Introduction to Biometrics

# Final Presentation

## Disentangled Representation Learning in Face Recognition

İpek Erdoğan

# Outline

- Problem definition
- Methodology
- Datasets
- Preliminary Results

# What is the problem?

The result gets affected from the redundant attributes of subjects.

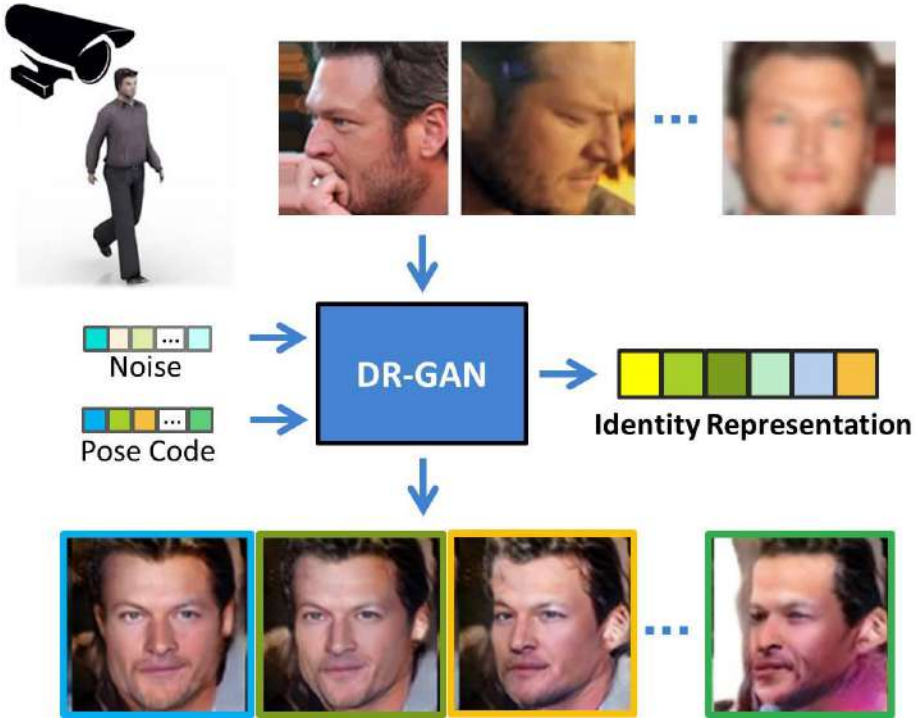
“Focusing on the wrong side”

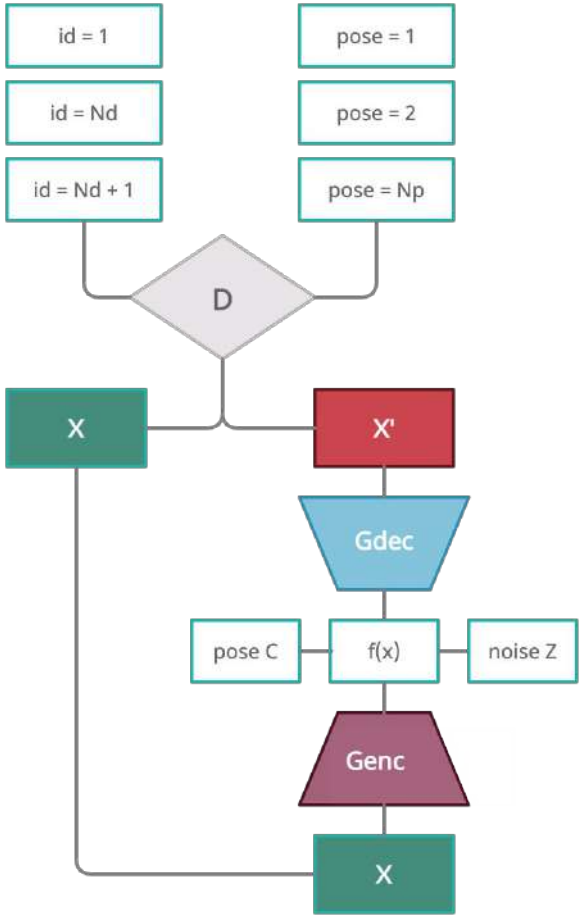
Solution: Learning representations in more specific ways for specific tasks.

- View invariant pose estimation
- Identity invariant action recognition
- Appearance invariant gait recognition
- **Pose invariant face recognition**

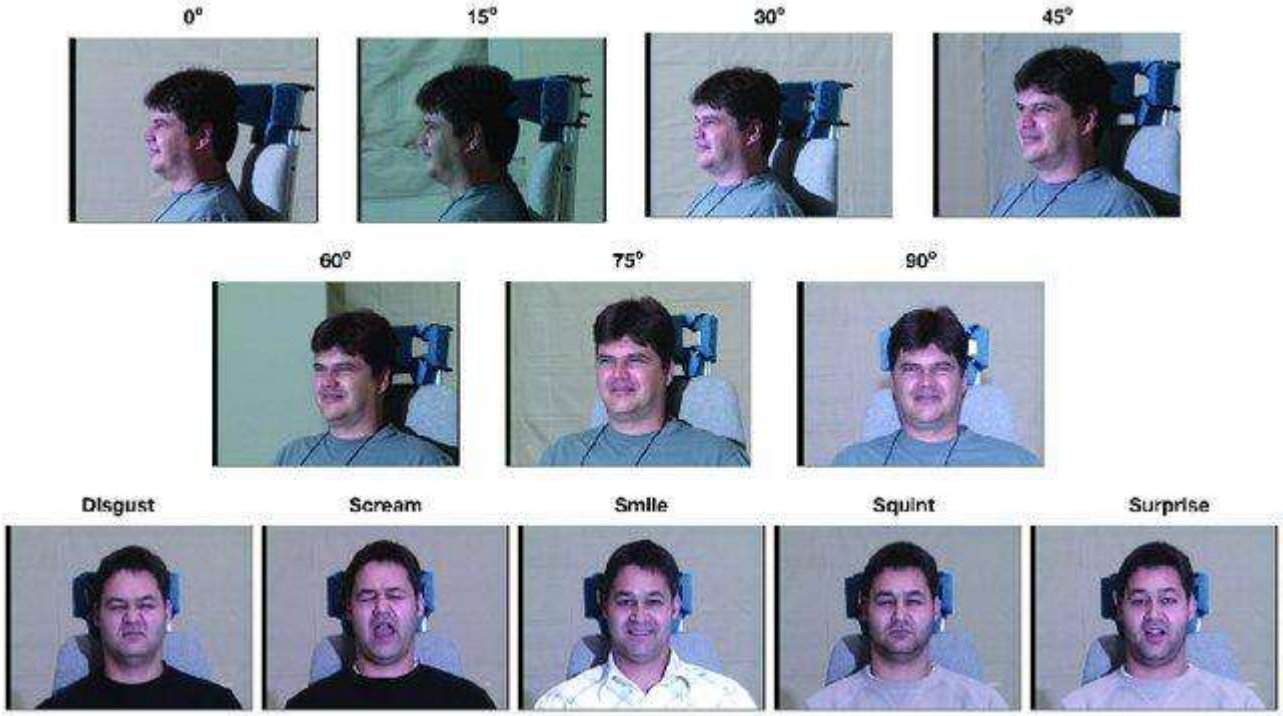
**“Disentangled Representation Learning GAN for Pose-Invariant Face Recognition”**

# Method


Noise
Pose Code
DR-GAN
Identity Representation


id = 1
id = Nd
id = Nd + 1
pose = 1
pose = 2
pose = Np
D
X
X'
Gdec
pose C
f(x)
noise Z
Genc
X

# Dataset (Training) (Multi-PIE)



# Dataset (Training) (CASIA-WebFace)

# Dataset (Testing) (Celebrities in Frontal-Profile in the Wild)

# Dataset (Testing) (IJB-A)

# Qualitative Results

Original Input

Cropped

Frontalized

# Qualitative Results

Original Input

Cropped

Frontalized

# Preliminary Results (Qualitative)

Original Input

Cropped

Frontalized

# Qualitative Results


Original Input

Cropped

Frontalized

# Qualitative Results

Original Input

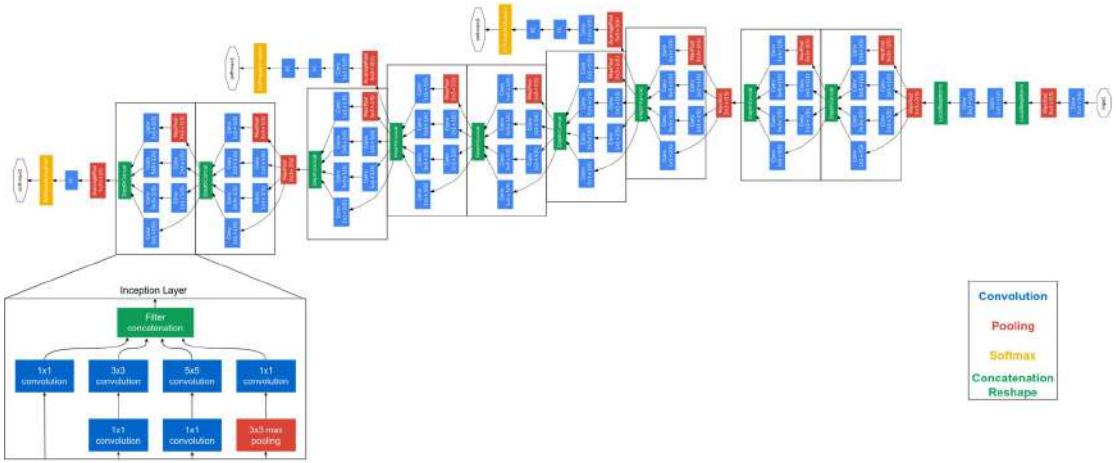
Cropped

Frontalized

# There were 2 targets:

- Recognition with frontalized images
- Recognition with 320-dimensional representations

=>FaceNet for both!



# Recognition with Frontalized Images

Pretrained Facenet on VGGFace2.

Choose data points from VGGFace2 dataset (to be able to classify and compare the results)


5223

1512

4024

# Recognition with Frontalized Images

4851

545

7412

5619

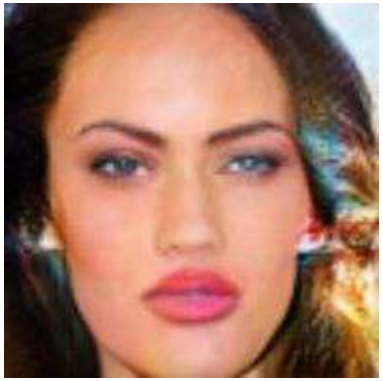
5223

931

# Recognition with Frontalized Images

5223

8019



6506

5223

1512

6974

# Recognition with Frontalized Images

5223

6406

6617

6382


5523

1253

# Recognition with Frontalized Images

8019

6903

5223

5698

5448

5960

# Recognition with Frontalized Images

7473

7779

8066

# Recognition with 320-dimensional representations

Finetuning face recognition models?

> Image dimension constraint

> Representation dimension constraint

> Dataset constraint

How about using very basic distance metrics (like Euclidian dist.) to measure the quality of representations?

# References

[1] L. Tran, X. Yin and X. Liu, "Disentangled Representation Learning GAN for Pose-Invariant Face Recognition," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1283-1292, doi: 10.1109/CVPR.2017.141.

[2] Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2010). Multi-PIE. Proceedings of the ... International Conference on Automatic Face and Gesture Recognition. IEEE International Conference on Automatic Face & Gesture Recognition, 28(5), 807–813. https://doi.org/10.1016/j.imavis.2009.08.002

[3] Yi, D., Lei, Z., Liao, S., & Li, S. (2014). Learning Face Representation from Scratch. ArXiv, abs/1411.7923.

[4] S. Sengupta, J.C. Cheng, C.D. Castillo, V.M. Patel, R. Chellappa, D.W. Jacobs, Frontal to Profile Face Verification in the Wild, IEEE Conference on Applications of Computer Vision, 2016.

[5] B. F. Klare et al., "Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1931-1939, doi: 10.1109/CVPR.2015.7298803.