

# Supplementary Material for: Automatic Construction of Deformable Models In-The-Wild

Epameinondas Antonakos, Stefanos Zafeiriou  
Department of Computing, Imperial College London  
180 Queen’s Gate, SW7 2AZ, London, U.K.  
{e.antonakos, s.zafeiriou}@imperial.ac.uk

In the following sections, we supply additional qualitative experimental results. Section 1 shows the facial statistical shape model (Point Distribution Model) used in all experiments. Section 2 provides additional results regarding the convergence of the proposed algorithm. Finally, in Section 3 we report some qualitative fitting results that supplement the experiments of Sec. 3.3 of the original manuscript.

## 1. Facial Shape Model

As explained in Sec. 3.2 of the main manuscript, we construct our 2D facial statistical shape model by applying Principal Component Analysis on 50 annotated shapes from the LFPW database [1]. These shapes are selected in such a way that they demonstrate various facial poses and expressions. We employ the annotations provided by [3, 4], which utilize a facial mask with 68 landmark points. Figure 1 shows the four principal components of the resulting shape model. Note that the generation of the training shapes could also be achieved by deforming a 3D shape model (e.g. [2]) and projecting the examples to the 2D camera plane.

## 2. Convergence of AAM Automatic Construction

In Sec. 3.2 of the main paper, we present the convergence of the proposed method for automatic construction of a generative and a discriminative AAM. We apply the discriminative model once and report the convergence curves with respect to the cost function and point-to-point normalized RMSE per iteration. Figure 2 shows the evolution of the fitted shapes for eight images during the automatic building procedure. Starting from the bounding boxes (first row), the final result of the last generative model (last row) is very accurate. This figure also highlights the importance of the discriminative model. Even though the fitted shapes that it provides are not accurate, because its discriminative nature requires carefully annotated data, however, it manages to move the generative model’s shapes from the point where

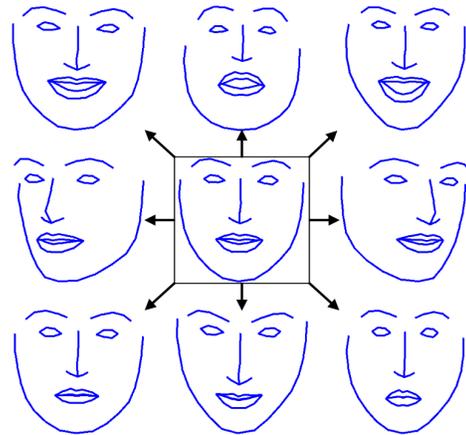


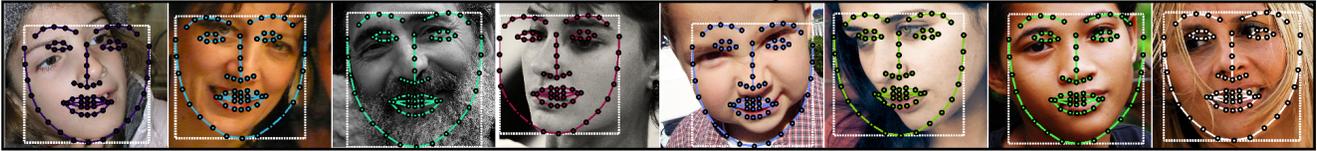
Figure 1: Four principal components of the employed 2D statistical shape model.

they stuck. We believe that the final fitted shapes shown at the last row of Fig. 2 are very impressive, given the automatic nature of the proposed method. Moreover, Fig. 3 shows the eight fitted shapes with the worst RMSE error, that were estimated automatically with the proposed procedure. As can be seen, even in the worst cases, the method provides decent shapes.

## 3. Fitting Results

In Sec. 3.3 of the main paper, we compare the performance between our automatically constructed generative and discriminative models and other models trained on carefully annotated data. Herein, we visualize some fitting results using the generative and discriminative AAMs trained both automatically and on manual annotations. Figures 4 and 5 show such results for the AFW dataset and the union of LFPW and Helen databases respectively. Again, we strongly believe that these results are very promising, especially considering the fact that our method’s models were constructed by starting with just a bounding box per face.

### Initialization from Bounding Box



### First Generative Model

*Iteration 1*



*Iteration 25*



*Iteration 50*



### Discriminative Model



### Final Generative Model

*Iteration 1*



*Iteration 25*



*Iteration 50*



Figure 2: Automatic construction of AAM with a single application of the discriminative model. The figures show the evolution of the fitted shapes for 8 images, starting from the bounding boxes. Each automatically trained generative model is performed for 50 iterations.

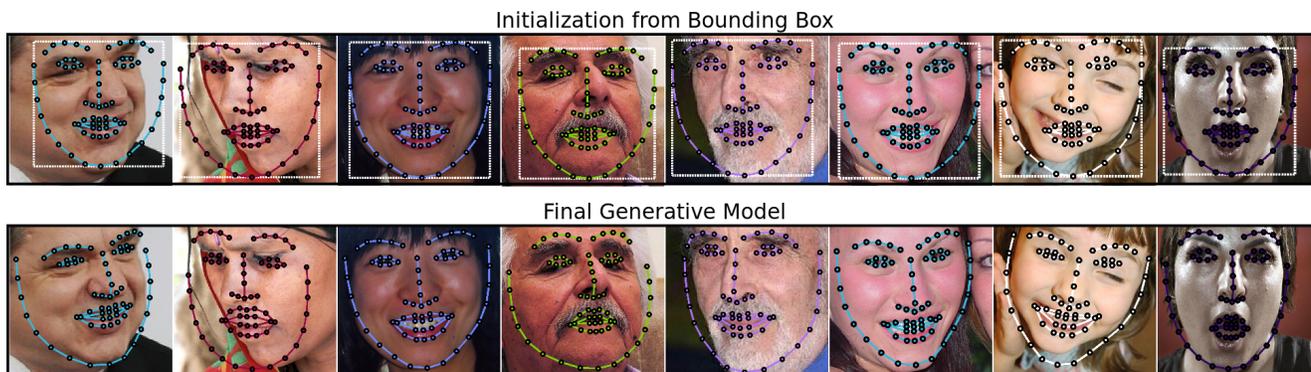
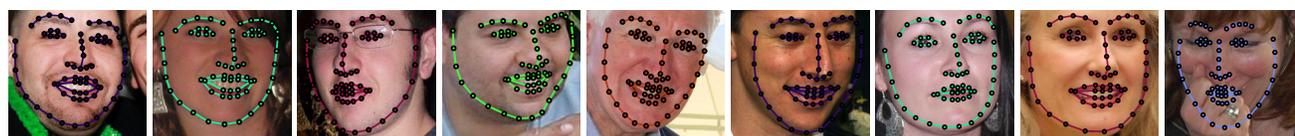


Figure 3: The 8 worst fitted shpes during the automatic construction of AAM with a single application of the discriminative model.



(a) Automatically trained generative model.



(b) Generative model trained on manual annotations.



(c) Automatically trained discriminative model.

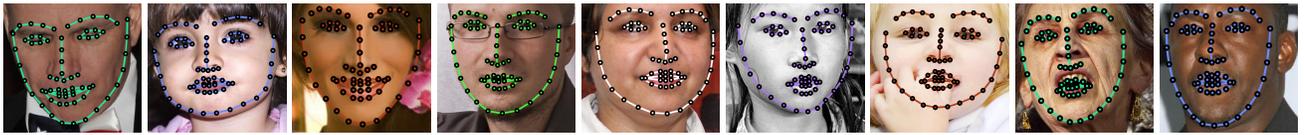


(d) Discriminative model trained on manual annotations.

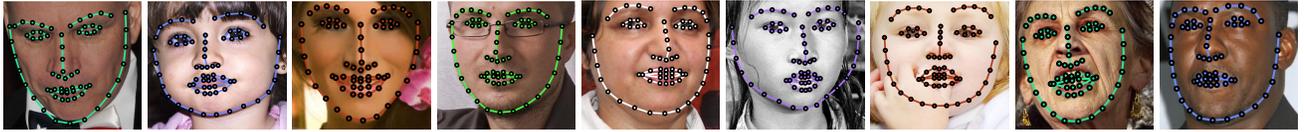
Figure 4: Fitting results on AFW database.

## References

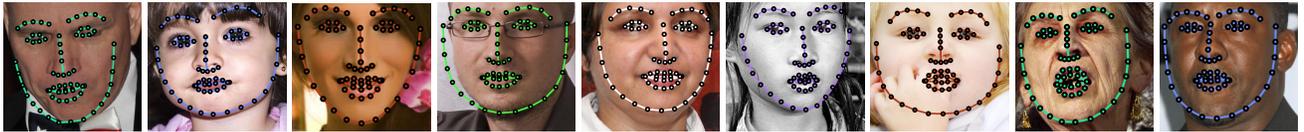
- [1] P. N. Belhumeur, D. W. Jacobs, D. J. Kriegman, and N. Kumar. Localizing parts of faces using a consensus of exemplars. In *IEEE CVPR*, 2011. 1
- [2] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter. A 3d face model for pose and illumination invariant face recognition. In *IEEE AVSS*, 2009. 1
- [3] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. 300 faces in-the-wild challenge: The first facial landmark localization challenge. In *IEEE ICCV'W*, 2013. 1
- [4] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. A semi-automatic methodology for facial landmark annotation. In *IEEE CVPR'W*, 2013. 1



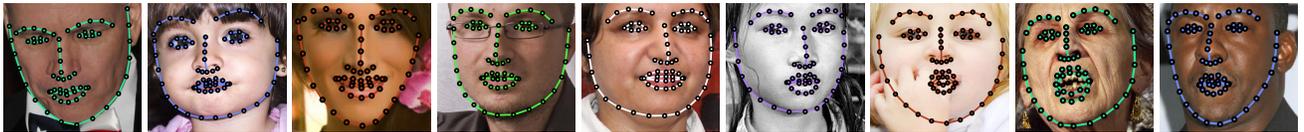
(a) Automatically trained generative model.



(b) Generative model trained on manual annotations.



(c) Automatically trained discriminative model.



(d) Discriminative model trained on manual annotations.

Figure 5: Fitting results on LFPW and Helen testing databases.