

OpenFace

A general-purpose face recognition library with
mobile applications

M2019170
허성실



0

Abstract

- Cameras are being ubiquitous and can use face recognition technology
- But, large accuracy gap between today's publicly available face recognition systems and the state-of-the-art private face recognition systems.
- Provides near-human accuracy on The LFW benchmark and present a new classification benchmark for mobile scenarios



1

Introduction

- Recognizing people is a **primitive operation** in mobile computing
-> cognitive assistance, social events, speaker annotation in meetings
- Training **new** classification models now become important

That model Is **OpenFace**

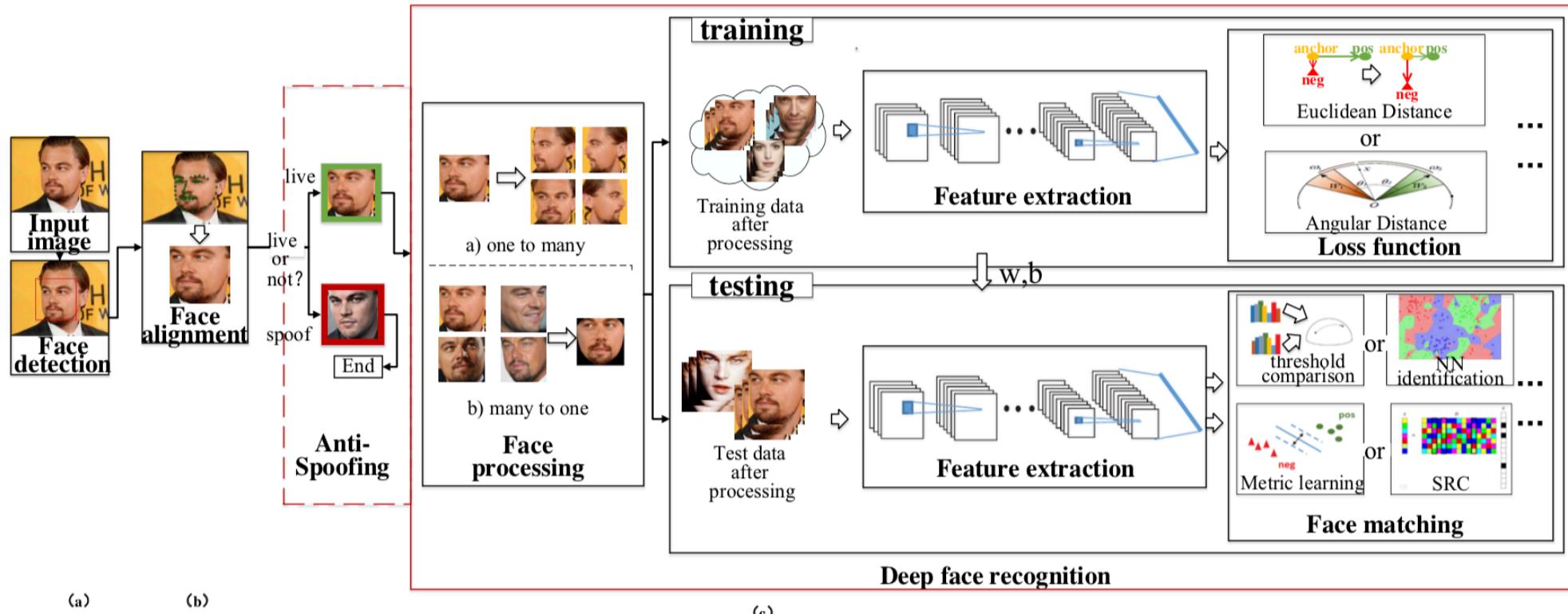


2

Background & Related Work

2.1. Face Recognition

- Face Recognition System



1. Face detection

2. Face processing

3. Face recognition

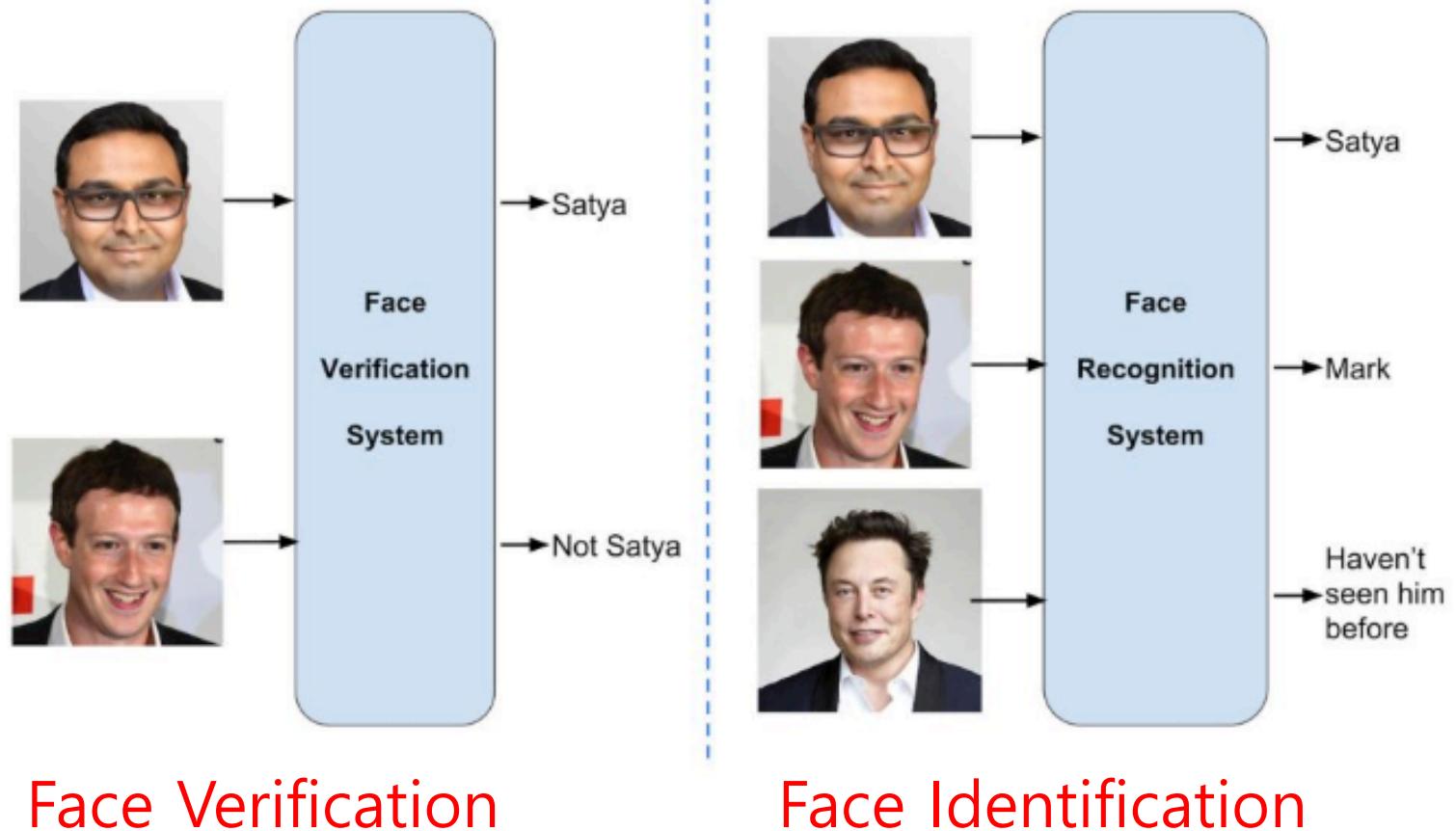
Face Recognition

The Gallery: Target IDs

The Problem: Test ID

Verification: One-to-one

Identification: One-to-many

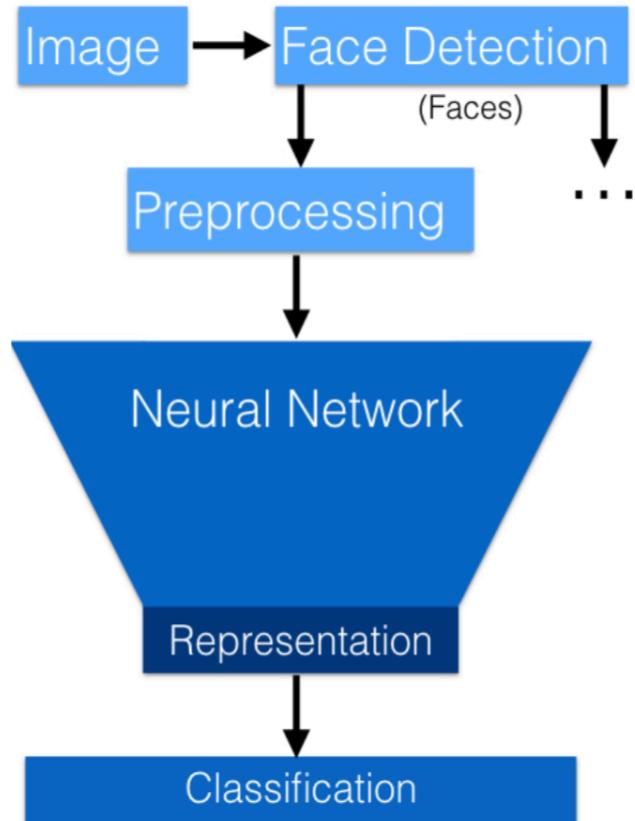


출처: <https://www.learnopencv.com/face-recognition-an-introduction-for-beginners/>

Face Recognition vs Object Recognition

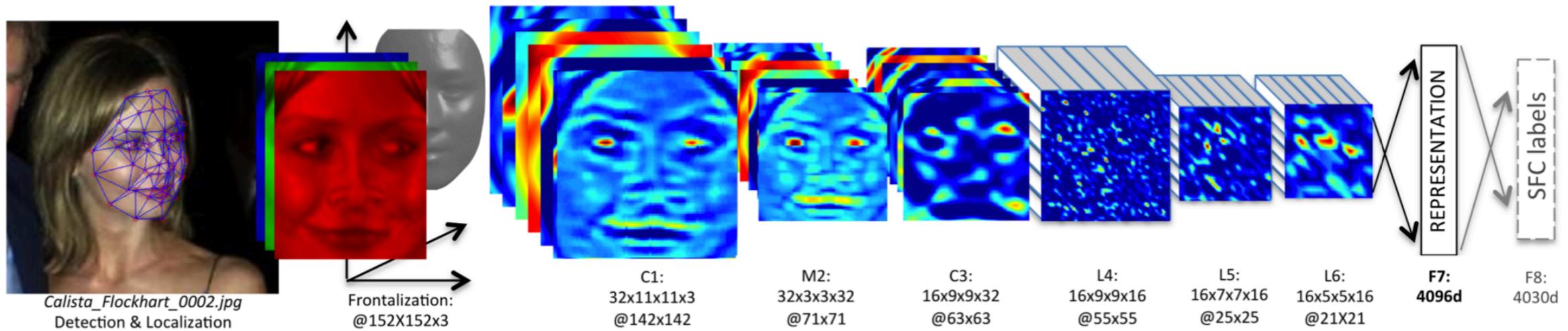
- Intra-variations(pose, age) could be larger than inter-differences
- Low resolution
- The softmax loss is not sufficiently effective for FR

- 2.2 Face Recognition with Neural Networks
(DeepFace, FaceNet etc...)



- Face detection is another active research topic
- **Low-dimensional** representation is key
-> efficiently used in classifiers or clustering techniques

2.2.1 Deepface



DeepFace's innovations comes from:

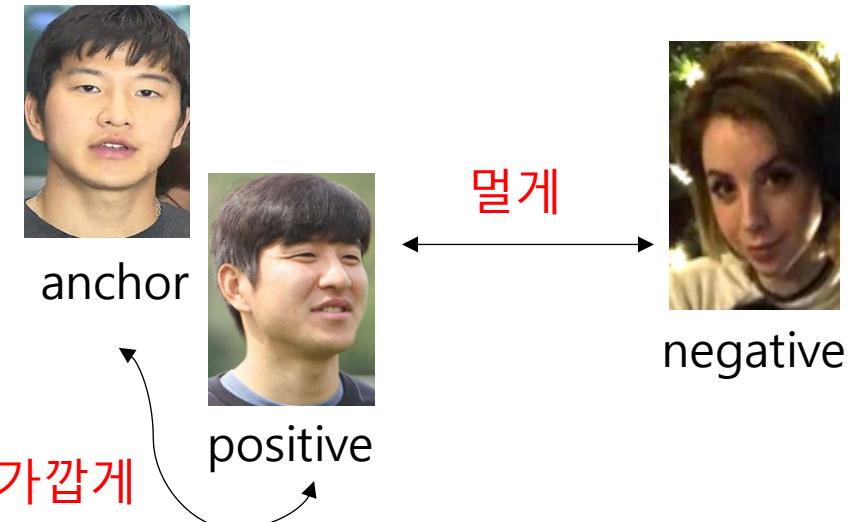
- (a) The **3D alignment**
- (b) A neural network structure with **120 million parameters**
- (c) Training with **4.4 million** labeled faces

But, the representation. Is difficult to use. Because faces. Of the same person **aren't necessarily clustered**

2.2.2 FaceNet

Triplet Loss 사용

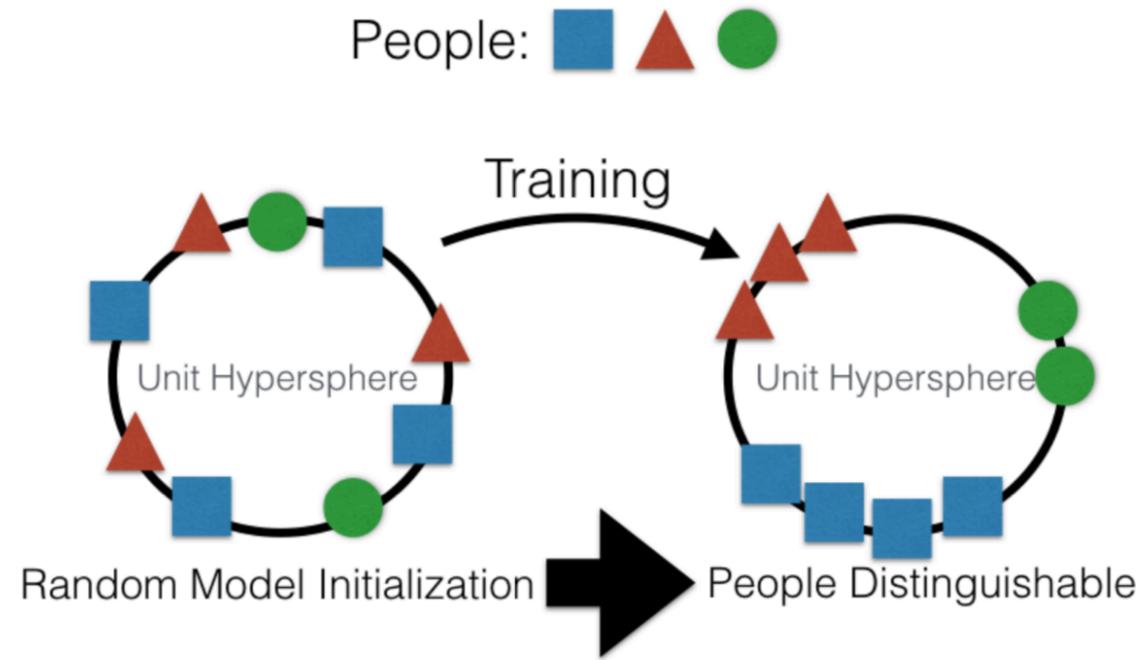
- 같으면 가깝게, 다르면 멀게



$$\|f(x_i^a) - f(x_i^p)\|_2^2 + \alpha < \|f(x_i^a) - f(x_i^n)\|_2^2 ,$$

$$\forall (f(x_i^a), f(x_i^p), f(x_i^n)) \in \mathcal{T} .$$

2.2.2 FaceNet



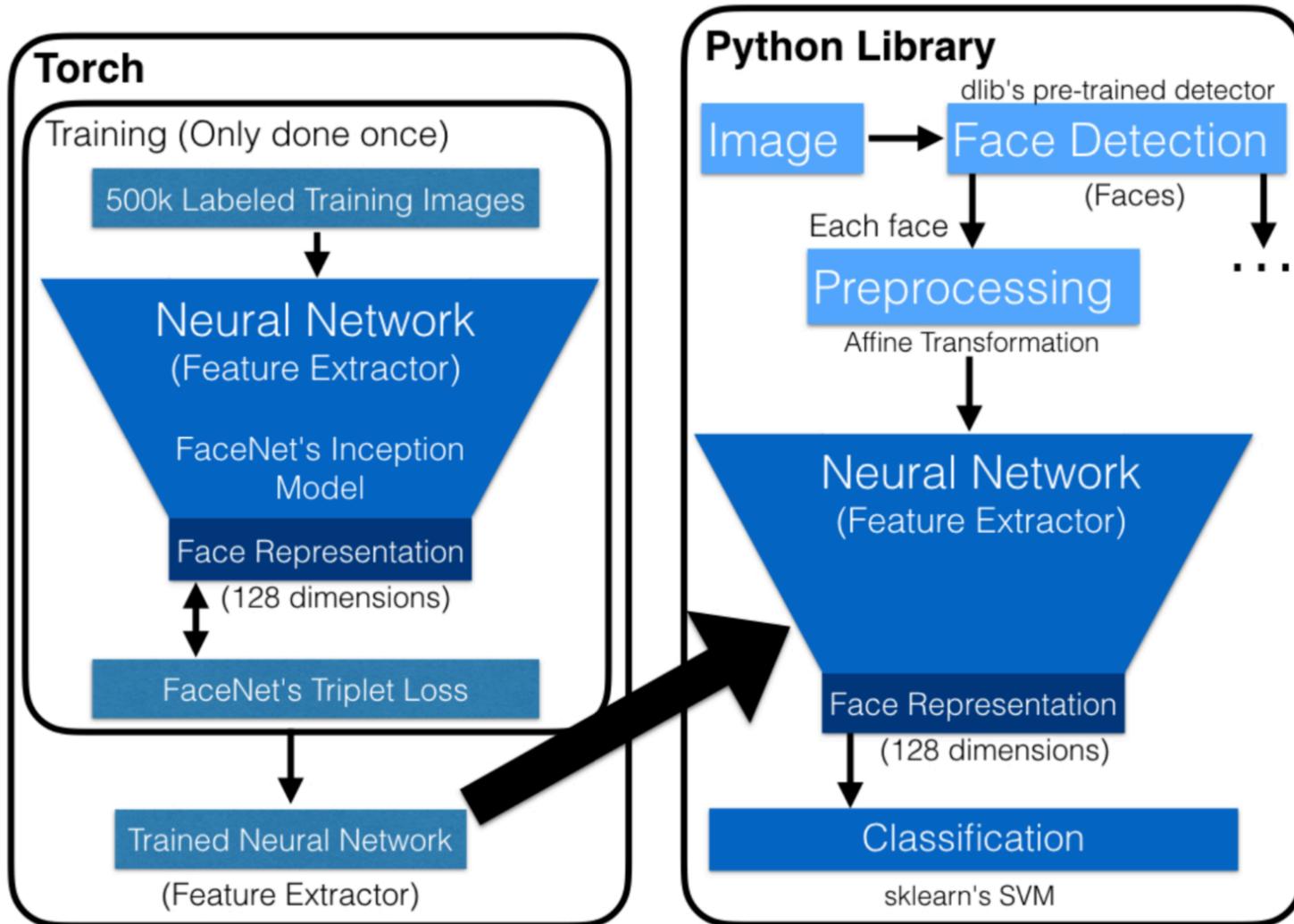
- FaceNet's innovation comes from:
 - (a) The **triplet loss**
 - (b) Their **triplet selection** procedure
 - (c) Training with **100 million to 200 million** labeled images

- 2.3 Face Recognition in Mobile Computing
- Due to lack of availability, it has less accuracy than the state-of-the-art
- Used Eigenface-based face recognition system
- Has been a rise of efficient GPU architectures for mobile devices



3

Design & Implementation



- Our key design consideration is a system that gives **high accuracy with low training and prediction times**

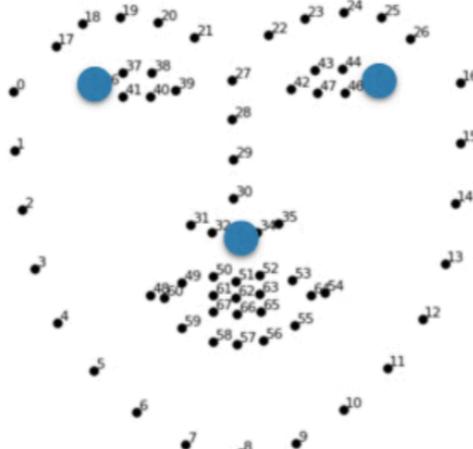
Figure 4: OpenFace's project structure.

- 3.1 Preprocessing: Alignment with an Affine Transformation
- The face detection portion returns a **list of bounding boxes** around the faces
- They are all **different pose and different illumination conditions**
- FaceNet handle this with a large training dataset
- But it can't be with a small dataset -> **need a 2D affine transformation**

Input Face
with Landmarks



Mean Landmarks



Affine Transformation



2D affine transformation

- **68 landmarks** are detected with dlib's face landmark detector
- Move eye corners and nose close to the **mean locations**
- **Resizes and crops** to 96 x 96. pixels

- 3.2 Training the Face Representation Neural Network

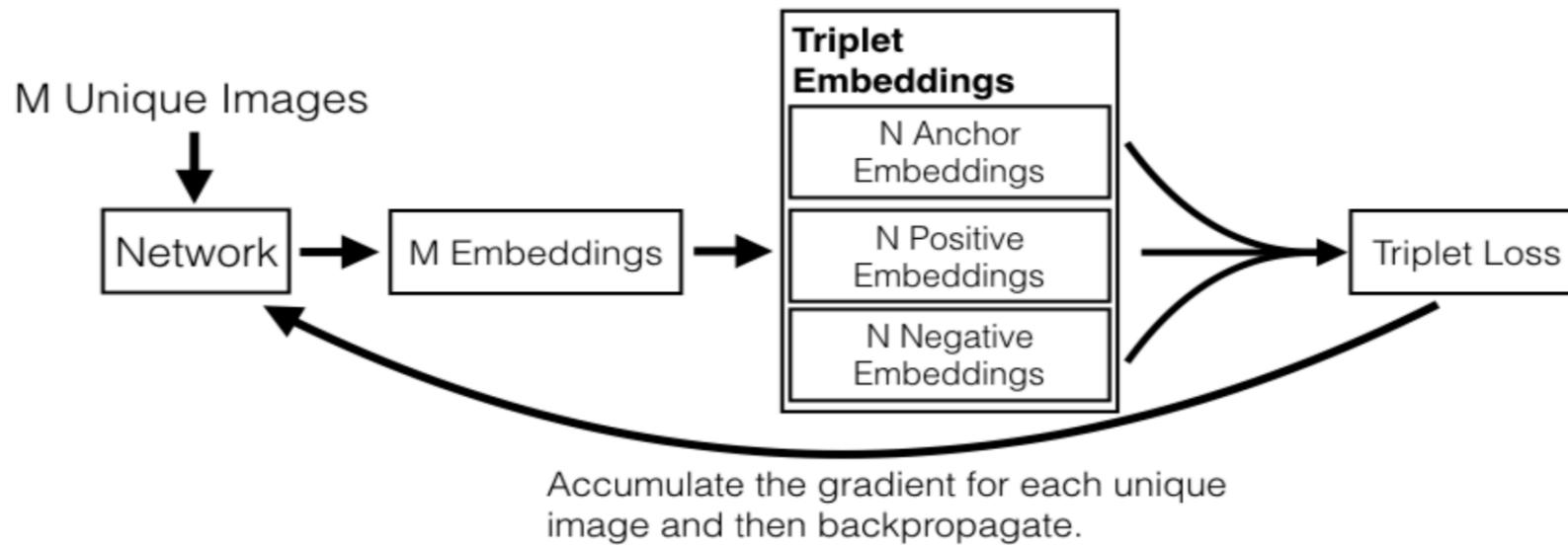


Figure 6: OpenFace's end-to-end network training flow.

- Used 500k images. From combining CASIA-WebFace and FaceScrub
- OpenFace uses a modified version of FaceNet's nn4 network based on the GoogLeNet



4

Evaluation

- Used LFW dataset

Technique	Accuracy
Human-level (cropped) [KBBN09]	0.9753
Eigenfaces (no outside data) [TP91] ³	0.6002 ± 0.0079
FaceNet [SKP15]	0.9964 ± 0.009
DeepFace-ensemble [TYRW14]	0.9735 ± 0.0025
OpenFace (ours)	0.9292 ± 0.0134

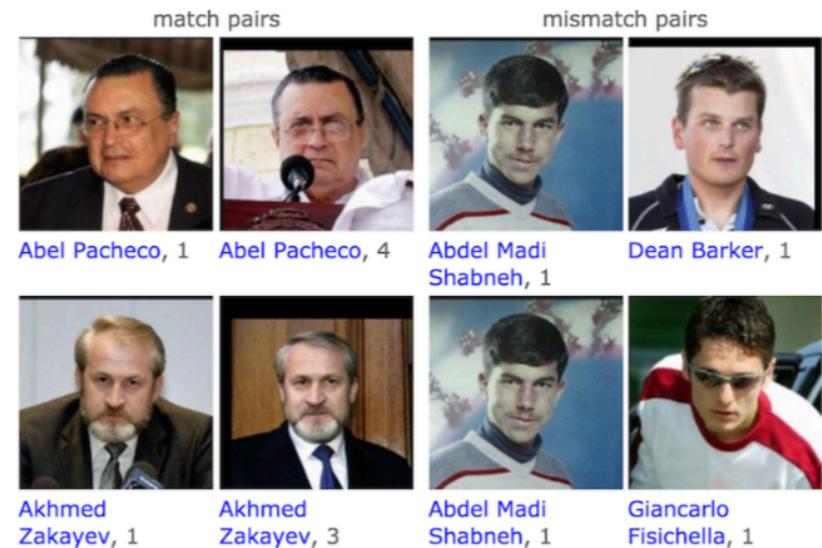


Figure 7: LFW accuracies and example image pairs.

- 4.1 LFW verification

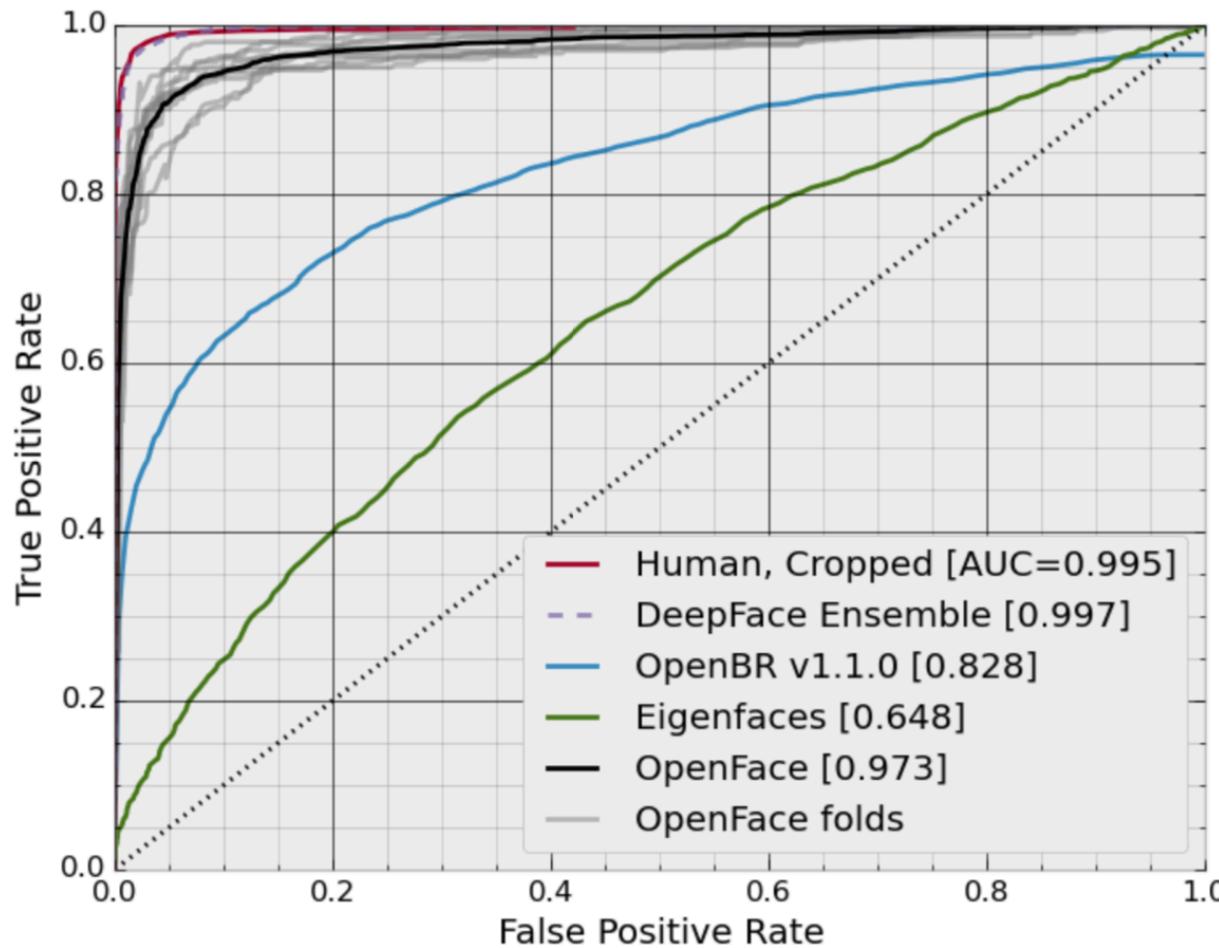
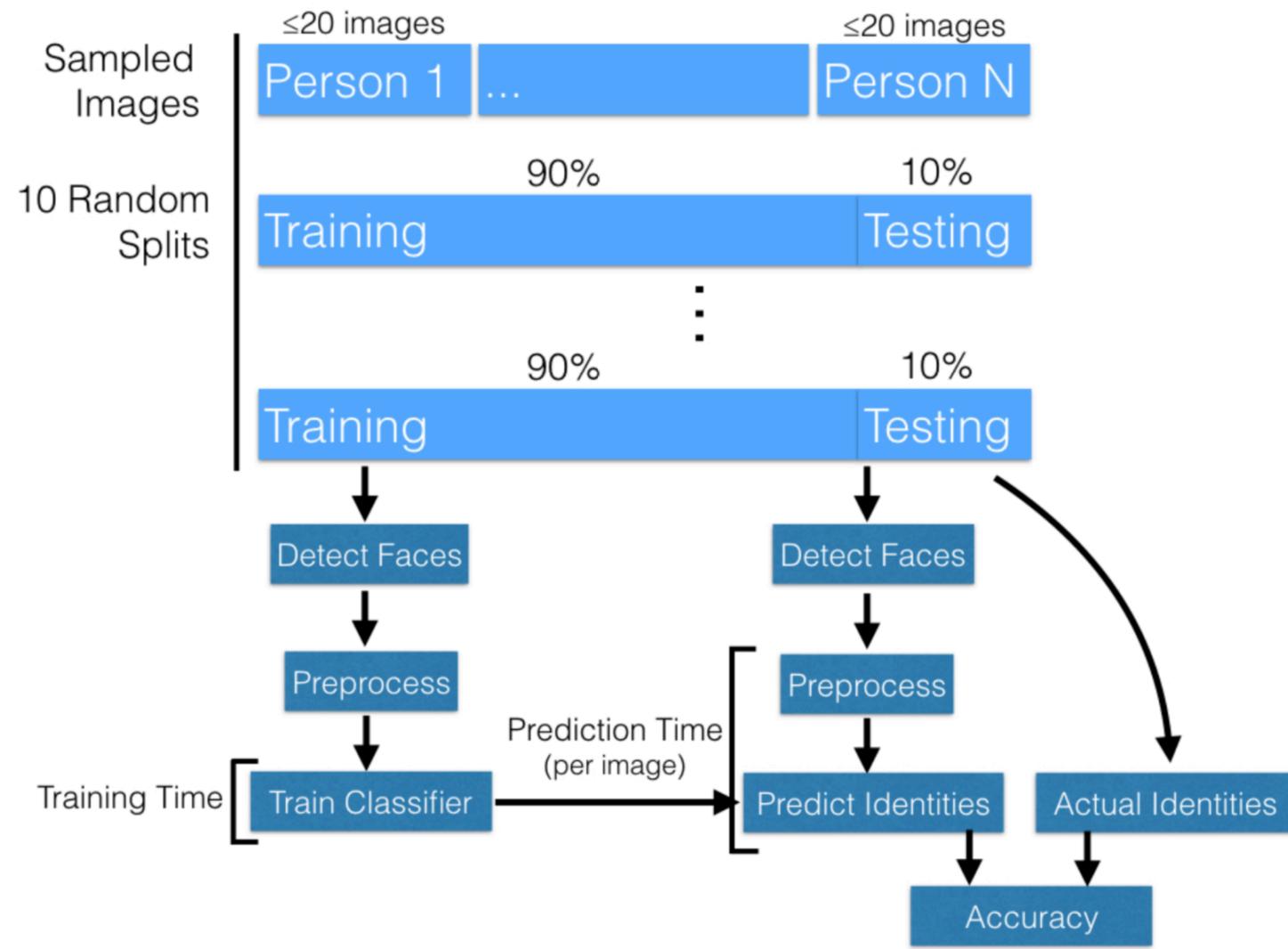
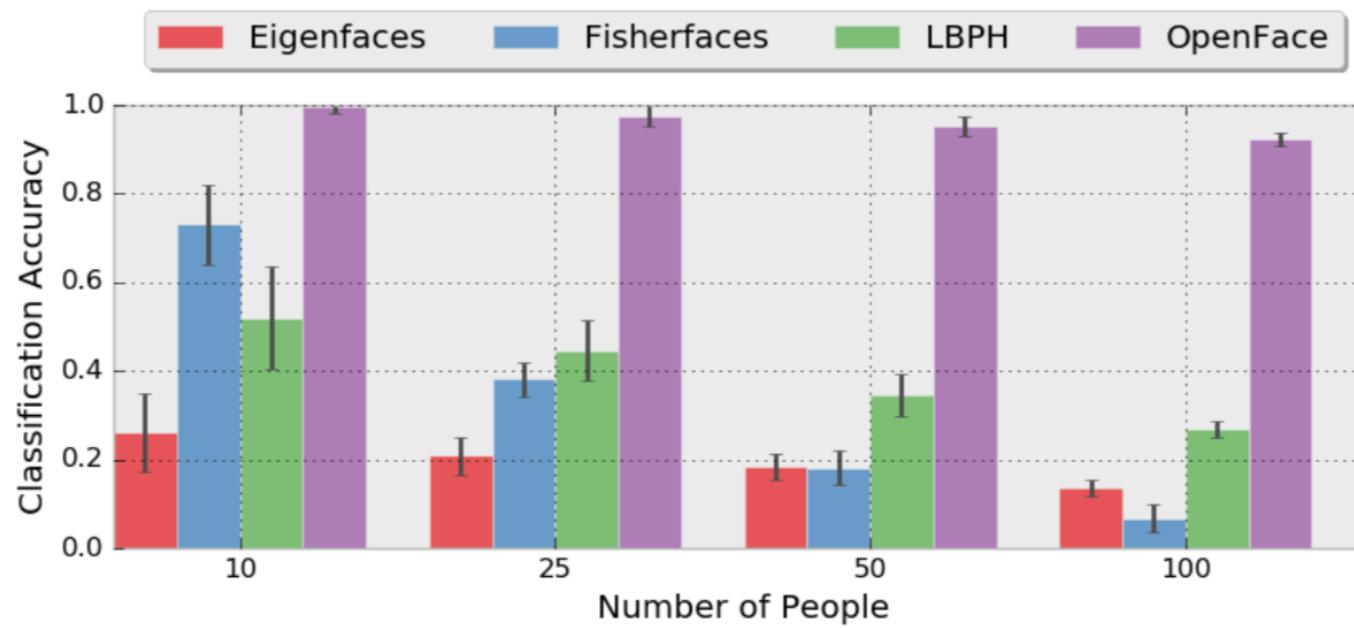


Figure 8: ROC curve on the LFW benchmark with area under the curve (AUC) values.

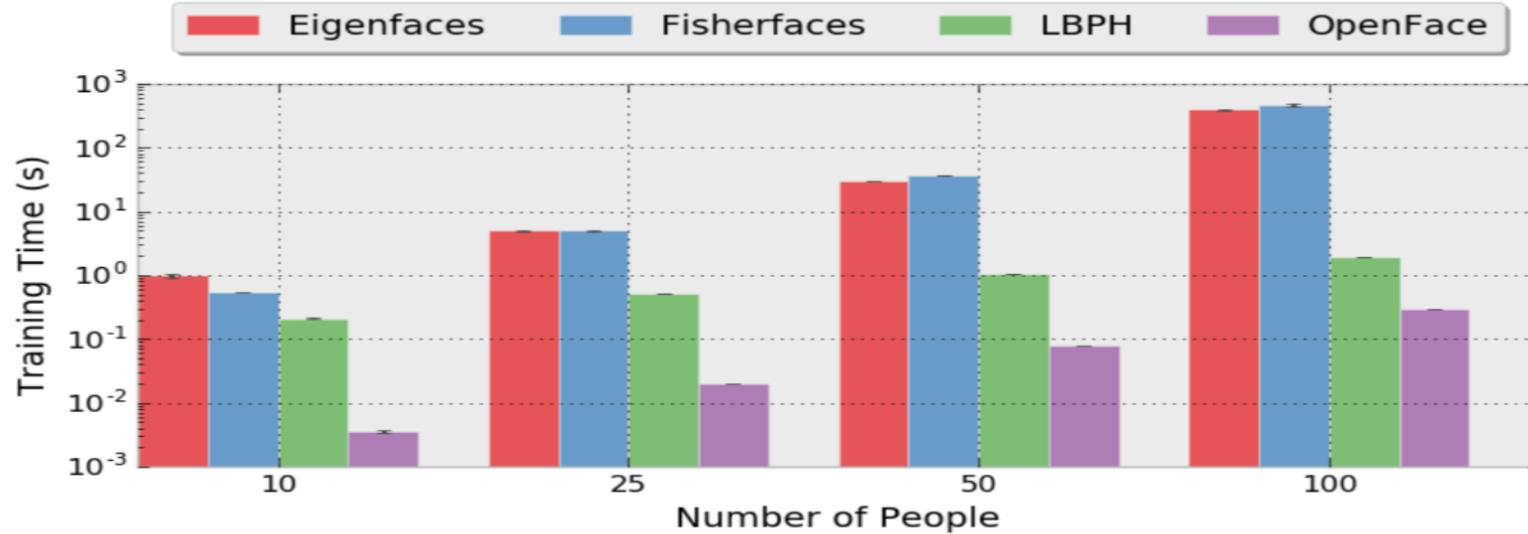
- 4.1 LFW classification



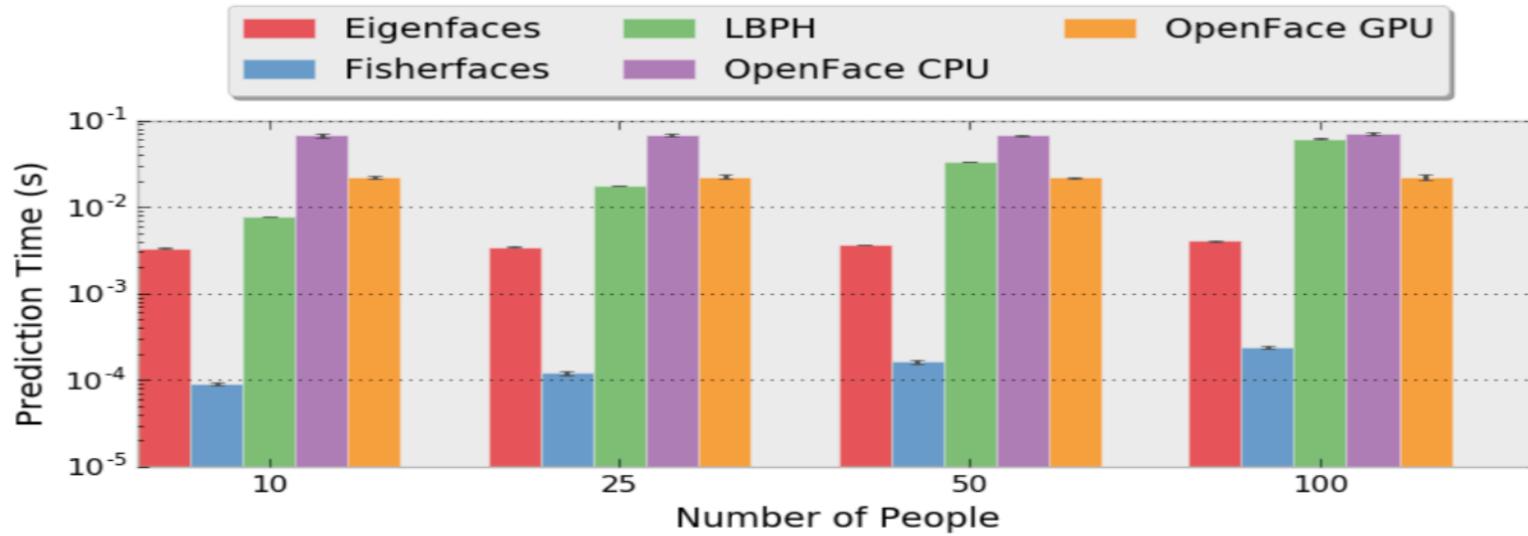
- Measuring the runtime performance of training classifiers and predicting who new images belong to is an important consideration for mobility
- User also want to **add or remove identities** from their recognition system for transient scenarios
- OpenFace classification. Uses a **linear SVM**



(a) Classification accuracy.



(b) Training times.



(c) Per-image prediction times.



5

Conclusion

- Trained a network with data smaller than DeepFace and FaceNet
- Show competitive accuracy and performance result on the LFW verification benchmark