Supplementary materials - DeepVisage: Making face recognition simple yet with powerful generalization skills

Abul Hasnat[1], Julien Bohné[2], Jonathan Milgram[2], Stéphane Gentric[2], and Liming Chen[1]

[1]Laboratoire LIRIS, École centrale de Lyon, 69134 Ecully, France.
[2]Safran Identity & Security, 92130 Issy-les-Moulineaux, France.
md-abul.hasnat@ec-lyon.fr, {julien.bohne, jonathan.milgram, stephane.gentric}@safrangroup.com, liming.chen@ec-lyon.fr

1. Additional analysis

1.1. Error analysis

Now, we provide examples of the error cases observed from the face verification experiments on different datasets, such as LFW [2], IJB-A [3], YouTube Faces [4] and CACD-VS [1].

LFW [2]: Figure 1 provides the examples of the failure cases on the LFW [2] benchmark. The ratio of false accept vs reject is ***1:1.56***. Note that our method achieves 99.62% accuracy on LFW. In Figure 1(b) three pairs are marked with red colored rectangles. These pairs are erroneously labeled in the dataset, which means our method makes correct judgment on them and hence the accuracy further increases to 99.67% by considering them as correct match.

CACD-VS [1]: Figure 2 provides the examples of the failure cases on the CACD [2] dataset. The ratio of false accept vs reject is ***1:6***. Note that our method achieves 99.13% accuracy on CACD.

YTF [4]: Figure 3 provides few examples of the failure cases on the YTF [4] dataset. The ratio of false accept vs reject is ***1:2.2***. In Figure 3, we only show top three mistakes (sorted based on their similarity score) in terms of false accept and reject. Note that our method achieves 96.24% accuracy on YTF.

IJB-A [3]: Figure 4 provides few examples of the failure cases on the IJB-A [3] dataset. The ratio of false accept vs reject is ***1:5.15***. In Figure 4, we only show top three mistakes (sorted based on their similarity score) in terms of false accept and reject. Note that our method achieves following results on IJB-A: 0.887 at TAR@FAR=0.01% and 0.824 at TAR@FAR=0.001%. From the falsely rejected template pairs, we observe that: (a) one pair has only one image in the template; (b) the pre-processor fails to detect face as well as landmarks and (c) the images in the template have very high pose and large occlusion which causes important face attributes to be absent.

References

[1] Bor-Chun Chen, Chu-Song Chen, and Winston H Hsu. Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset. *IEEE Trans. on Multimedia*, 17(6):804–815, 2015.

[2] Gary B Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical report, Technical Report 07-49, University of Massachusetts, Amherst, 2007.

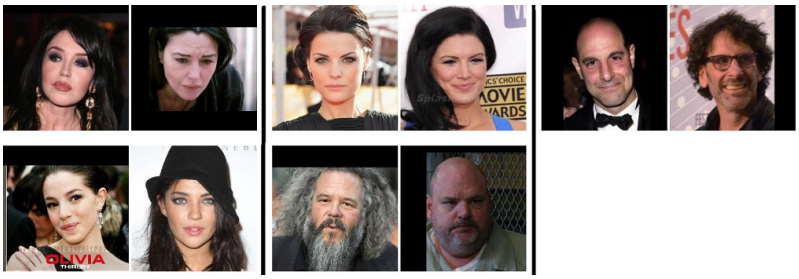
[3] Brendan F Klare, Ben Klein, Emma Taborsky, Austin Blanton, Jordan Cheney, Kristen Allen, Patrick Grother, Alan Mah, Mark Burge, and Anil K Jain. Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. In *Proc. of IEEE CVPR*, pages 1931–1939, 2015.

[4] Lior Wolf, Tal Hassner, and Itay Maoz. Face recognition in unconstrained videos with matched background similarity. In *Proc. of IEEE CVPR*, pages 529–534, 2011.

(a)

(b)

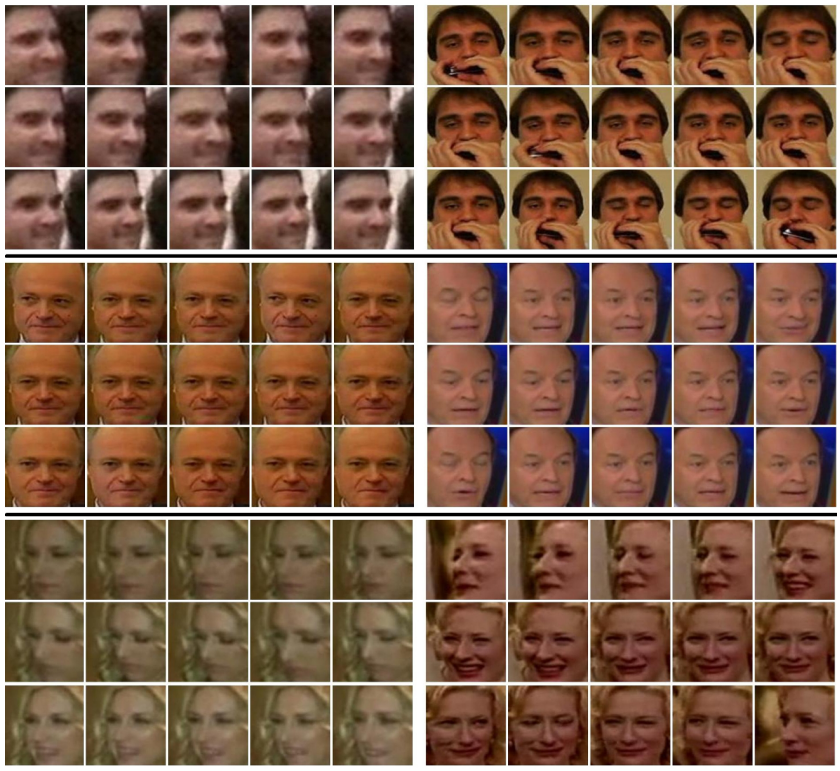
Figure 1: Illustration of the false accepted/rejected image pairs from the LFW [2] benchmark. (a) false accepted pairs and (b) false rejected pairs. The red colored rectangles indicate the examples which were erroneously labeled in the dataset.



(a)

(b)

Figure 2: Illustration of the false accepted/rejected image pairs from the CACD-VS [1] dataset. (a) false accepted pairs and (b) false rejected pairs.



(a)

(b)

Figure 3: Illustration of the false accepted/rejected video pairs from the YTF [4] dataset. (a) false accepted pairs and (b) false rejected pairs.

(a) (b)

Figure 4: Illustration of the false accepted/rejected template pairs from the IJB-A [3] dataset. (a) false accepted pairs and (b) false rejected pairs.