outer cheek border points h = vertical distance between chin location and eye center points.
- Skin segmentation is then applied to the predefined remove
- Average of the pixel values of each grid is calculated, and then concatenated into a sequence of T for C channels.

A. Das et al., 'Robust Remote Heart Rate Estimation from

Proposed methodology: data augmentation

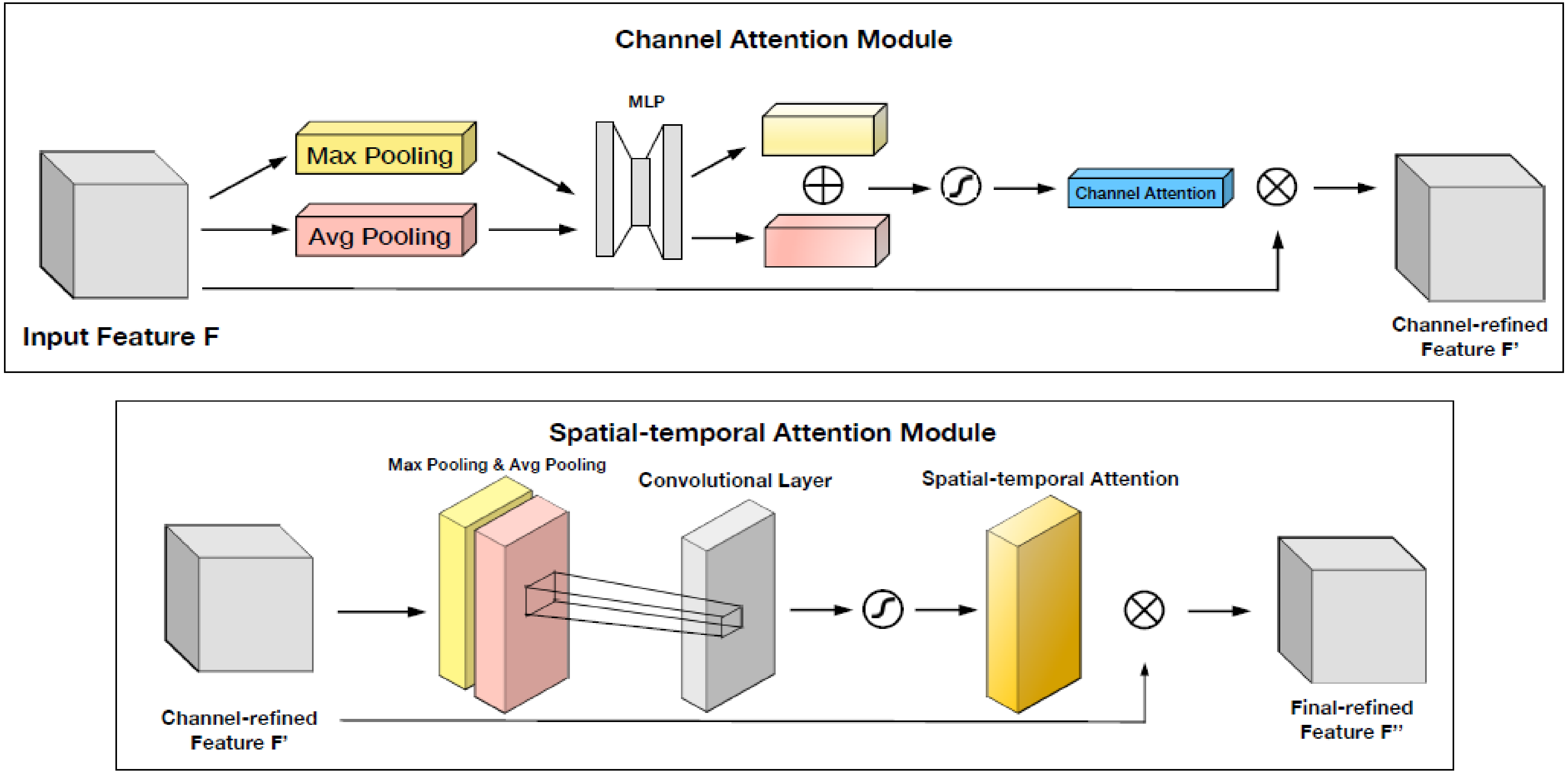
- Videos which the ground-truth HRs range between 60 bpm and 110 bpm, were down-sampling with sampling rate of 0.67
- Videos with a ground-truth HRs range between 70 bpm and 85 bpm the sampling rate is 1.5





A. Das et al., 'Robust Remote Heart Rate Estimation from

Proposed methodology: attention mechanism



A. Das et al., 'Robust Remote Heart Rate Estimation from

A. Das et al., "Robust Remote Heart Rate Estimation from Face Videos Utilizing Channel and Spatial-temporal Attention"

Experimental results: dataset and setup details

	No. Subj.	No. Vids.	Video Length	Protocol
MMSE -HR[1]	40	102	30s	cross -database
VIPL -HR[2]	107	2,378	30s	five-fold

Performance measure used

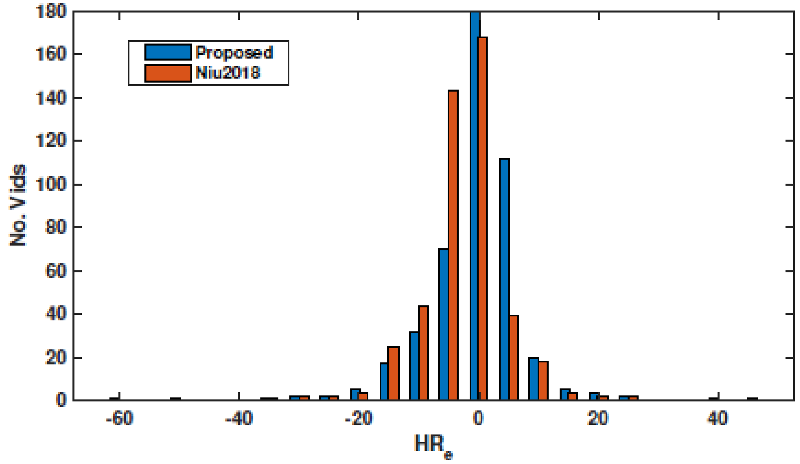
- Mean (HR_{me}) of the HR error
- Standard deviation (HR_{std}) of the HR erro
- Mean absolute HR error (HR_{mae})
- Root mean squared HR error (HR_{rmse})
- Mean of error rate percentage (HR_{mer})
- Parsons correlation coefficients r

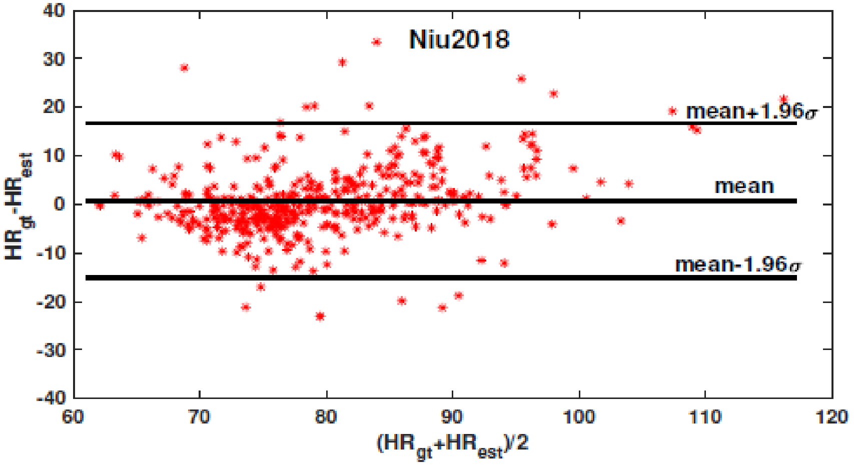
- The model is implemented based on the PyTorch4 framework.
- For the proposed approach face ROI were divide it into 5×5 block.
- The number of maximum iteration epochs employed is 50, and the batch size is 100.
- The model was first trained from scratch with learning rate of 0.001 with Adam solver and the trained model is further trained including the attention with learning rate to 0.0015.

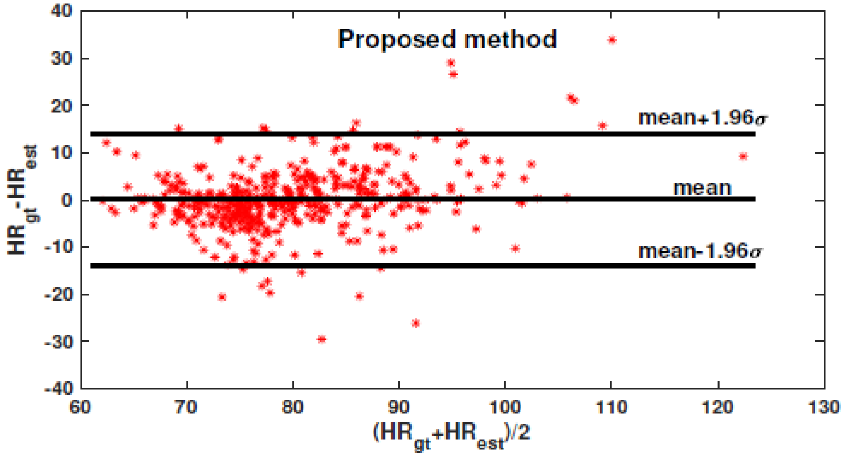
A. Das et al., 'Robust Remote Heart Rate Estimation from

Experimental results: on VIPL-HR dataset

Method	HRme (bpm)	HRsd (bpm)	HRmae (bpm)	HRrmse (bpm)	HRmer	r
Haan2013 [3]	7.63	15.1	11.4	16.9	17.8%	0.28
Tulyakov2016 [1]	10.8	18.0	15.9	21.0	26.7%	0.11
Wang2017 [4]	7.87	15.3	11.5	17.2	18.5%	0.30
Niu2018 (ResNet-18) [2]	1.02	8.88	5.79	8.94	7.38%	0.73
ResNet-18 + DA	-0.08	8.14	5.58	8.14	6.91%	0.63
Proposed	-0.16	7.99	5.40	7.99	6.70%	0.66



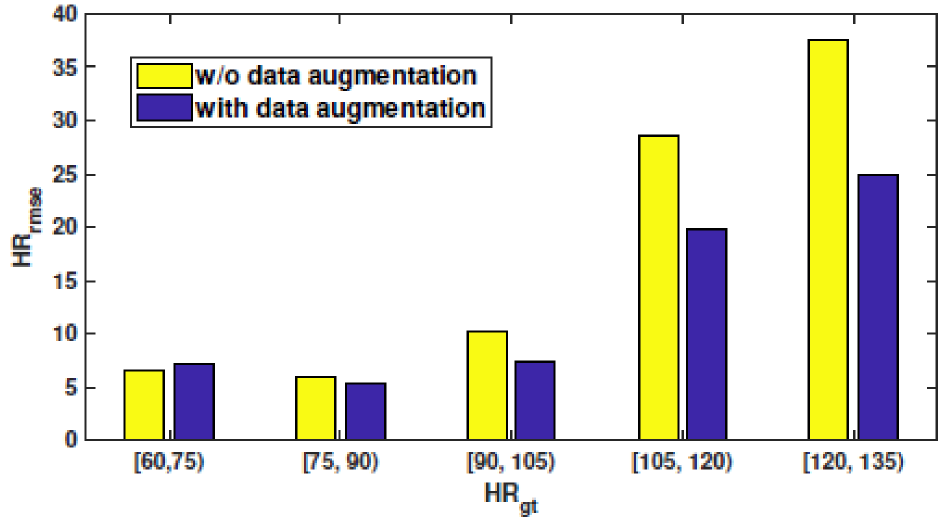


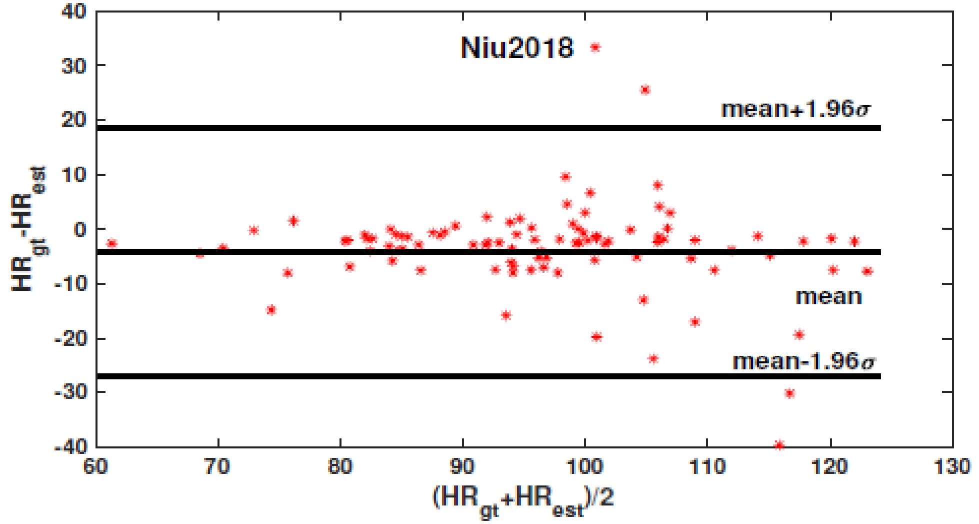


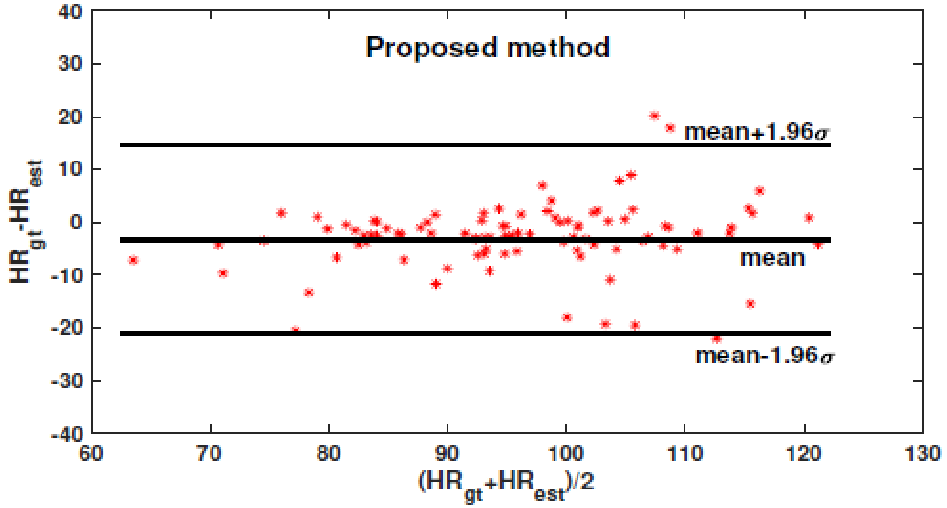
bust Remote Heart Rate Estimation from

Experimental results: on MMSE-HR dataset

Method	HRme (bpm)	HRsd (bpm)	HRrmse (bpm)	HRmer	r
Li2014 [5]	11.56	20.02	19.95	14.64%	0.38
Haan2013 [3]	9.41	14.08	13.97	12.22%	0.55
Tulyakov2016 [1]	7.61	12.24	11.37	10.84%	0.71
Niu2018 [2]	-2.26	10.39	10.58	5.35%	0.69
Proposed	-3.10	9.66	10.10	6.61%	0.64







Conclusions and future scopes on HR estimation

- We propose an end-to-end learning network for HR estimation based on channel and spatial-temporal attention.
- We also design an effective video augmentation method to overcome the limitation of training data.
- Experimental results on the VIPL-HR and MMSE-HR datasets show the effectiveness of the proposed method.
- Future work includes the expansion of the work onto continuous HR measurement.
- Additionally, remote measurement of further physical signals, such as breath rate, heart rate variability, will be studied.

Reference

[1] S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, J. F. Cohn, and N. Sebe, “Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions,” in Proc. IEEE CVPR, 2016, pp. 2396–2404.

[2] X. Niu et al., VIPL-HR: A multi-modal database for pulse estimation from less-constrained face video,” in Proc. ACCV, 2018.

[3] G. De Haan and V. Jeanne, “Robust pulse rate from chrominancebased rppg,” IEEE Trans. Biomed. Eng., vol. 60, no. 10, pp. 2878– 2886, 2013.

[4] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, “Algorithmic principles of remote ppg,” IEEE Trans. Biomed. Eng., vol. 64, no. 7, pp. 1479–1491, 2017.

[5] X. Li, J. Chen, G. Zhao, and M. Pietikainen, “Remote heart rate measurement from face videos under realistic situations,” in Proc. IEEE CVPR, 2014, pp. 4264–4271.

[6] A. Das et al., “Robust Remote Heart Rate Estimation from Face Videos Utilizing Channel and Spatial-temporal Attention”, submitted to FG 2019.

A. Das et al., ‘Robust Remote Heart Rate Estimation from

Bias in face analysis

Introduction: bias in face analysis

- Biasness in training and evaluation data (in automatic face recognition)
- Problems
  - Produces skewed results [1]
  - Young individuals (18-30 years) exhibit low accuracy for face recognition. (US-based law enforcement)
  - Algorithms performed worse for females than males (National Institute for Standards and Technology -NIST[3])

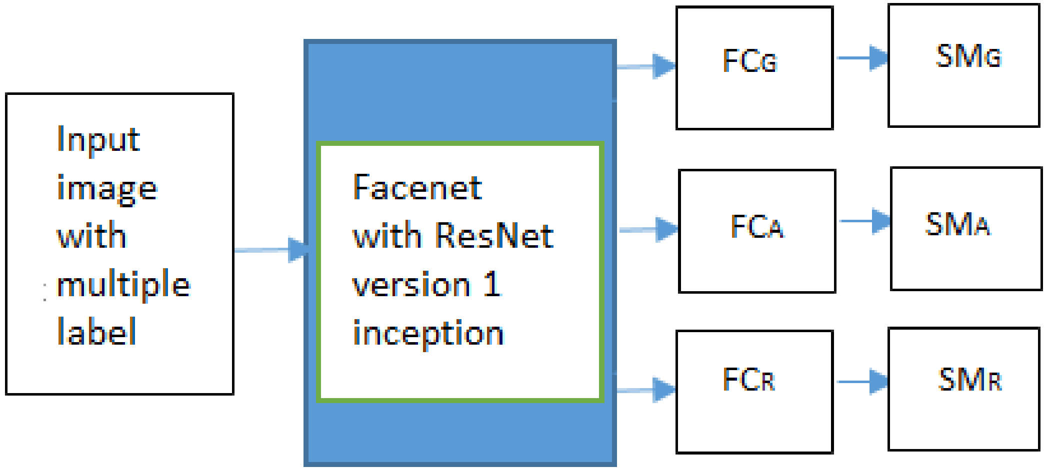
[1] J. Buolamwini, T. Gebru, : Gender shades: Intersectional accuracy disparities in commercial gender classification. In: Conference on Fairness, Accountability and Transparency. (2018) 77-91

[2] B. F. Klare et al. Face recognition performance: Role of demographic information. IEEE Transactions on In- formation Forensics and Security 7(6) (2012) 1789{1801

[3] M. Ngan, et al.,.: Face recognition vendor test (FRVT) performance of automated gender classification algorithms. US Department of Commerce, National Institute of Standards and Technology (2015)

Proposed methodology: bias in face analysis

- Multi-Task CNN (MTCNN) Joint dynamic weight loss.
- Facenet with ResNetV1 inception was the base model employed.
- Summed initial weight for the each classification task was obtained by brute-force search on the validation set.
- Mini-batch Stochastic Gradient Descent was employed to solve the above optimization problem of loss weight followed by weights averaged for each batch



$$loss_{total} = W_A * loss_A + W_G * loss_G + W_R * loss_R$$

$$W_A + W_G + W_R = 1$$

Experimental result: datasets

- ***UTKFace*** consist of over 20,000 face images with long age span ranged from 0 to 116 years.
- Specially, annotation for age includes following classes: baby: 0-3 years, child: 4-12 years, teenagers: 13-19 years, young: 20-30 years, adult: 31-45 years, middle aged: 46-60 years and senior: 61 and above years.
- The dataset additionally contains the labeling for gender (male and female) and races (White, Black, Asian, Indian and other race).
- Bias Estimation in Face Analytics **(*BEFA*)** challenge dataset is contains 13431 test images.

[1] https://sites.google.com/site/eccvbefa2018/home?authuser=0

[2] Z. Zang, et al.,.: Age progression / regression by conditional adversarial autoencoder. In: IEEE Conference on Computer Vision and Pattern Recognition

Experimental result: on UTKFace

	Race	Gender	Age
Facenet	85.1	91.2	56,9
Finetuned Facanet (FFNet)	86.1	96.1	64
Proposed MTCNN	90.1	98.23	70.1

	FFNet	Proposed
Male	61,5	69.1
Female	66,8	70.9
White	61,8	69.6
Black	59,2	68.9
Asian	78,7	80.1
Indian	63,6	69.5
Others	64.5	66.7

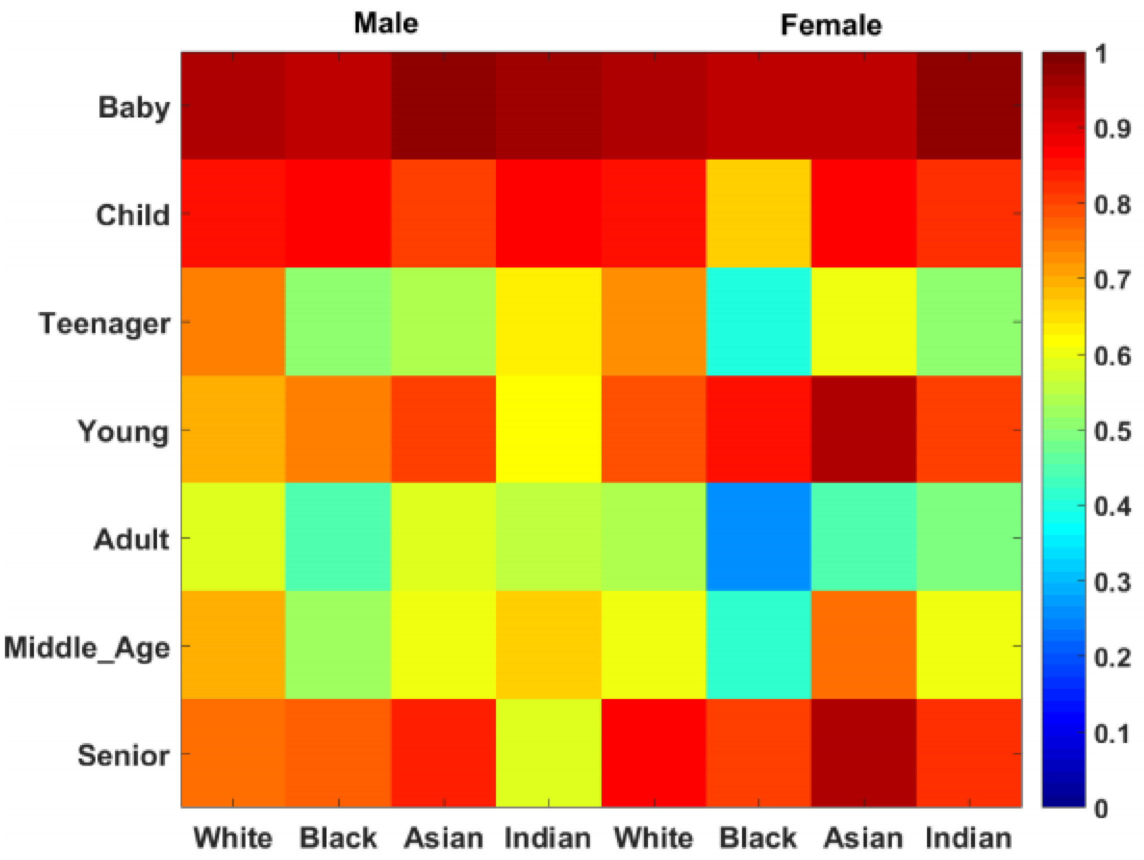
	FFNet	Proposed
Male	87	90.9
Female	84.3	89.1
Baby	100	100
Child	84.3	88.8
Teenager	85.8	89.1
Young	84.9	88.9
Adult	88.1	91.5
Middle	87.7	90.7
Senior	81.9	87.7

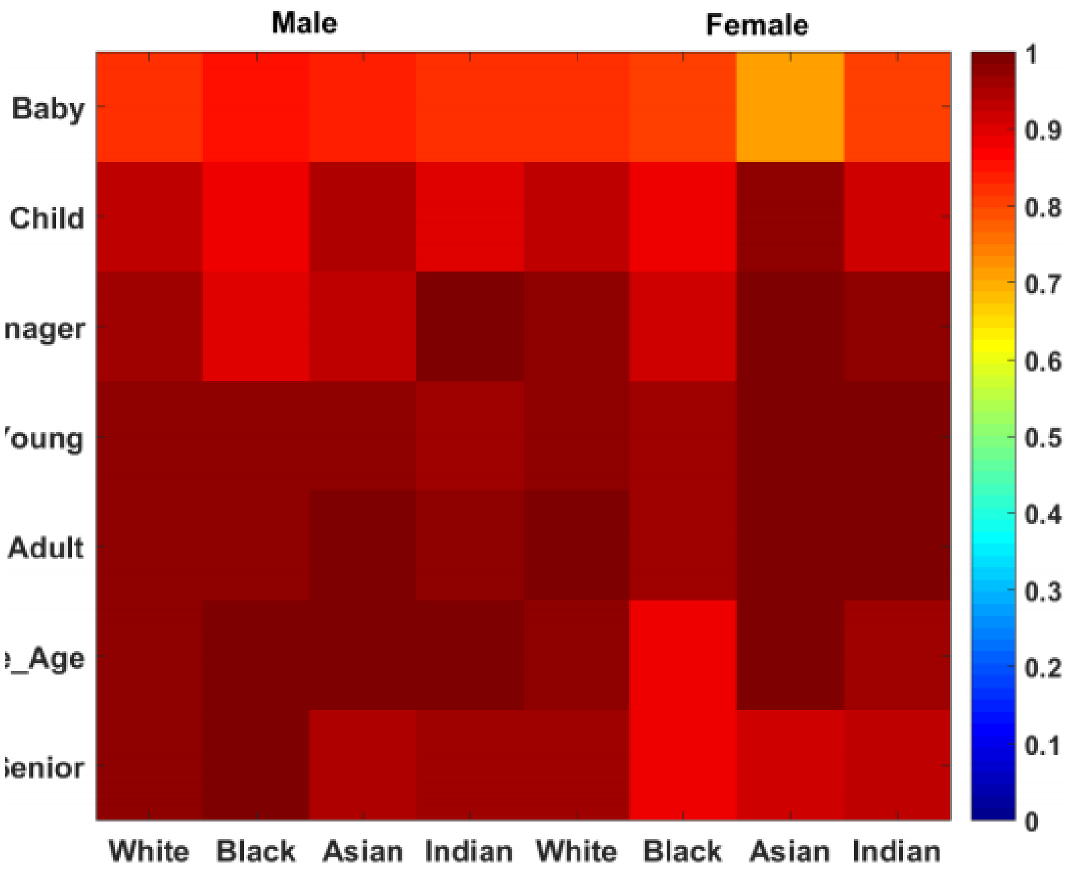
	FFNet	Proposed
Baby	70	80.5
Child	79,6	96.7
Teenager	92	95.
Young	96,8	97
Adult	97,7	98.3
Middle	96,6	97.7
Senior	95,1	96.5
White	97	98.7
Black	95	98.6
Asian	97,5	99.3
Indian	97,8	99.1
Others	97.5	99

A. Das et al., "Mitigating Bias in Gender, Age and Ethnicity Classfication: a Multi-Task Convolution Neural Network Approach" EECW 2018.

Experimental result: on BEFA

	All Attributes	Race	Gender	Age
MTCNN	56.37	84.29	93.72	71.83

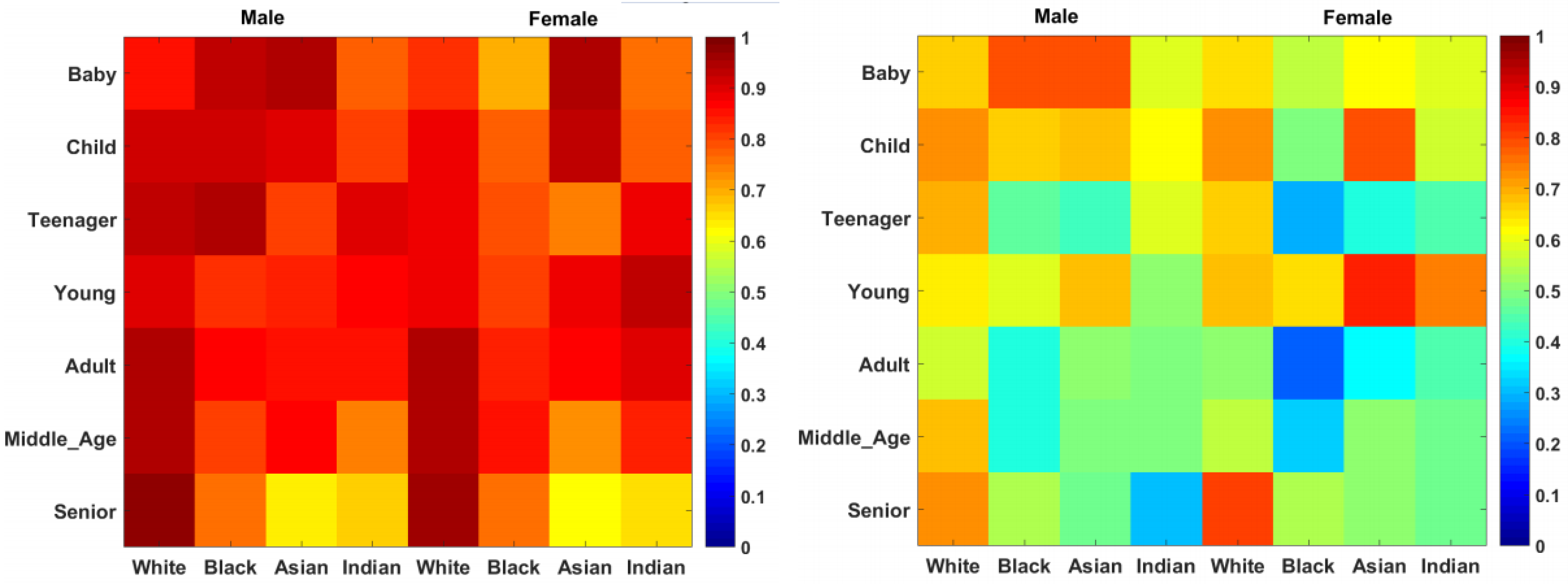




A. Das et al., 'Robust Remote Heart Rate Estimation from

A. Das et al., "Mitigating Bias in Gender, Age and Ethnicity Classfication: a Multi-Task Convolution Neural Network Approach" EECW 2018.

Experimental result: on BEFA



A. Das et al., 'Robust Remote Heart Rate Estimation from

A. Das et al., "Mitigating Bias in Gender, Age and Ethnicity Classfication: a Multi-Task Convolution Neural Network Approach" EECW 2018.

Conclusion and future scope: bias in face analysis

- Presented an approach for gender, age and race classification targeted to minimize inter-class bias.
- The proposed multi-task CNN approach utilized joint dynamic loss, providing promising.
- The proposed algorithm was ranked 1st in the BEFA-challenge of the ECCV 2018.
- Intend to extend the current study onto other facial attributes and face recognition.

A. Das et al., 'Robust Remote Heart Rate Estimation from

Team and collaborators

STARS, Inria

- Antitza Dantcheva
- Francois Bremond

CAS, China

- Xuesong Niu
- Xingyuan Zhao
- Hu Han
- Shiguang Shan
- Xilin Chen

A. Das et al., 'Robust Remote Heart Rate Estimation from

Thank you!!!

A. Das et al., "Robust Remote Heart Rate Estimation from