

Neural Photofit: Gaze-based Mental Image Reconstruction - Supplementary Material

Florian Strohm¹, Ekta Sood¹, Sven Mayer², Philipp Müller³, Mihai Bâce¹, Andreas Bulling¹

University of Stuttgart {florian.strohm, ekta.sood, mihai.bace, andreas.bulling}@vis.uni-stuttgart.de

LMU Munich info@sven-mayer.com

German Research Center for Artificial Intelligence (DFKI) philipp.mueller@dfki.de

1. FaceMaker Sliders

Table 1 lists all used FaceMaker sliders. Each slider in FaceMaker allows us to continuously manipulate one specific facial feature. As such, FaceMaker generates human-like faces and at the same time provides fine control over important features like the eyes. For a description of how the sliders manipulate the face in general, as well as the influence of each specific slider, please refer to Schwind et al. [1]. In our work we focused on the four most important features for identification (eyes, nose, mouth and jaw). We defined these features by grouping the FaceMaker sliders as defined in their paper. The subscripts in **Table 1** indicate which groups each feature is assigned to: *e* for eyes, *n* for nose, *m* for mouth and *j* for jaw.

List of FaceMaker slider we used		
Eyes Colour _e	Eyes Height _e	Eyes Size _e
Eyes Opening _e	Eyes Rotation _e	Eyes Shape _e
Eyes Distance _e	Eyebrows Line _e	Eyebrows Shape _e
Eyebrows Colour _e	Lips Volume _m	Lip Size Ratio _m
Mouth Shape _m	Mouth Width _m	Mouth Height _m
Nose Length _n	Nose Shape _n	Nose Width _n
Nose bridge _n	Cheeks _j	Jaw Shape _j
Chin Length _j	Chin Shape _j	Ears Size
Throat Size	Gender	Details
Skin Colour	Forehead	Hair Colour

Table 1. The names of all used facial features in FaceMaker that can be controlled using sliders. The subscripts *e*, *n*, *m*, *j* indicate which groups each feature is assigned to (eyes, nose, mouth and jaw).

2. Activation Maps

The scoring network compares the human attention to neural attention and predicts how relevant the corresponding image feature is for the mental image. For a better intuition on what an activation maps is, **Figure 1** shows exam-

ples for four image features overlaid over the corresponding image. As can be seen, neural attention is highest for areas influenced the most by the corresponding slider, e.g., there is high attention on the eyes for the *Eyes Size* slider.

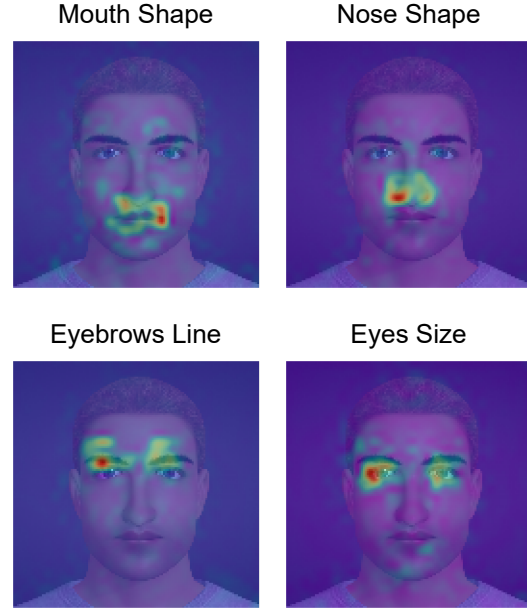


Figure 1. Example activation maps for four image features overlaid over the corresponding image.

3. Test Set Reconstructions

The dataset of joint image and gaze data we collected was split into a train, validation, and test set. The test set contains 36 images for all of which the following figures show the target faces, reconstructions synthesised by our method, baseline reconstructions as well as a colour-coded indication of the reconstruction quality of the different facial regions. The colour-coding was calculated by assigning each feature group to one of three equidistant bins according to their mean absolute slider distance (MASD).

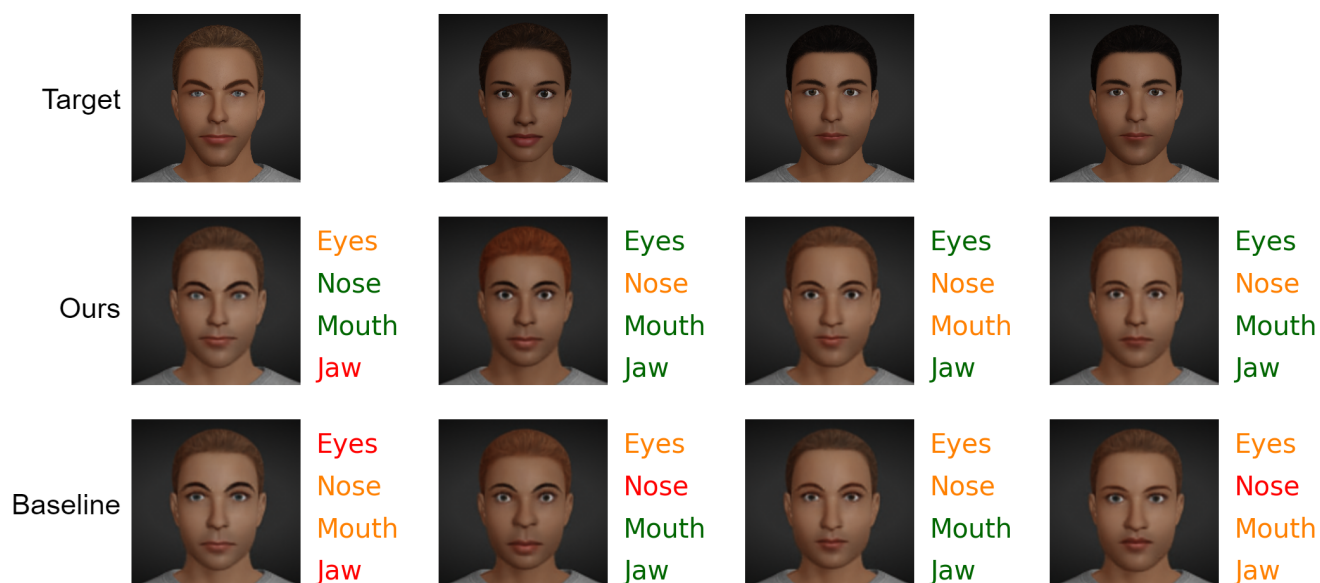
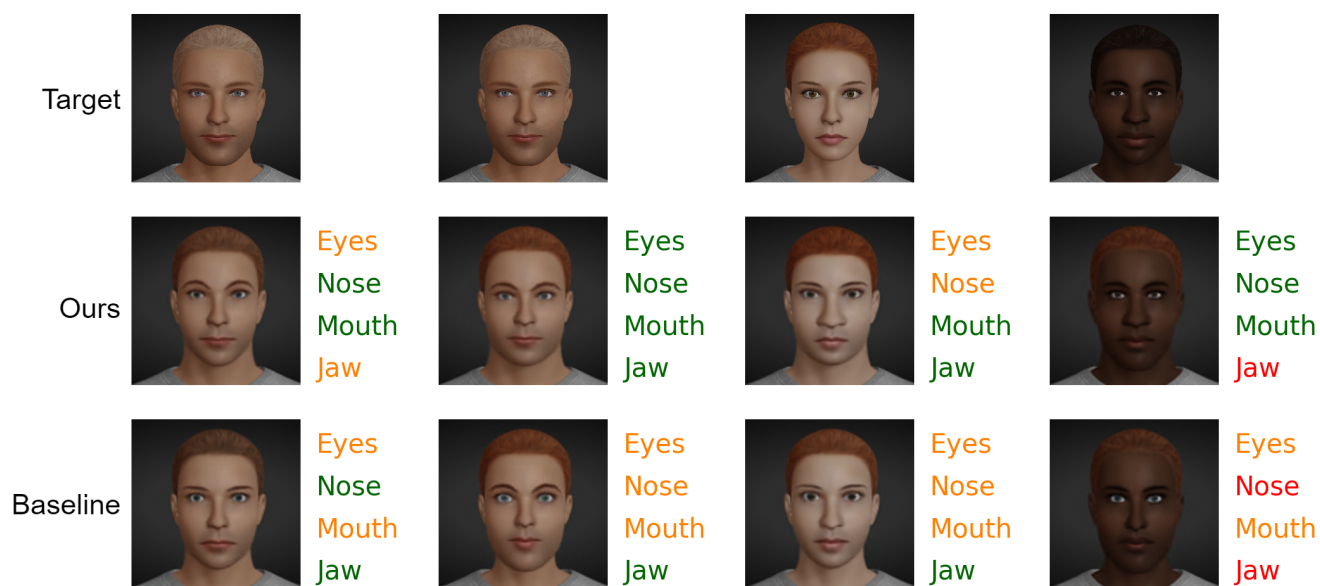
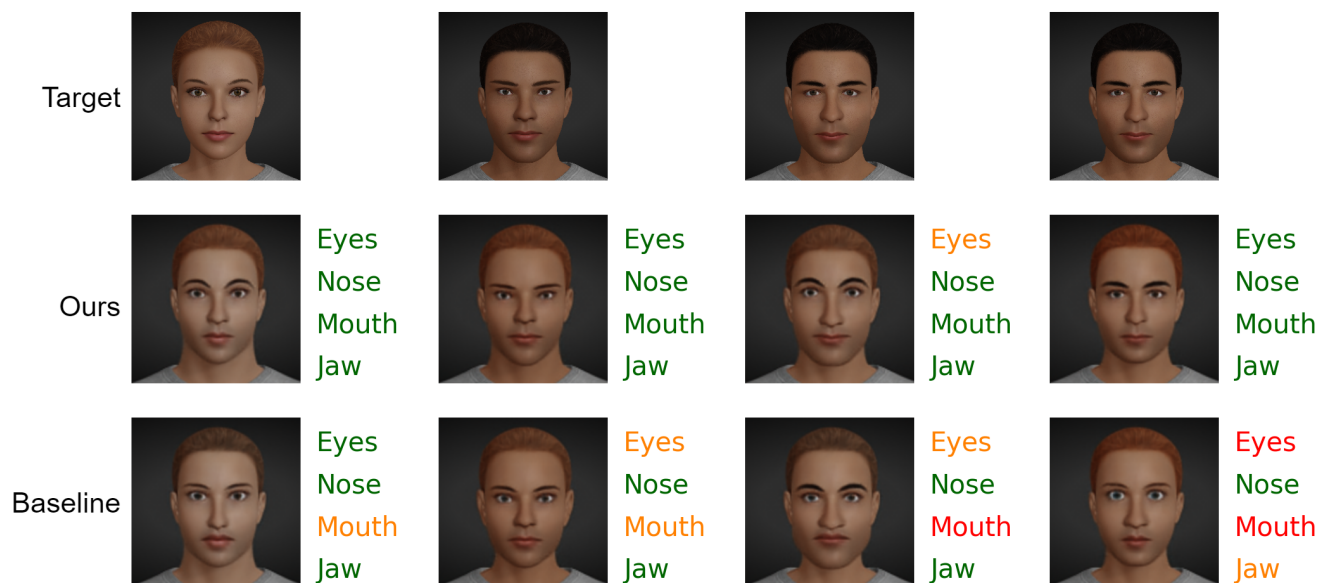
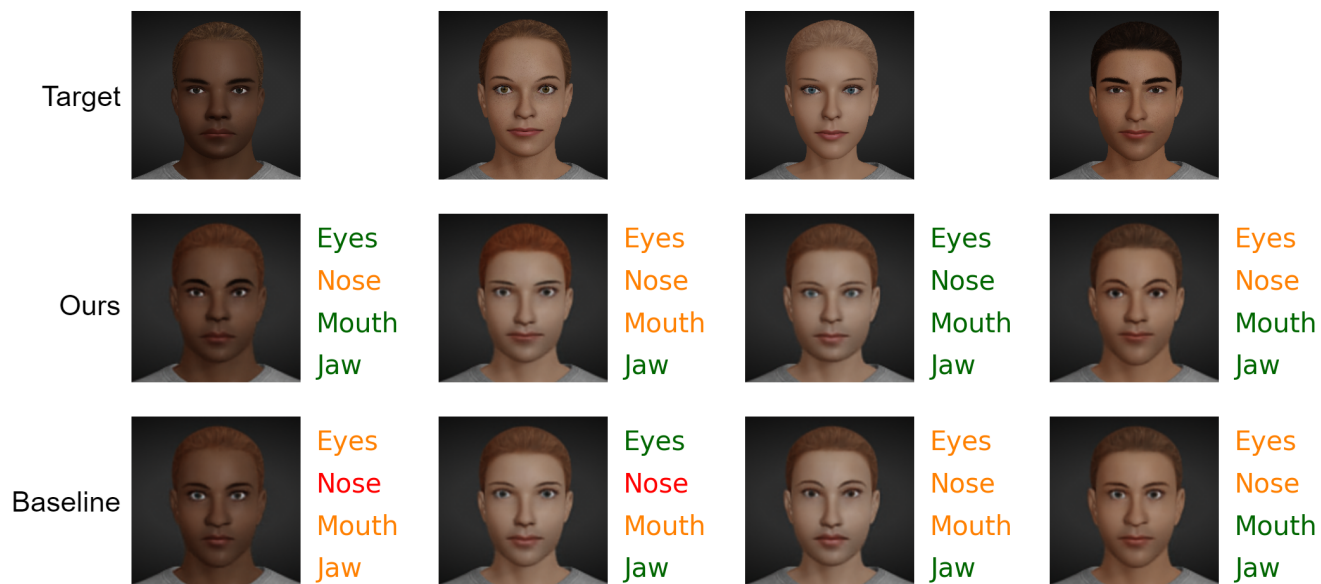
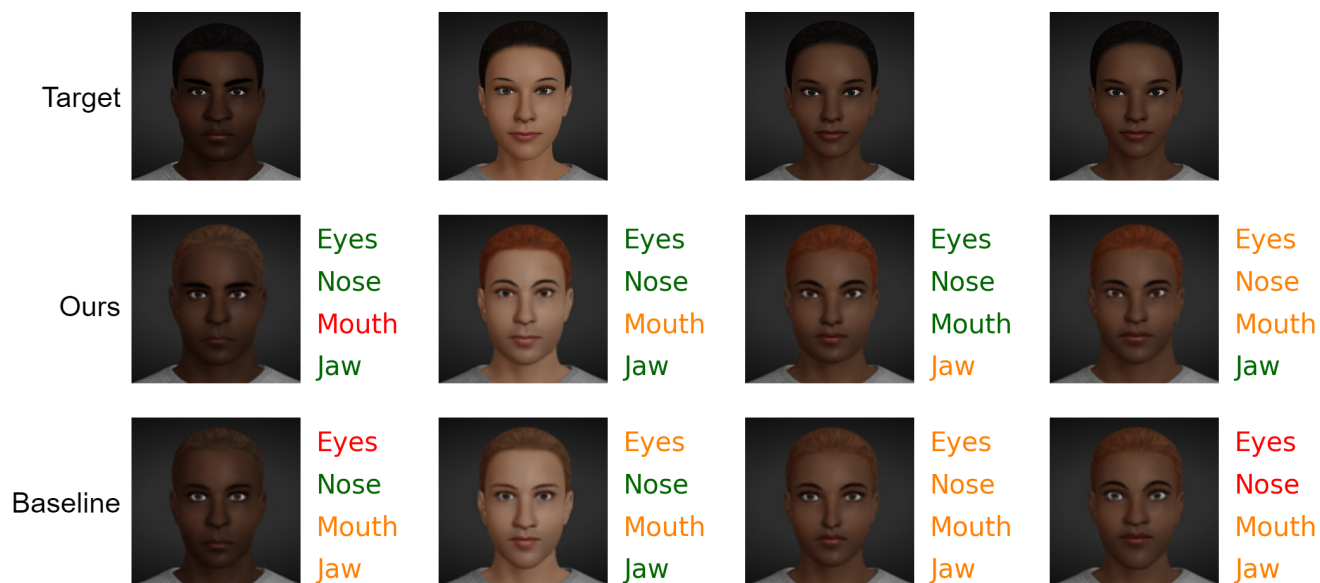
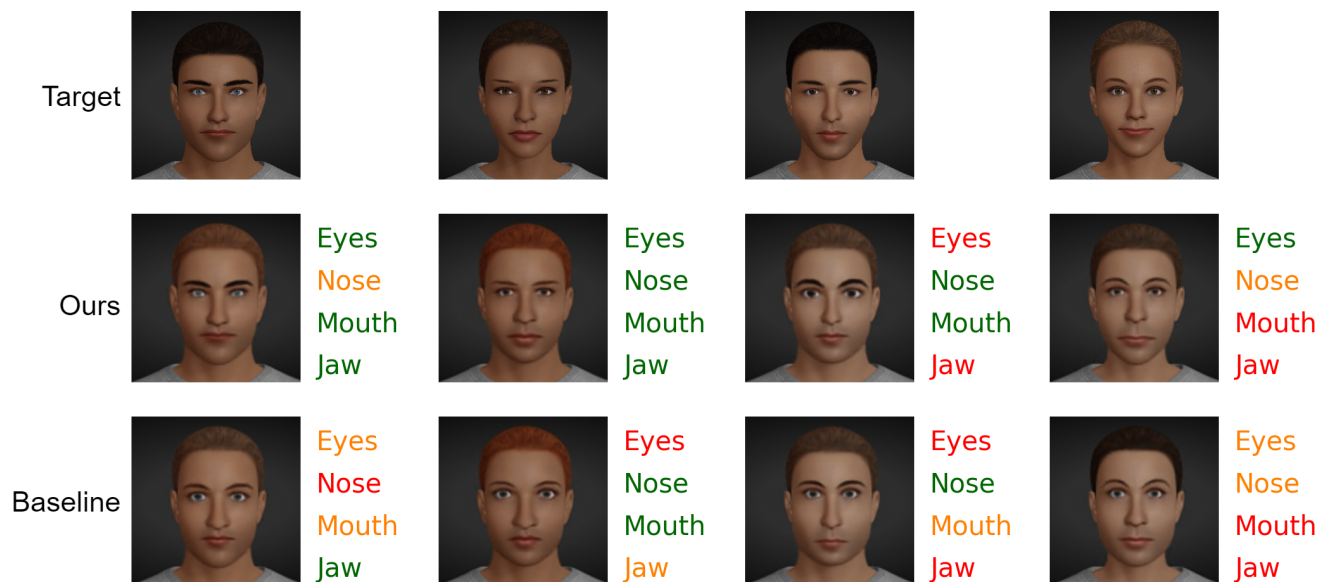
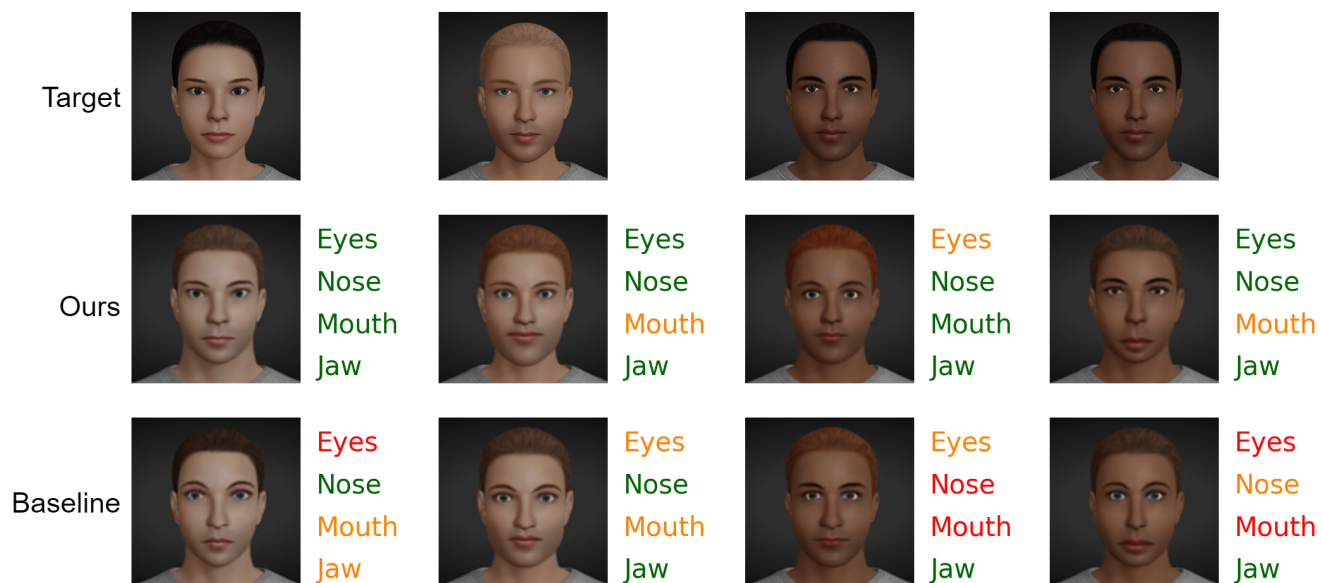
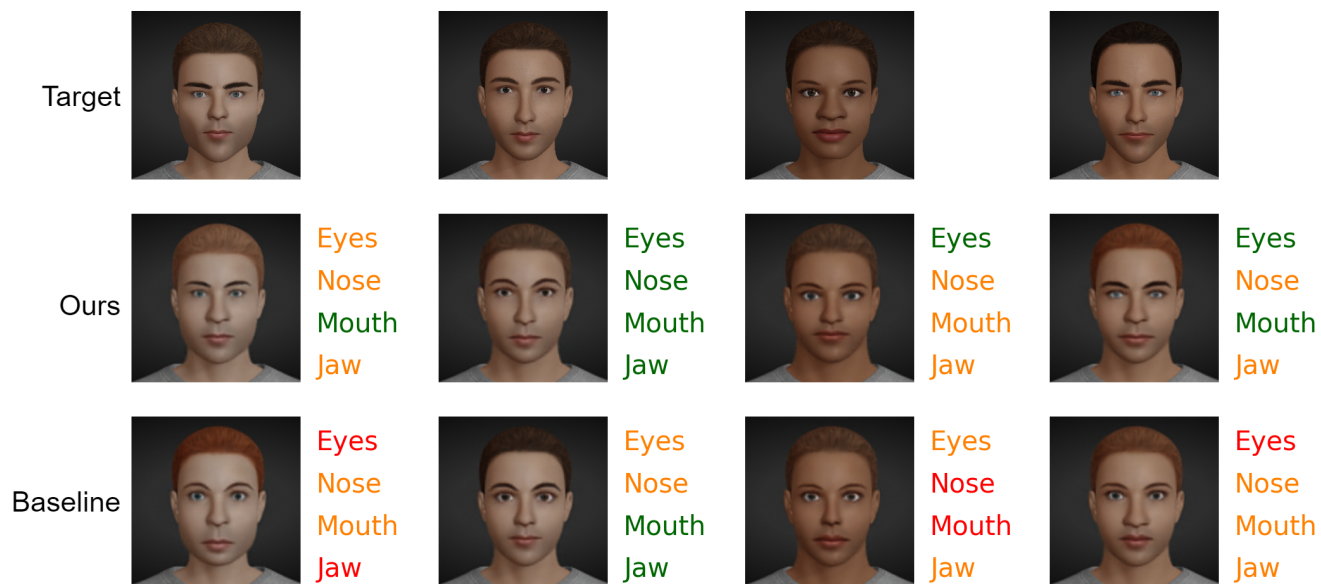


Figure 2. Test images reconstructed by our method (Ours) in comparison with the respective target and baseline. The colour-coded labels for the different facial regions indicate reconstruction quality (high, medium, or low). We calculated these by assigning each feature group to one of three equidistant bins according to their mean absolute slider distance (MASD). Columns A and B show two of the best, while C and D show two of our worst mental image reconstructions.









Target



Ours



Eyes
Nose
Mouth
Jaw



Eyes
Nose
Mouth
Jaw



Eyes
Nose
Mouth
Jaw



Eyes
Nose
Mouth
Jaw

Baseline



Eyes
Nose
Mouth
Jaw



Eyes
Nose
Mouth
Jaw



Eyes
Nose
Mouth
Jaw



Eyes
Nose
Mouth
Jaw

References

- [1] Valentin Schwind, Katrin Wolf, and Niels Henze. *FaceMaker - A Procedural Face Generator to Foster Character Design Research*, volume Game Dynamics: Best Practices in Procedural and Dynamic Game Content Generation, pages 95–113. Springer International Publishing, Cham, 2017. [1](#)