

Deep features for recognizing disguised faces in the wild

Shuang Sha

June 18, 2018

1. Learning Deep face Representations

Authors describe their approach to face verification and present results on the DFW challenge dataset. They build an ensemble of two deep CNNs for the task. The two networks are trained on the same dataset. An overview of a typical face verification pipeline is shown in figure 1.

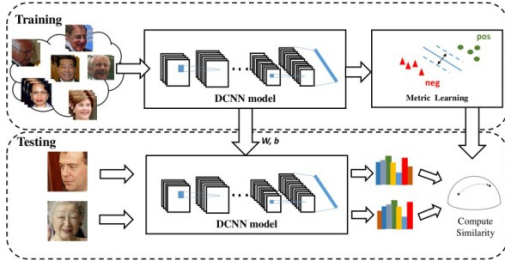


Figure 1. A typical face verification system pipeline.

1.1. Dataset

They used the Universe Face dataset, which is a combination of curated MS-Celeb-1M [3], UMDFaces [1], and frames from UMDFaces-Videos. They removed subject overlaps from the DFW dataset while creating the Universe Face dataset.

1.2. Loss Function

They gave a brief description of the L_2 -constrained softmax (L2SM) loss function used to train our networks. The aim behind L2SM is to improve the softmax loss to give high similarity scores for positive pairs and low similarity scores for negative pairs. L2SM adds an L_2 -constraint on the feature descriptor, forcing them to lie on a hypersphere of a fixed radius.

This has two advantages. First, it forces both good quality and bad quality images to have the same norm, unlike the softmax loss which gave good quality images a higher norm than poor quality images. This means that L2SM gives similar attention to both good and bad quality faces, unlike softmax. Second, L2SM forces the features from the same subject to be closer and features from different subjects to

be far from each other. Therefore, it maximizes the margin between positive and negative pairs.

1.3. Results

Firstly, researchers evaluated their approach on the relatively simple Disguised and Makeup Faces Database [4]. The images in this dataset are mostly celebrities with different disguises and makeups. This method achieves significant performance improvements over the baseline results reported in [4]. The method achieves a true accept rate (TAR) of 92% at a false accept rate (FAR) of 0.0001, and a TAR of 96.4% at FAR 0.001. This shows that our method can recognize people with make-up and disguises with high confidence [2].

Secondly, they evaluated their approach on the recently announced Disguised Faces in the Wild (DFW) challenge. The DFW challenge provides about 7,800 test images for about 600 identities containing both disguises and impersonations. Each identity in the test set contains a normal image, some validation images, a few images with the subject in disguise, and a few images of impersonators i.e. other people who look like the subject under consideration. The aim of this challenge is to recognize disguised faces as belonging to the subject under consideration and reject the impersonators. The evaluation criterion is a standard ROC curve which plots the True Acceptance Rate (TAR) against False Acceptance Rate (FAR).

Figure 2 shows the ROC curves for both our networks and for the final fused scores. Table 1 gives the TAR values at FAR = 0.01 and FAR = 0.001 for the three kinds of features.

Table 1. TAR (%) at different FAR values for the three different features used in this work.

Feature	FAR=0.001	FAR=0.001
ResNet-101	70.0	85.1
Inception ResNet-101	73.2	86.6
Fused	72.9	86.7

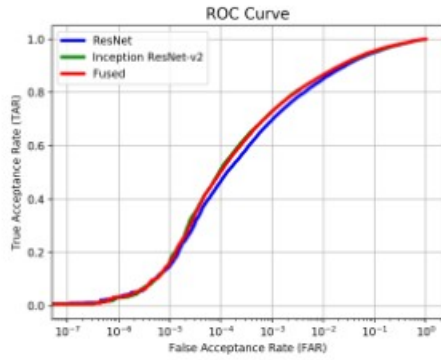


Figure 2. Results ROC.

References

- [1] A. Bansal, A. Nanduri, C. D. Castillo, R. Ranjan, and R. Chellappa. Umdfaces: An annotated face dataset for training deep networks. In *IJCB*, pages 464–473, 2017. [1](#)
- [2] A. Bansal, R. Ranjan, C. D. Castillo, and R. Chellappa. Deep features for recognizing disguised faces in the wild. In *CVPRW*, pages 10–16, 2018. [1](#)
- [3] Y. D. Guo, L. Zhang, Y. X. Hu, X. D. He, and J. F. Gao. MS-Celeb-1M: A dataset and benchmark for large-scale face recognition. In *ECCV*, pages 87–102, 2016. [1](#)
- [4] T. Y. Wang and A. Kumar. Recognizing human faces under disguise and makeup. In *ISBA*, pages 1–7, 2016. [1](#)