

Visual Heart Rate Estimation with Convolutional Neural Network

Radim Špetlík, Vojtěch Franc, Jan Čech, Jiří Matas
Center for Machine Perception, Czech Technical University in Prague

Heart Rate Estimation

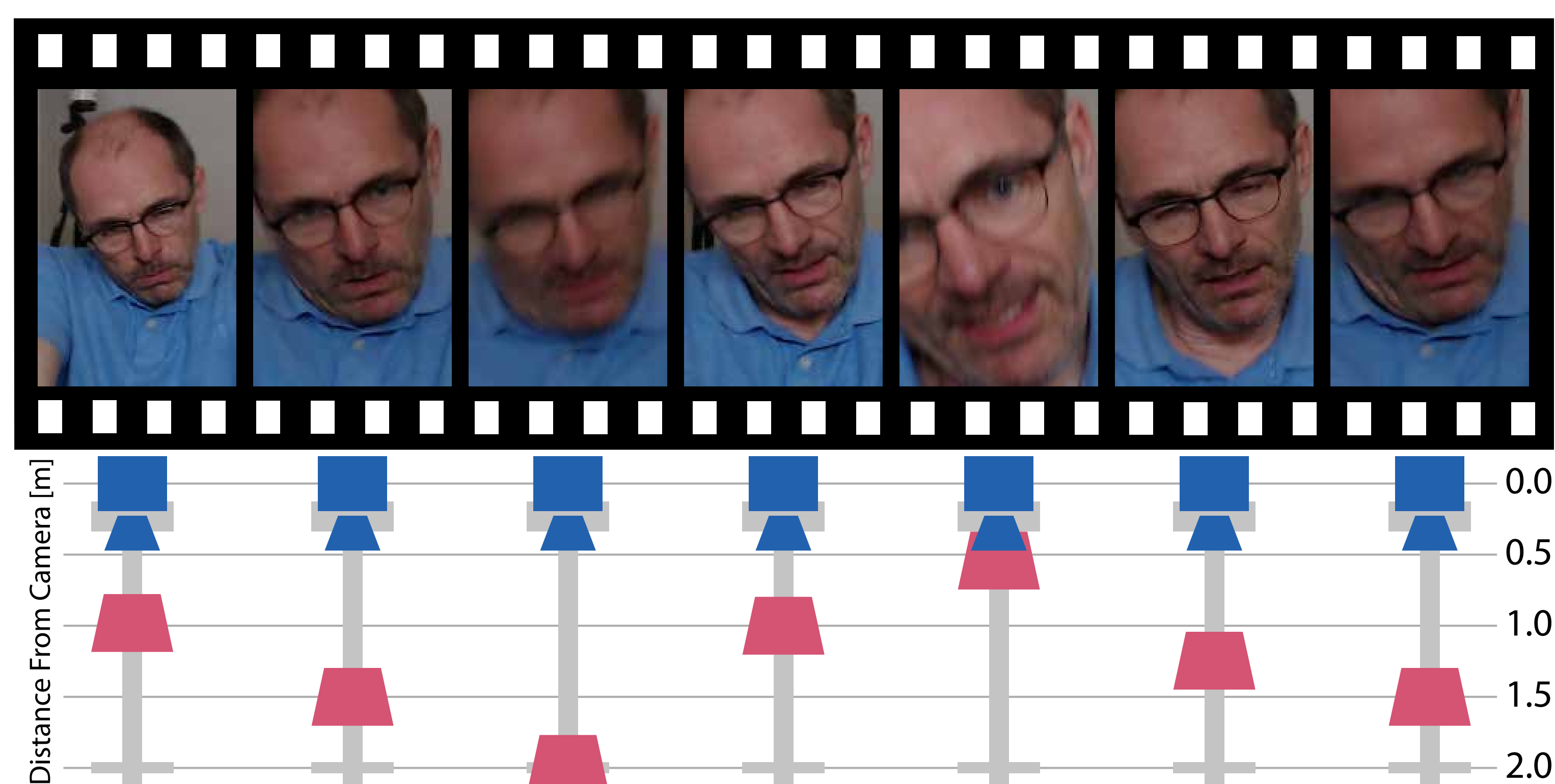
- Heart Rate (HR) is a basic parameter of cardiovascular activity.
- HR value is used broadly, its measurement is performed precisely with expensive contact devices (ECG, pulse oximeters, ...).
- Visual HR estimation, i.e. HR estimation from a video sequence, delivers precise measurement using consumer webcams.
- Accuracy of visual HR estimation depends on recording conditions – prior visual HR methods are sensitive to motion and light interference, requiring subject's cooperation.

ECG-Fitness Dataset

- New, publicly available dataset of subjects performing physical exercises – 1080p RAW 30fps videos + 100Hz ECG recording.
- 207 videos in total, 17 subjects, 6 videos per subject, 2 cameras.
- HR (bpm) – min: 56, max: 159, mean: 108.96, std: 23.33.



Two cameras: one attached to a tripod, the other to the fitness machine. 4 activities, 2 lighting setups (daylight, halogen).

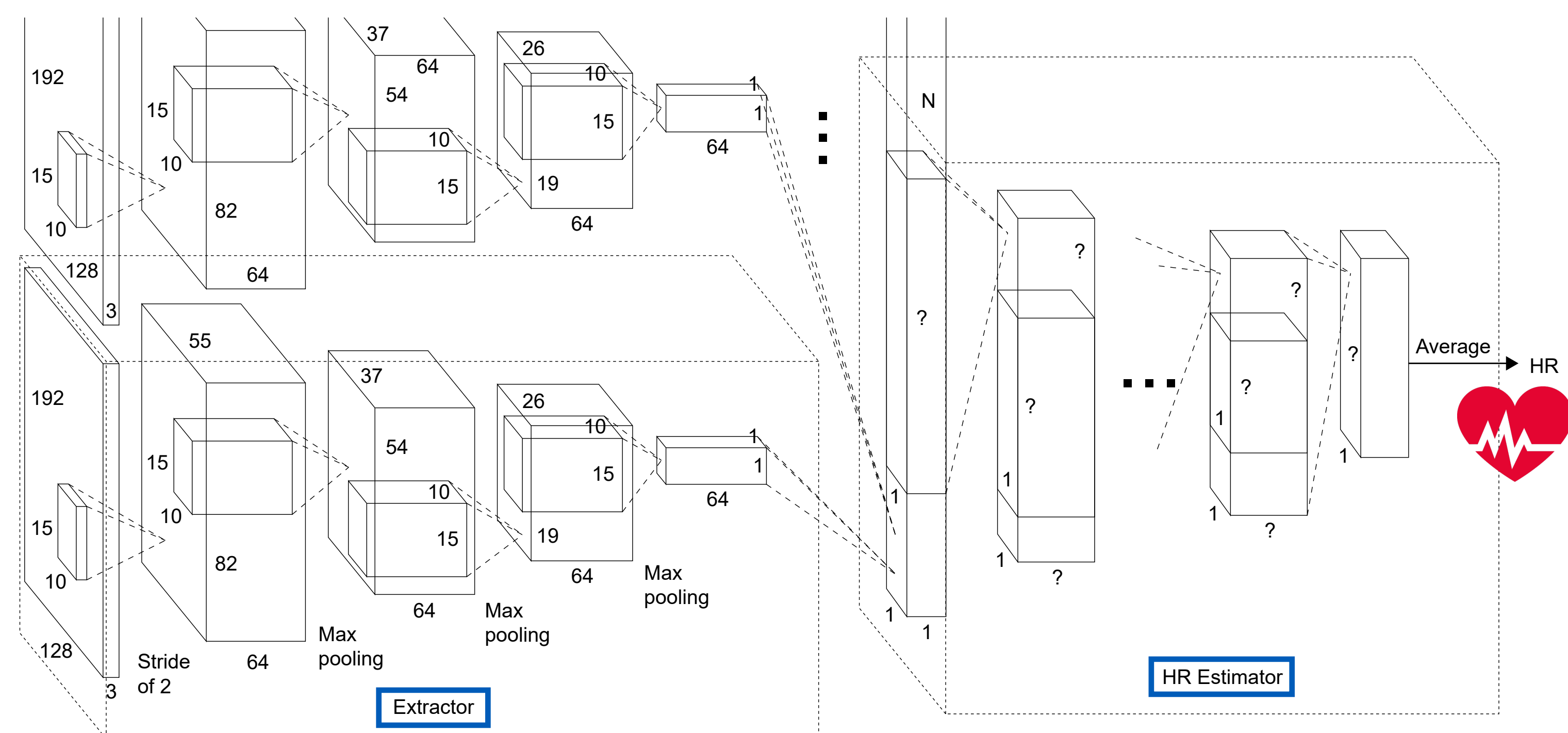


Facial images from a rowing session video of a subject from the ECG-Fitness dataset. Pink - the subject's position, blue - camera.

- Dataset challenges: (i) large subject's motion, (ii) motion blur, (iii) facial expressions, (iv) glasses, (v) non-uniform lighting, (vi) light interference, (vii) atypical non-frontal camera angles.

Proposed Method

- HR-CNN**: a two-step convolutional neural network.
 - Extractor* is run over an image sequence of faces.
 - Estimator* predicts the HR from the output of the Extractor.

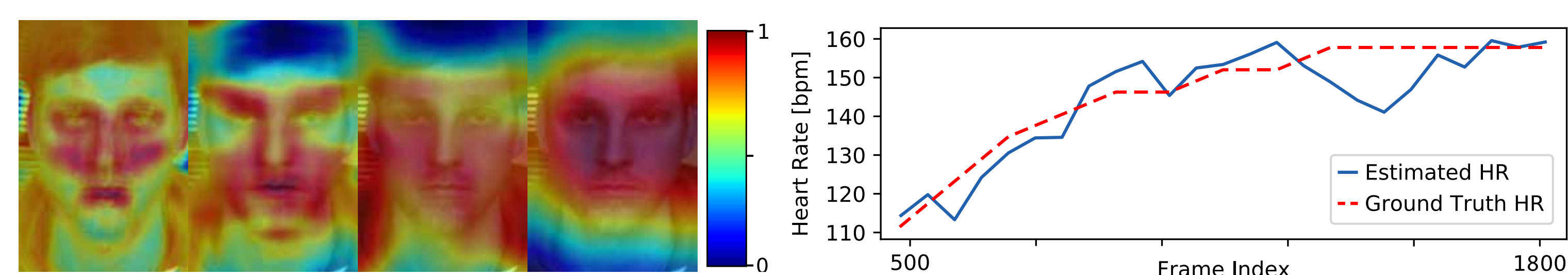


Experiments

| | | COHFACE | ECG-Fitness | MAHNOB | PURE | PURE MPEG-4 Visual |
|---|---------------|---------|-------------|--------|--------|-----------------------|
| MAE [bpm] | 2SR | 20.98 | 43.66 | 13.84 | 2.44 | 5.78 ① |
| | CHROM | 7.80 ① | 21.37 | 13.49 | 2.07 | 6.29 |
| | LiCVPR | 19.98 | 63.25 | 7.41 | 28.22 | 28.39 |
| | SAMC* | — | — | — | — | — |
| | HR-CNN | 8.10 | 14.48 ① | 7.25 ① | 1.84 ① | 8.72 |
| Pearson's correlation coefficient | 2SR | -0.32 | 0.06 | 0.14 | 0.98 | 0.43 |
| | CHROM | 0.26 | 0.33 | 0.21 | 0.99 ① | 0.55 |
| | LiCVPR | -0.44 | -0.02 | 0.45 | -0.38 | -0.42 |
| | SAMC* | — | — | 0.83 ① | — | — |
| | HR-CNN | 0.29 ① | 0.50 ① | 0.51 | 0.98 | 0.70 ① |

Test sets: COHFACE, ECG-Fitness, MAHNOB and PURE databases for three baseline methods and the proposed method.

Conclusion



Grad-Cam heatmaps of convolutional layers. Output example of the HR estimator for a challenging video.

- HR-CNN tested on three publicly available datasets against three baseline methods achieves state-of-the-art HR prediction.
- Challenging publicly available dataset ECG-Fitness with 60 s videos of people performing physical exercises was introduced.
- The proposed method yields significantly the best on the ECG-Fitness dataset.

