A Pipeline to Improve Face Recognition Datasets and Applications

IVCNZ 2018, 19-21 November 2018

Possible scenarios for Face Recognition

- Security issues
- HR management in companies
- Hi-Tech applications
- Taking roll at university

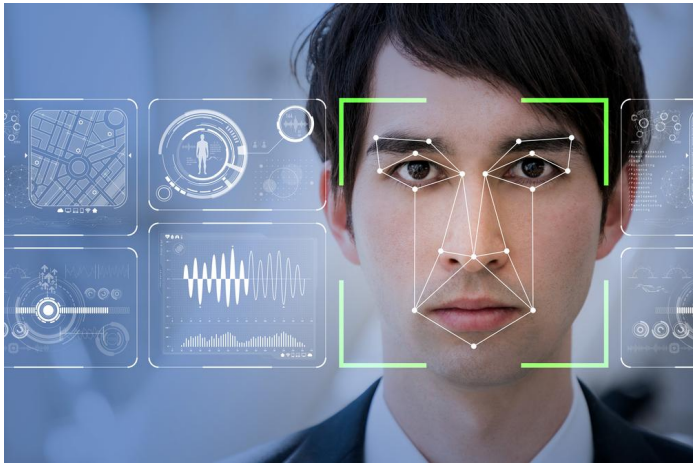
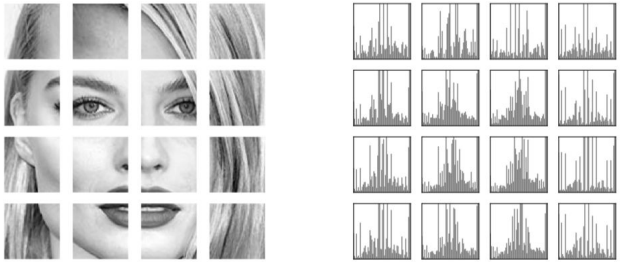


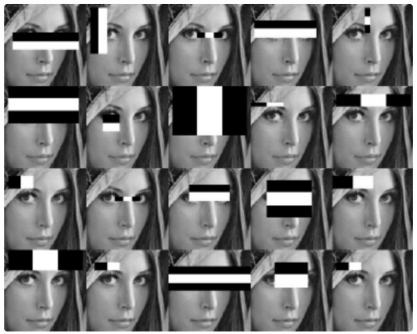
Image Classification before Deep Learning

1. Supervised Dataset: (X, Y)
2. Feature Extraction (HOG, SIFT, …)
3. Feature Selection
4. Training a model (SVM, MLP,...)



(a) (b)

Fig. 5: (a) Face image divided into 4×4 local regions. (b) Histograms of LBP descriptors computed from each local region.



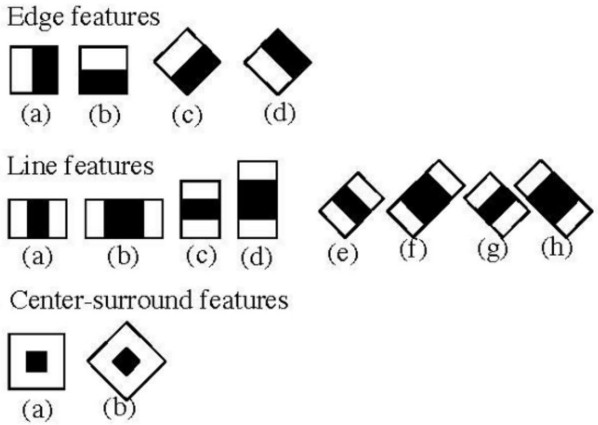
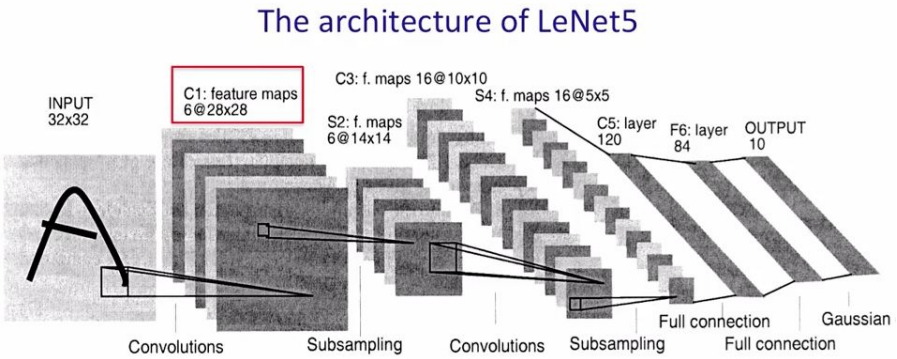
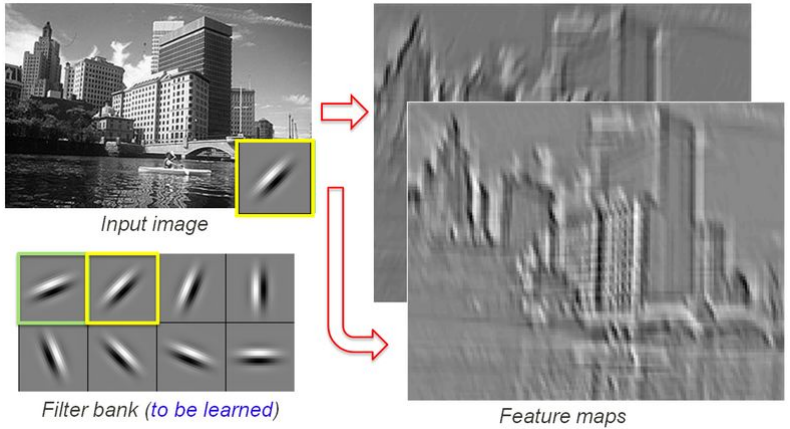


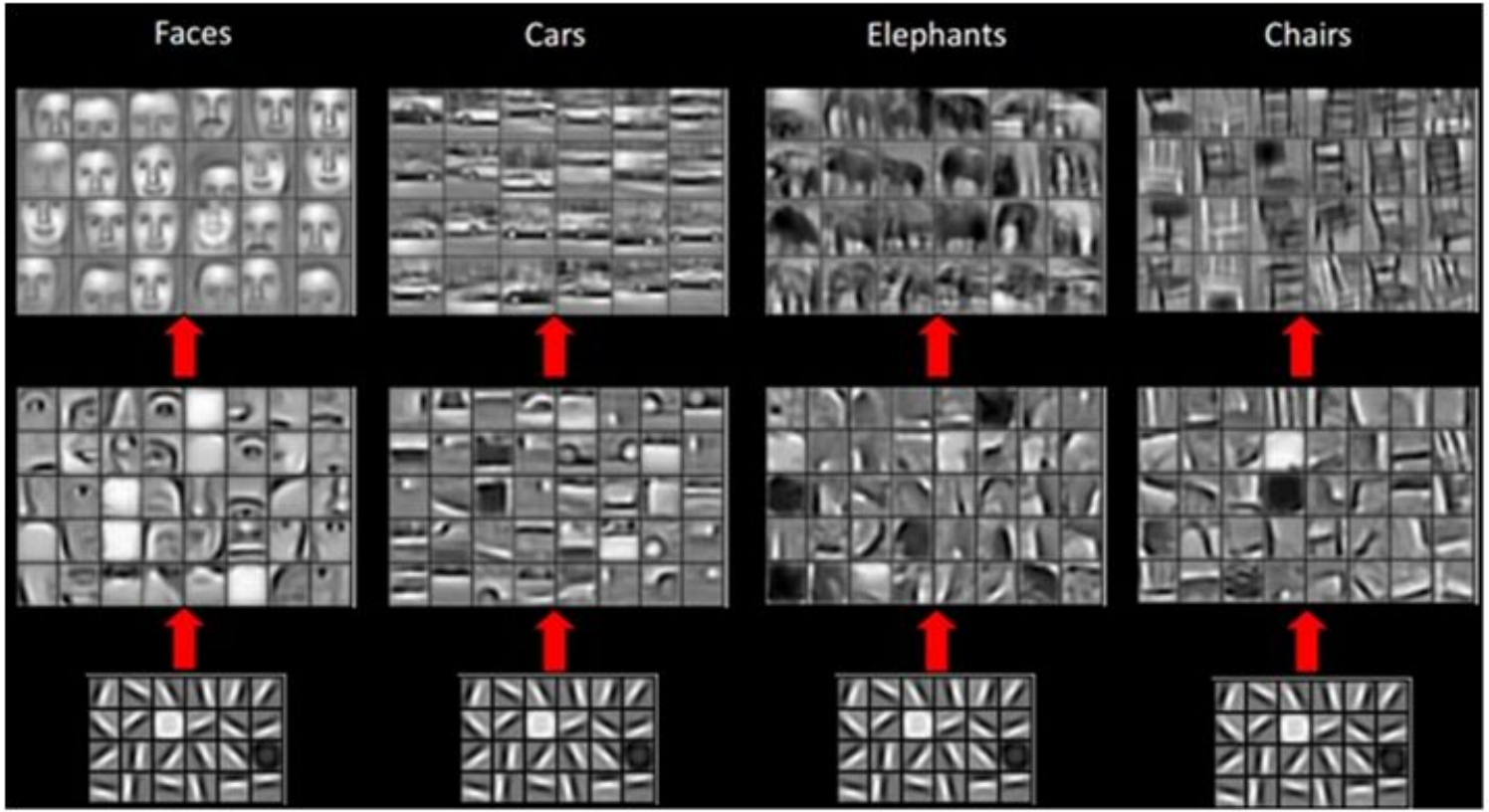
Image Classification with Deep Learning

1. Supervised Dataset: (X, Y)
2. Training a model (CNN)





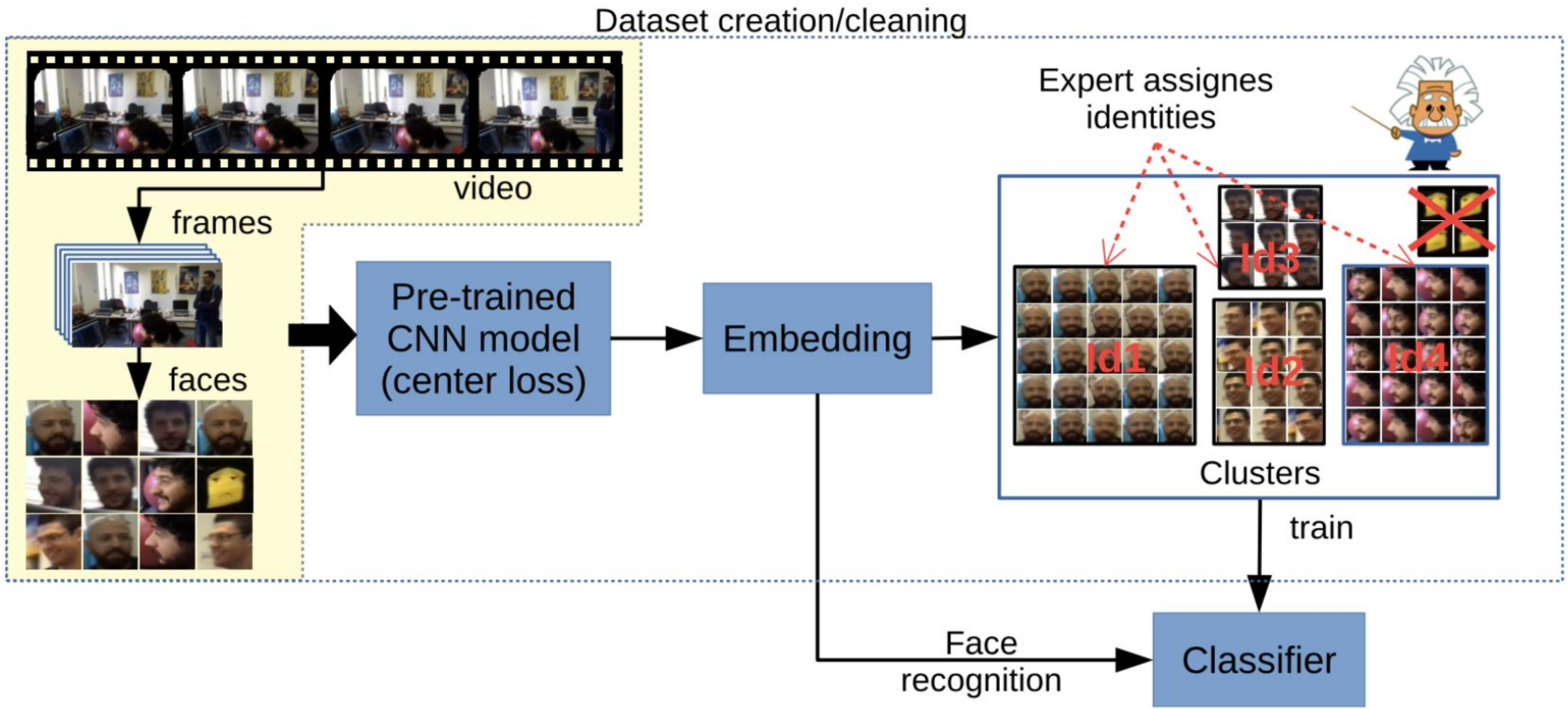
CNN Feature Maps



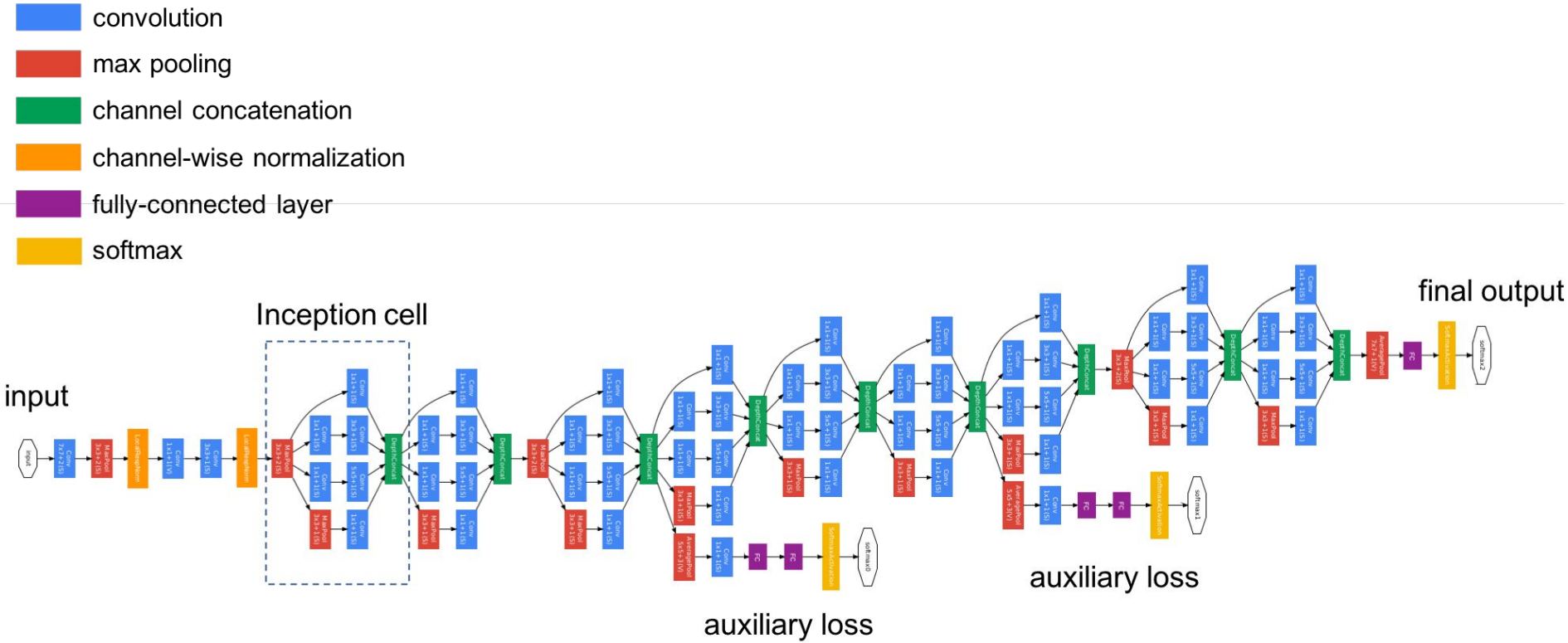
Objectives

- Dataset Creation or Cleaning
- Pipeline for Face Recognition

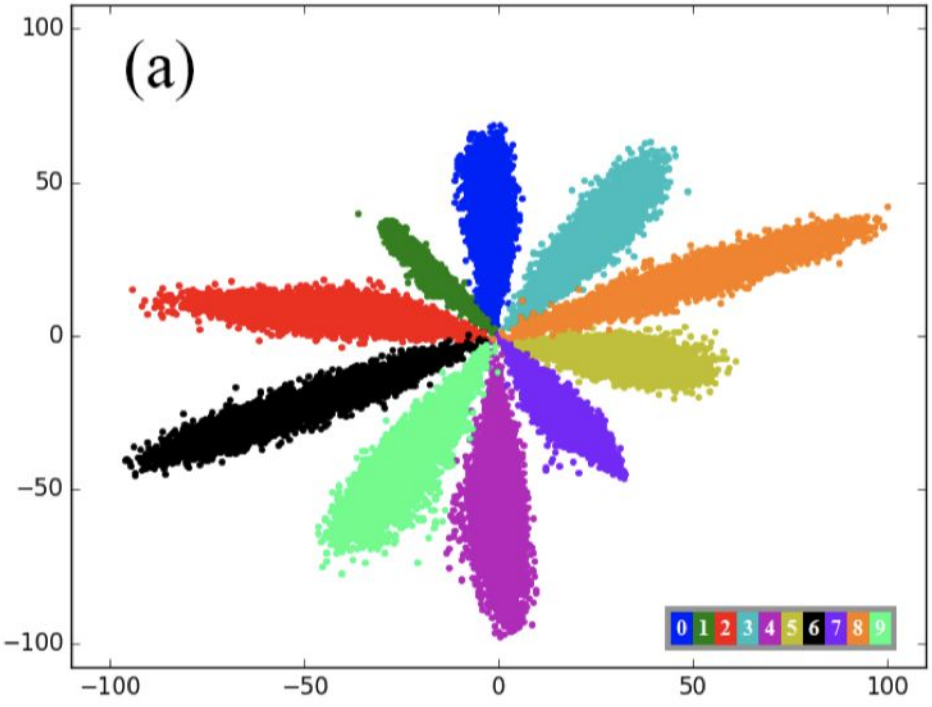
The Proposed Pipeline

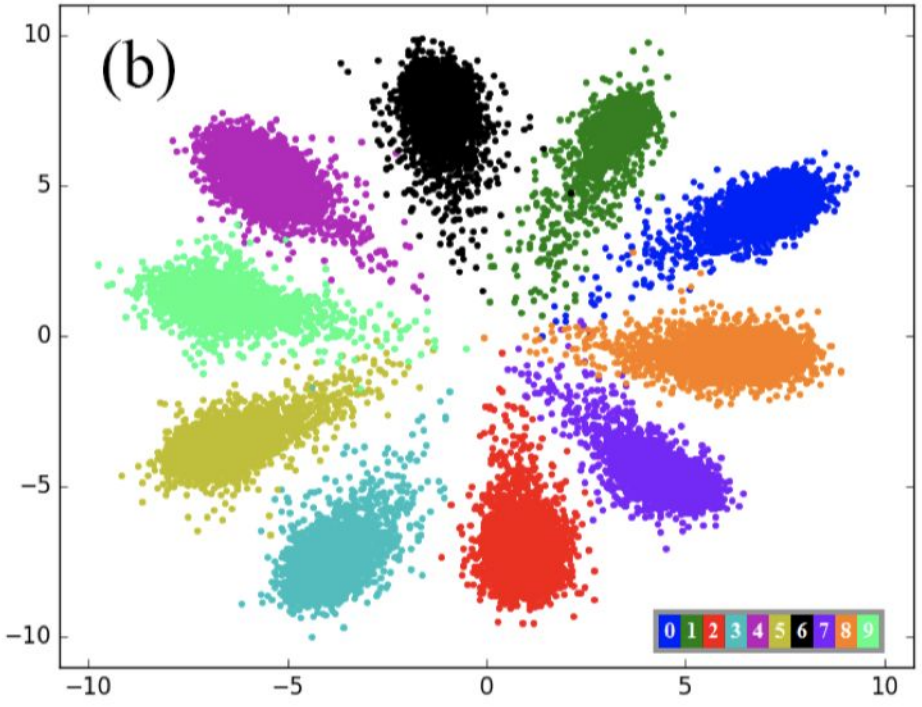


Inception ResNet-v1 model - GoogleNet (2012)



Softmax vs Center loss





Datasets

Datasets used in this paper are:

Still images

1. VGGFace2, 3.3M faces of 9000+ identities
2. MS-Celeb-1M, 10M faces of 100k identities
3. UMDFaces, 367k faces of 8.3k identities
4. Casia WebFace, 494k images of 10k identities

Video images

5. YouTube Faces, 3.4k videos of 1595 identities
6. 7Pixel dataset, 25 videos of 25 identities.

Datasets - Samples (1/3)

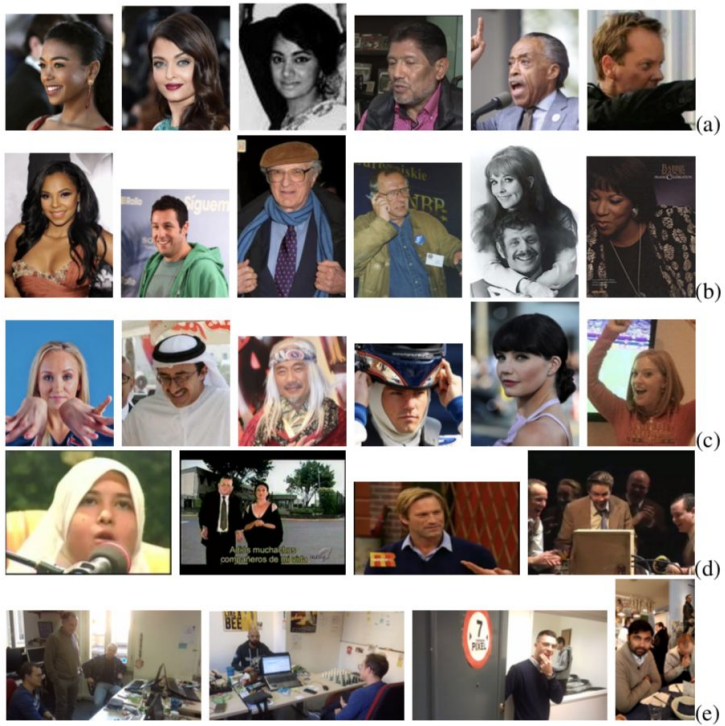


Fig. 3. In each row some examples of representative images/frames of datasets used in this paper: VGGFace2 (a), UMDFaces (b), MS-Celeb-1M (c), YouTube Faces (d) and 7Pixel-Face (e).

Datasets - Samples (2/3)

Fig. 2. Some aligned faces of a single cluster obtained merging 3 different videos from YouTube. In each video of Anthony Hopkins we have different environmental settings, qualities and lightings.

Datasets - Samples (3/3)

Fig. 4. In each line some problematic examples of faces extracted from videos that contain more than one identity of the 7Pixel-Face dataset. Blur and motion problems, in some cases, lead to clusters of faces with errors: the top row and the bottom row represent some errors found in two different clusters.

Face Recognition Classification process

- Given an image, we extract 128-dimensional embedding through an Inception ResNet-v1 model
- Embedding Classification with SVM

Results - Face Recognition (1/3)

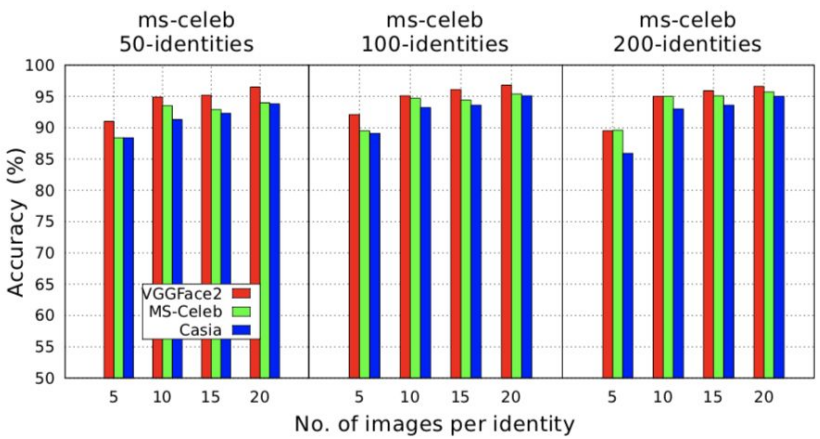


Fig. 5. Face recognition accuracy obtained from a trained SVM on the MS-Celeb-1M dataset varying the number of identities and the number of images per identity. The SVM receives embeddings obtained from a CNN trained on a Center loss function with VGGFace2, MS-Celeb and Casia datasets.

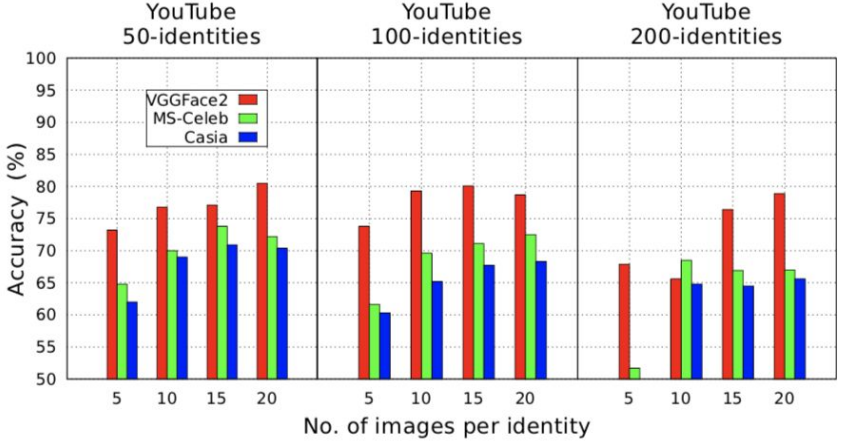


Fig. 6. Face recognition accuracy obtained from a trained SVM on the YouTube Faces dataset varying the number of identities and the number of images per identity. The SVM receives embeddings obtained from a CNN trained on a Center loss function with VGGFace2, MS-Celeb and Casia datasets.

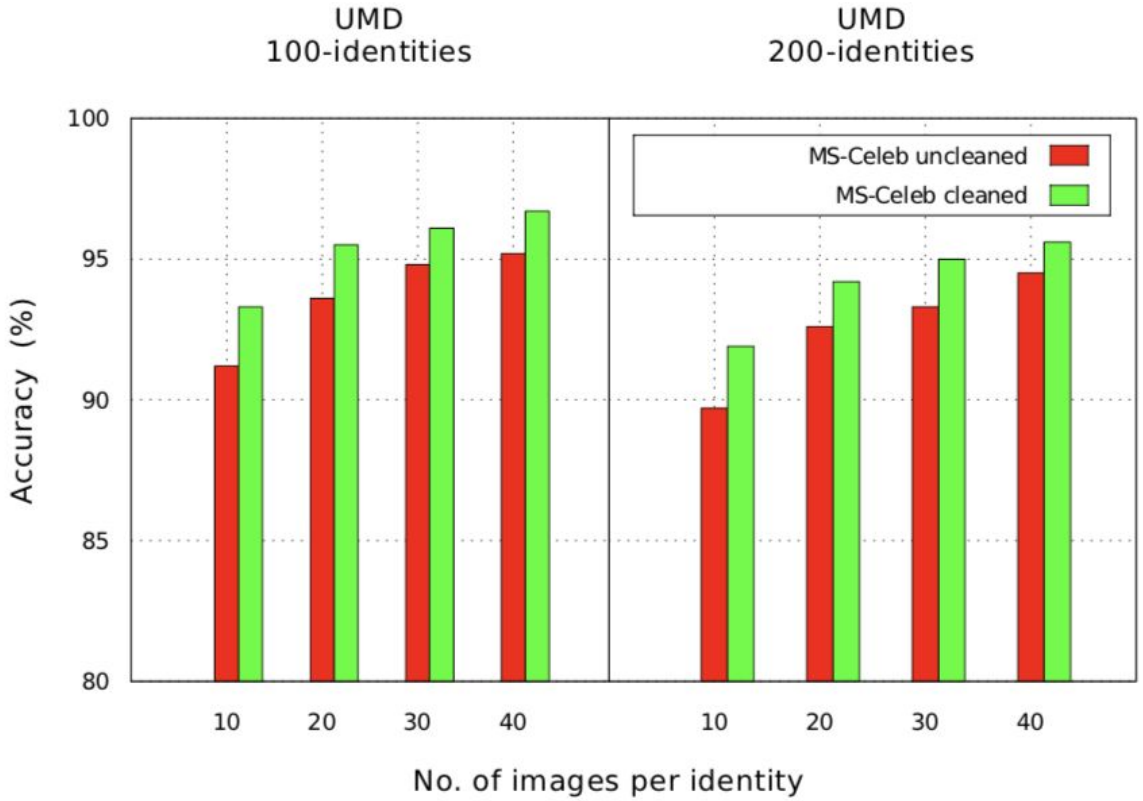
Results - Clustering (2/3)

TABLE I

ACCURACY (ACC), PRECISION (P) AND RECALL (R) OF THE CLEANING PROCESS APPLIED TO 50 RANDOMLY SELECTED IDENTITIES OF THE MS-CELEB DATASET. POSITIVE IMAGES BELONG TO A SELECTED IDENTITY, WHILE NEGATIVE ARE ALL REMAINING IMAGES.

	Positive	Negative	Acc=97.35%
Biggest cluster	2617	86	P=99.32%
Other clusters	18	1206	R=96.82%

Results - Clean vs Unclean (3/3)



Conclusion

- We proposed a semi-supervised solution based on a CNN model with Center loss, that speeds-up the faces labeling process from a video containing a set of identities.
  - A pipeline to create a dataset from scratch or cleaning an existing one
  - Face Recognition application

References

- Ignazio Gallo, Shah Nawaz, Alessandro Calefati and Gabriele Piccoli, A Pipeline to Improve Face Recognition Datasets and Applications, 2018 International Conference on Image and Vision Computing New Zealand (IVCNZ).
- Yandong Wen, Kaipeng Zhang, Zhifeng Li and Yu Qiao, A discriminative feature learning approach for deep face recognition, 2016 European conference on computer vision (ECCV).
- https://www.quora.com/How-was-facial-recognition-performed-before-without-deep-learning
- https://en.wikipedia.org/wiki/Softmax_function
- https://engmrk.com/lenet-5-a-classic-cnn-architecture/

Questions?