CNN-based Eye Landmark Estimation

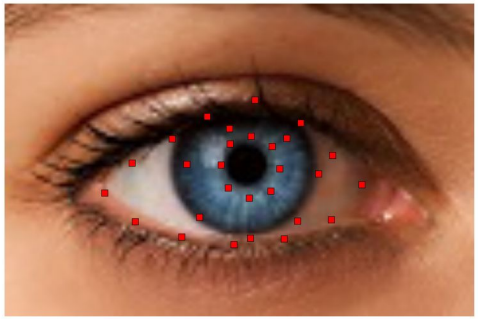
Dana **Kianfar**
Jose **Gallego**

Supervisor:
Roberto **Valenti**

SightCorp

Motivation

- Landmark Estimation
  - Eye-shape registration
  - Appearance-based gaze estimation
  - Market research, VR, driver eye tracking
- CNN vs. feature-based methods



Previous Work

- Zhang, Luo, et al. [1] and Wu, Hassner, et al. [2] explore facial landmark estimation with CNNs
- Lack of work on estimating eye features
- Our approach estimates accurate eye landmarks from a synthetic dataset

Syntheseyes

- 11,382 synthesized close-up images of left eyes - 80 x 120 pixels
- 2D and 3D eye landmarks generated from synthetic models
- Variation across head poses, gaze directions, illumination conditions, and facial features [3]

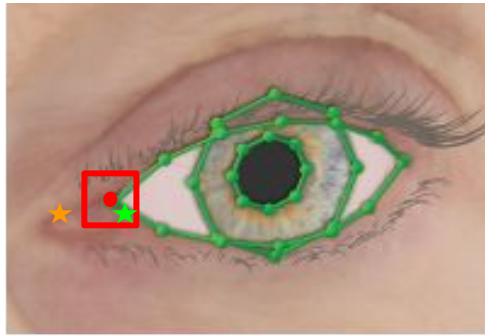
Application Issues

- Privacy
- Portability
- Robustness

Metrics

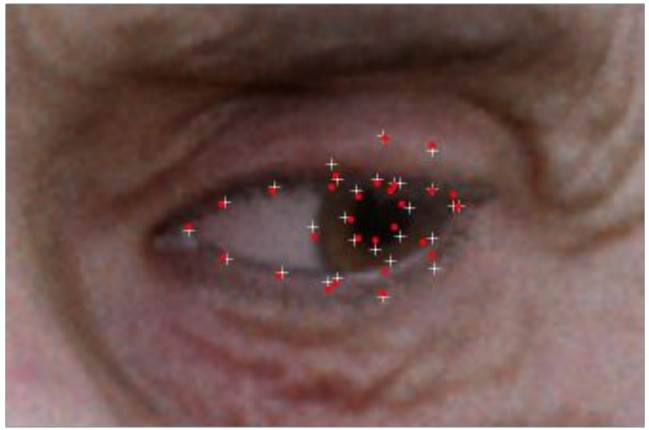
- MAE
- MSE
- Landmark Loss

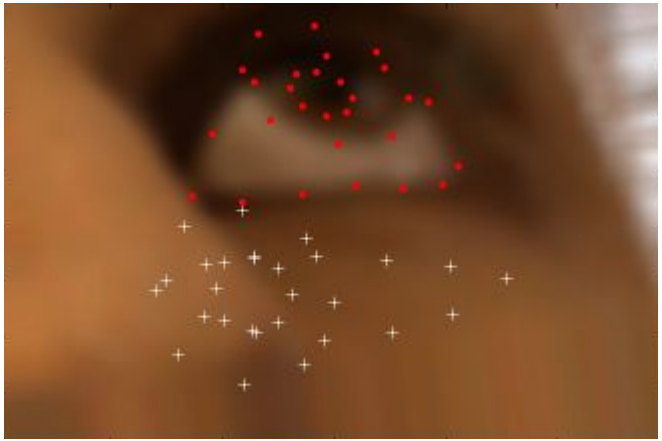
- 5x5 Landmark Accuracy



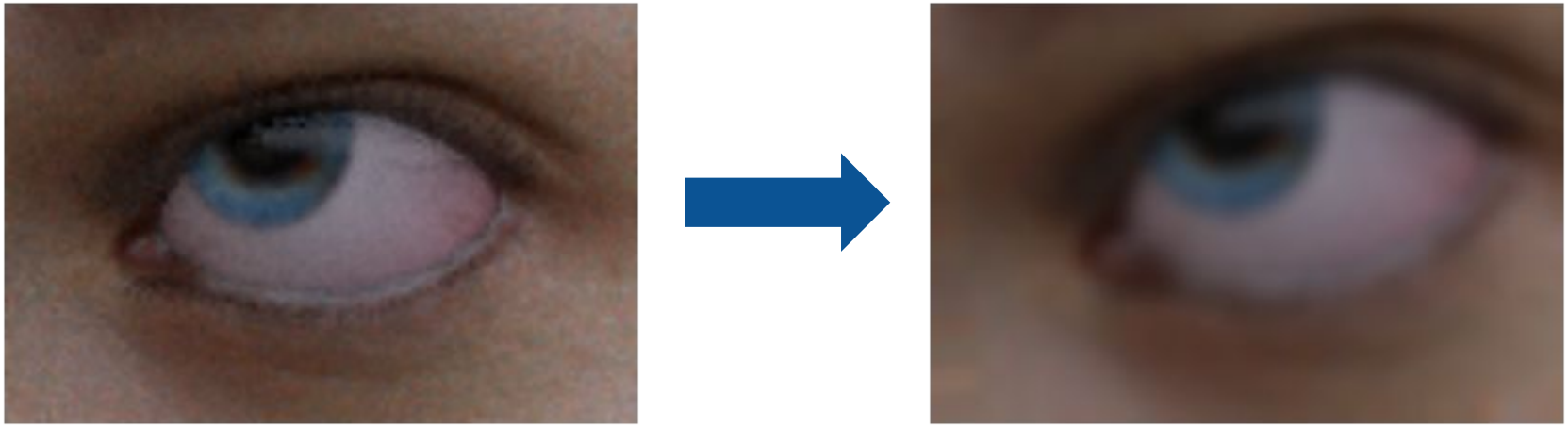
Robustness

Training on normal data not sufficient for generalization

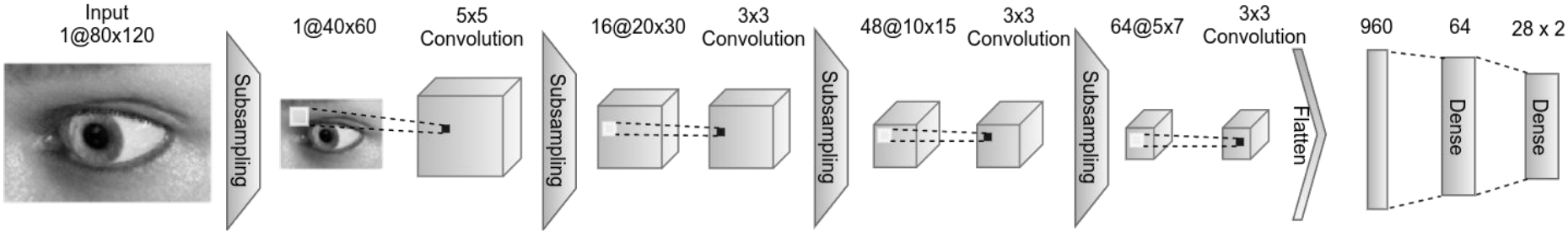




Data Augmentation

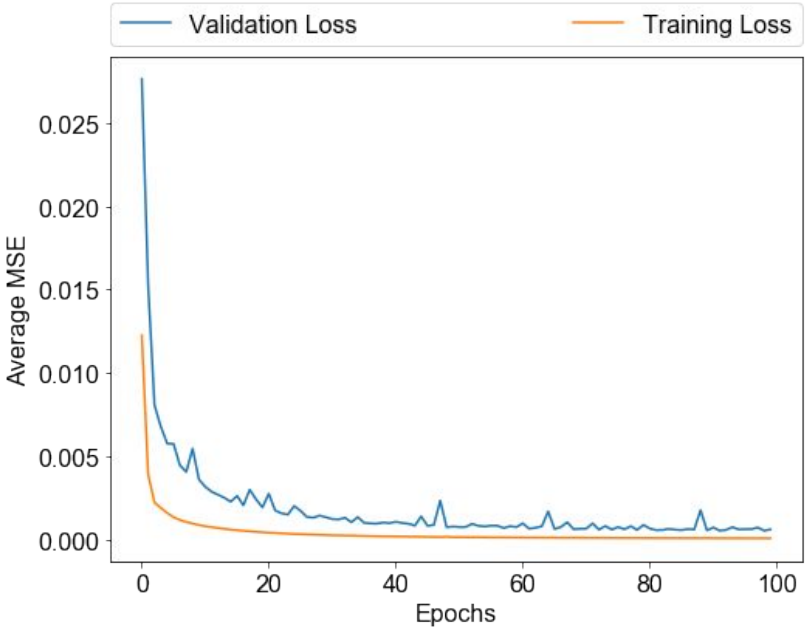


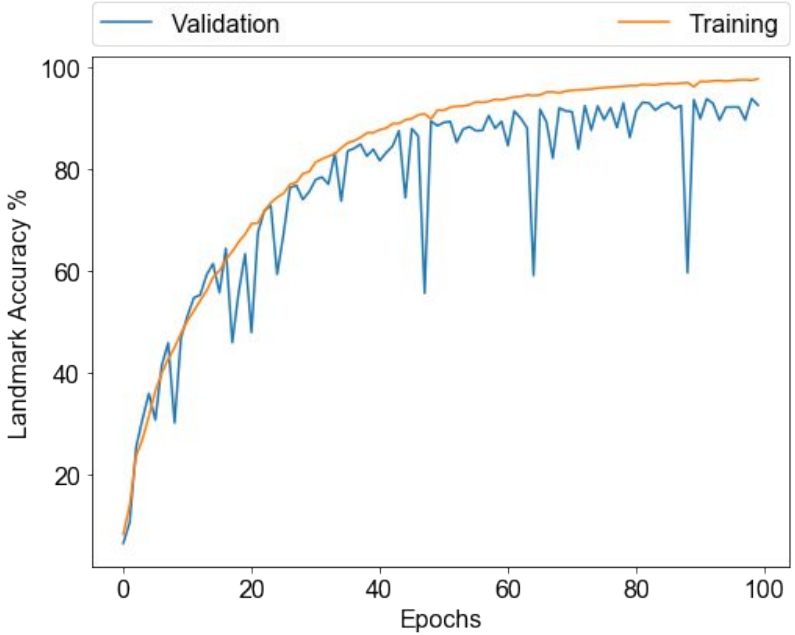
Model



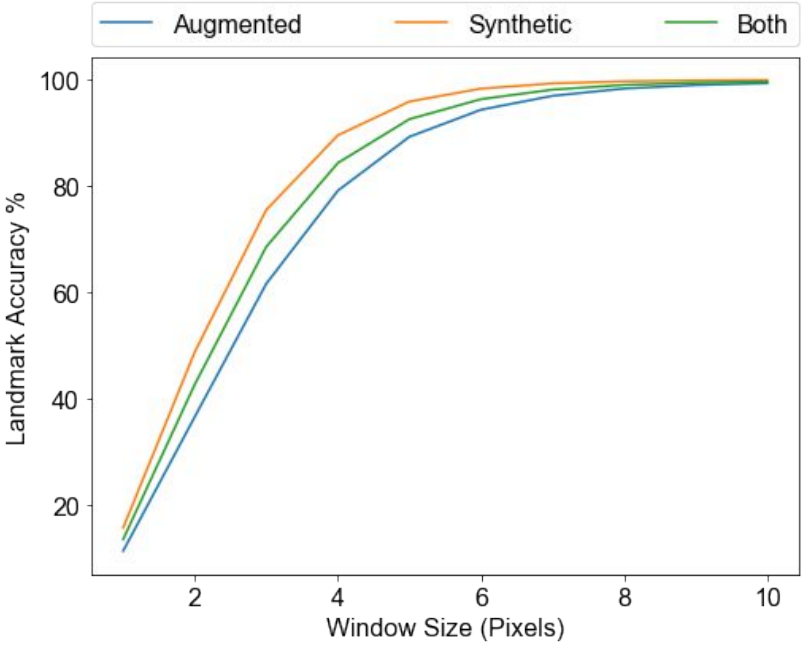
- 100 epochs with Adam
  - Normal data: **96.7%**
  - Full data: **92.5%**
- 80-20 train/test split
- 130K params
- 0.5s per 1K batch

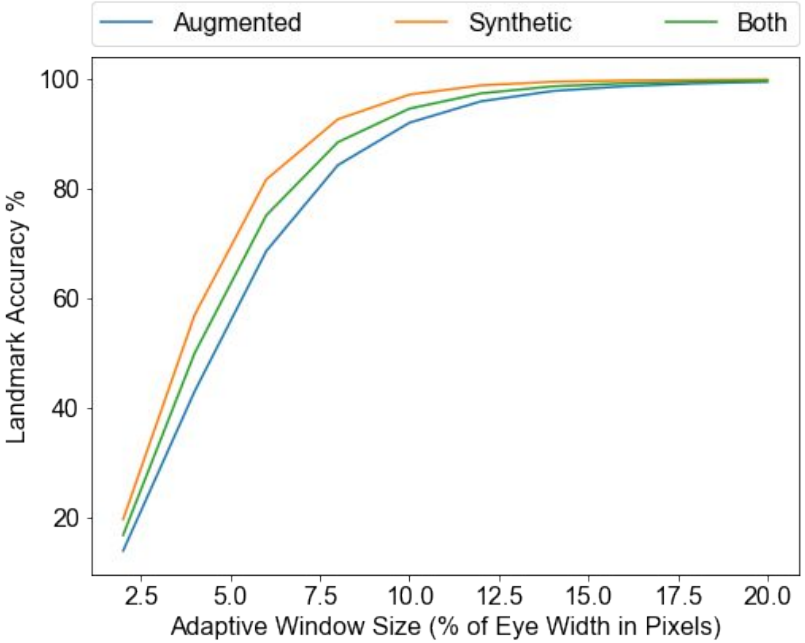
Results


Validation Loss
Training Loss
Average MSE
0.025
0.020
0.015
0.010
0.005
0.000
0
20
40
60
80
100
Epochs

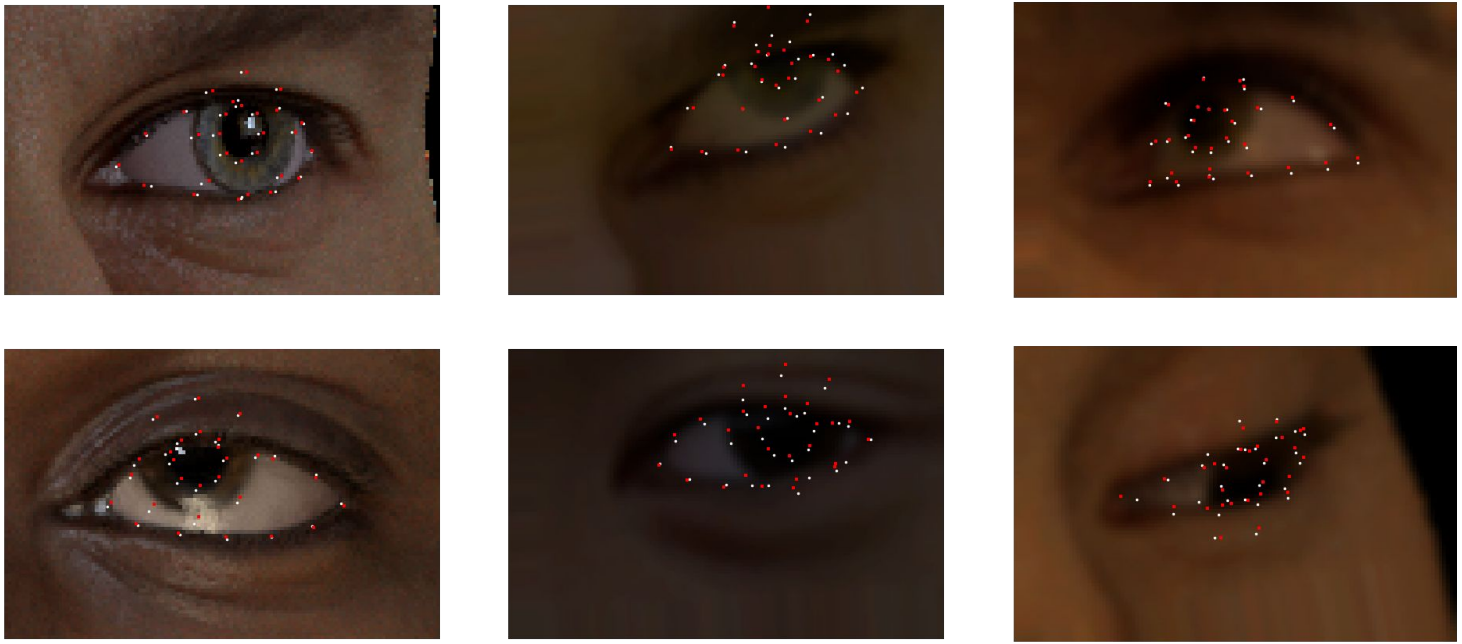

Validation
Training
Landmark Accuracy %
100
80
60
40
20
0
20
40
60
80
100
Epochs

Sensitivity

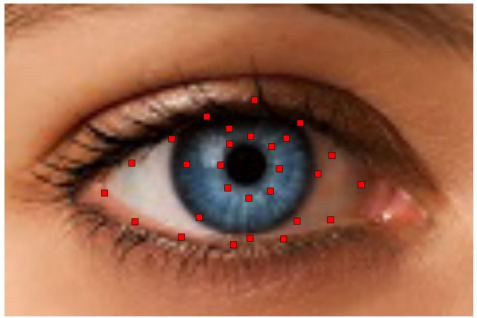
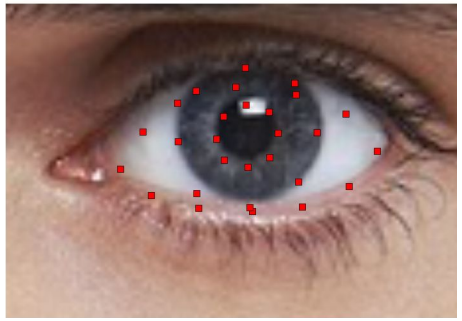
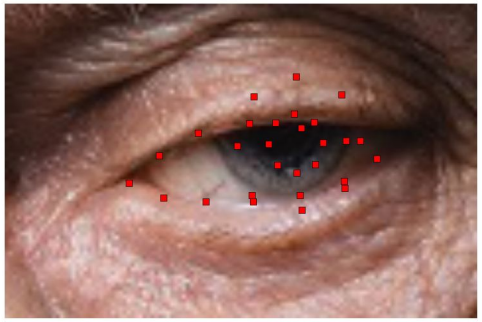
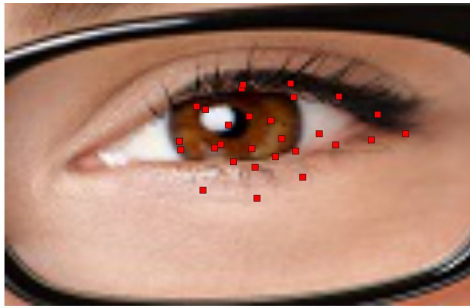
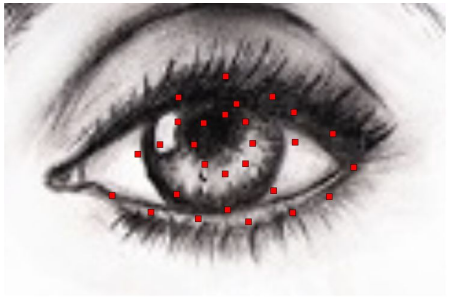
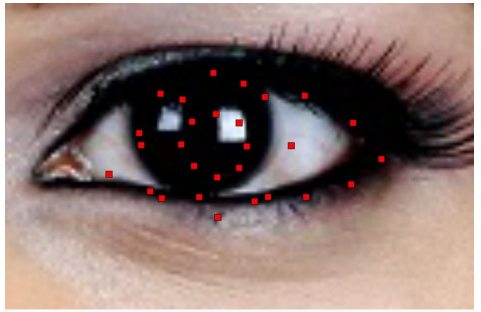




Validation Performance

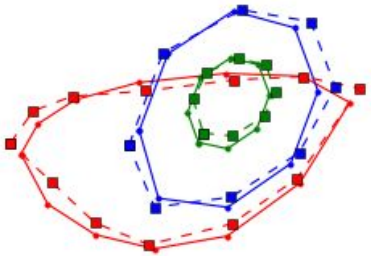


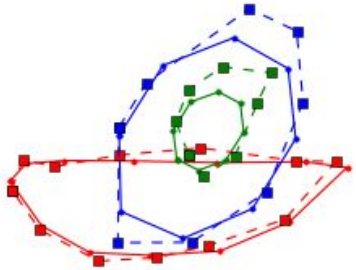
Real Eyes

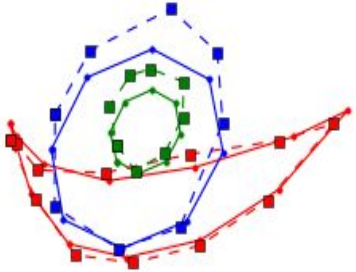







3D Landmarks

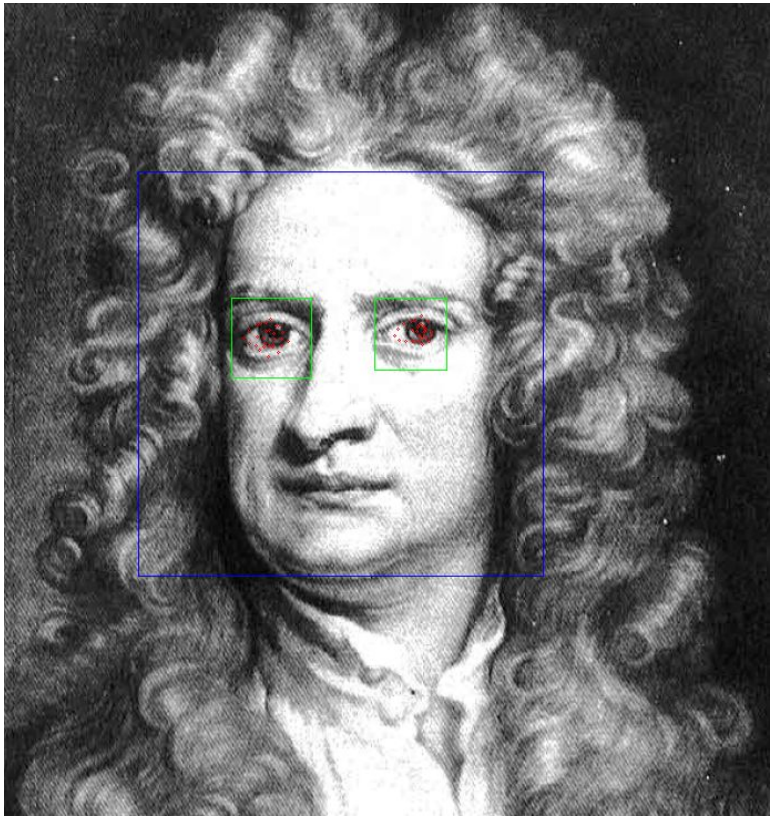
- Harder to quantify performance and visualize results
- Lack of depth information







Live Demo



Conclusions + Future Work

- Robust end-to-end eye landmark estimation from synthetic dataset
- Real-time performance
- Full face landmark detection
- Gaze estimation

References

[1] Zhanpeng Zhang, Ping Luo, Chen Change Loy, Xiaoou Tang. Facial Landmark Detection by Deep Multi-task Learning, in Proceedings of European Conference on Computer Vision (ECCV), 2014

[2] Yue Wu, Tal Hassner, Kang Geon Kim, Gerard Medioni and Prem Natarajan, Facial Landmark Detection with Tweaked Convolutional Neural Networks, arXiv preprint arXiv:1511.04031, 21 Mar 2016

[3] E. Wood, T. Baltrusaitis, X. Zhang, Y. Sugano, P. Robinson, A. Bulling, "Rendering of eyes for eye-shape registration and gaze estimation", ICCV, 2015.