PCA, Autoencoder, and FLD for Analyzing Human Faces

The human face is a very important pattern and has been extensively studied in the past 30 years in vision, graphics, and human computer interaction, for tasks like recognition, human identification, expression, animation, etc. An important step for these tasks is to extract effective representations from the face images. This project compares three types of representation in the context of dimension reduction: two generative methods, Principal Component Analysis (linear) and Autoencoder (non-linear), and one discriminative method, Fisher Linear Discriminants.

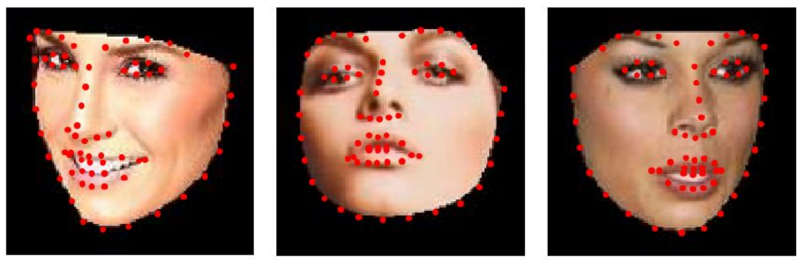

Figure 1: Example faces with 68 landmarks from CelebA. The data set contains 1000 images from CelebA, and they are cropped to 128×128 pixels by the OpenFace. These faces have different colors, illuminations, identities, viewing angles, shapes, and expressions.

<u>Principal Component Analysis (PCA)</u>

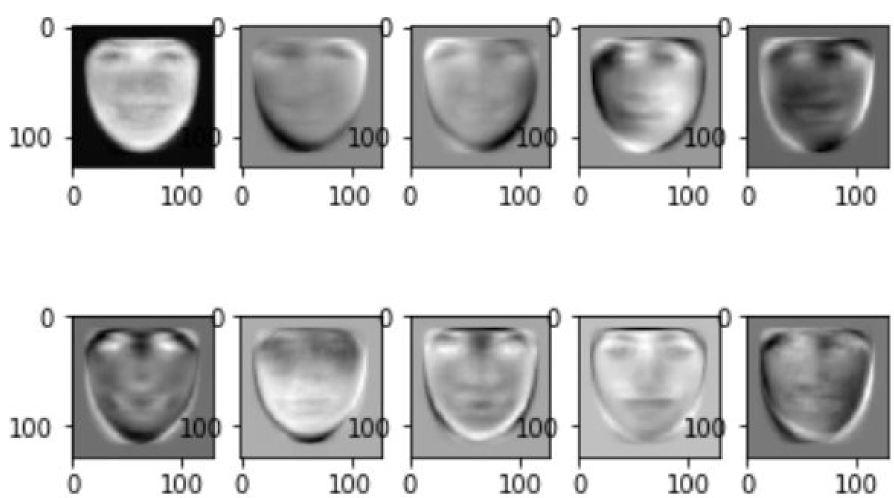

Figure 1: First 10 eigen-faces in grayscale for the training images with no landmark alignment

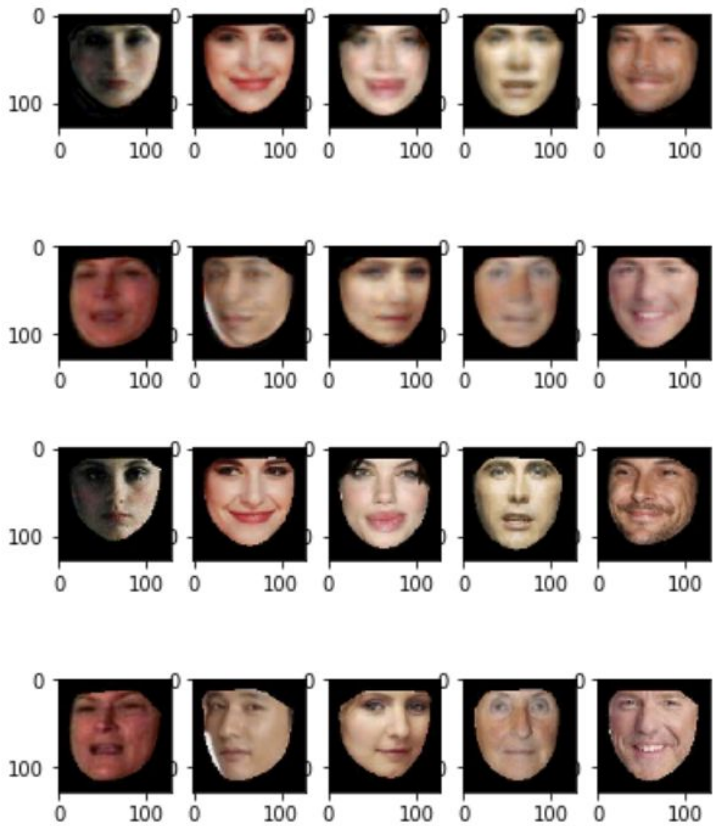

Figure 2: 10 reconstructed faces and the corresponding original faces

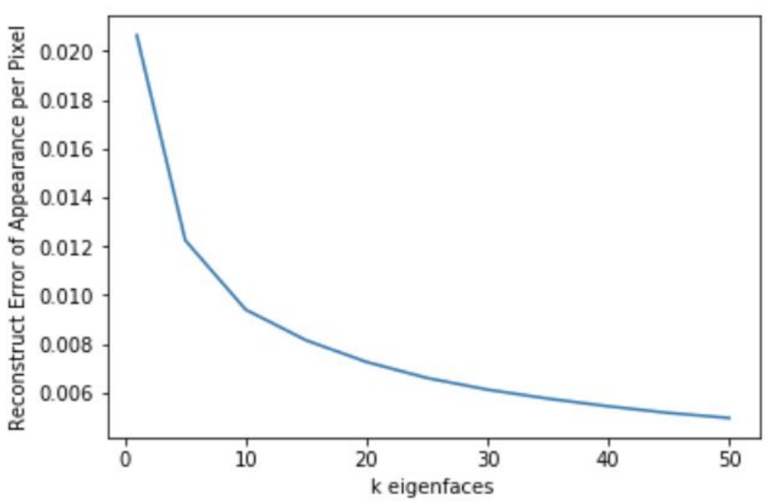

Figure 3: Total reconstruction error per pixel over number of eigen-faces k = 1, 5, 10, 15, …, 50

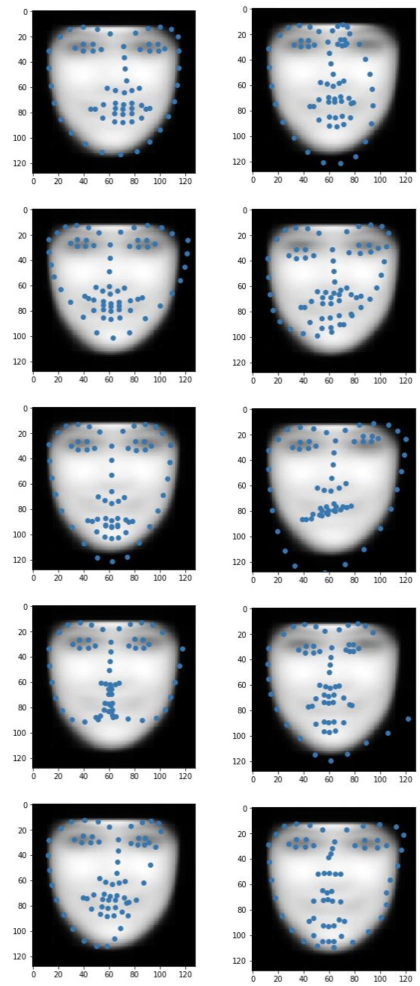

Figure 4: First 10 eigen-warpings of the landmarks for the training images

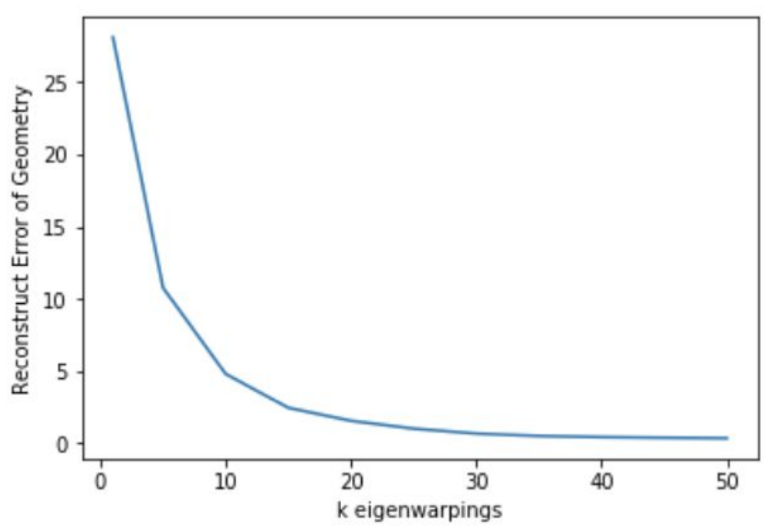

Figure 5: Reconstruction error (distance) over number of eigen-warpings k = 1, 5, 10, 15, …, 50

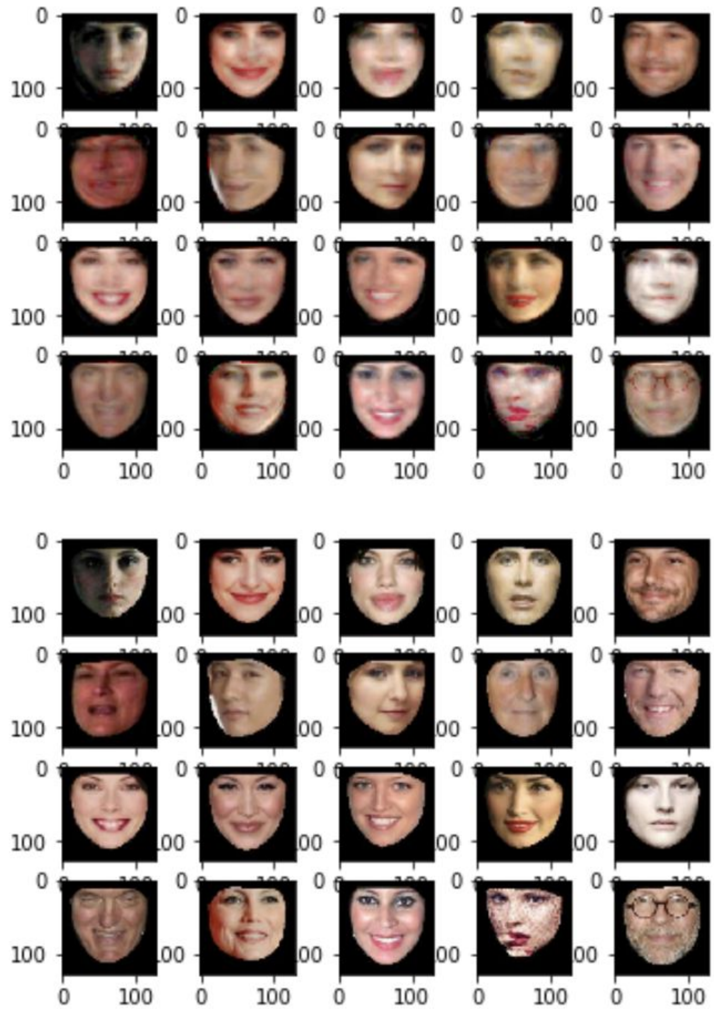

Figure 6: 20 reconstructed faces based on top 10 eigen-warpings and top 50 eigen-faces and corresponding original faces

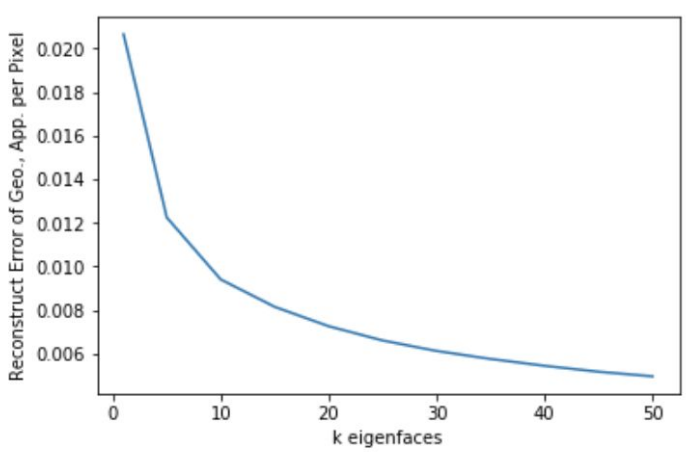

Figure 7: Reconstruction errors per pixel against number of eigen-faces k = 1, 5, 10, 15, …, 50

<u>Autoencoder</u>

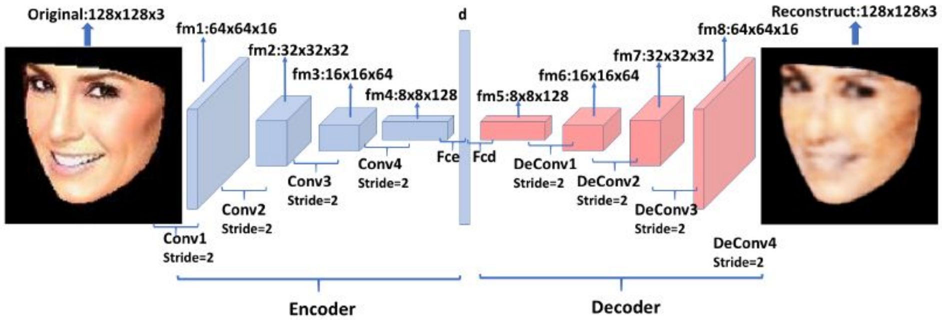

Figure 2: Illustration of Auto-encoder with convolutional architectures. An auto-encoder is a type of artificial neural network used to learn efficient data codings in an unsupervised manner.

Table 1: Suggested structure for appearance auto-encoder.

	Encoder	Decoder
0	Conv-(Channel 16, Kernel 5, Stride 2), LeakyReLU	Deconv-(Channel 128, Kernel 8, Stride 1), LeakyReLU
1	Conv-(Channel 32, Kernel 3, Stride 2), LeakyReLU	Deconv-(Channel 64, Kernel 3, Stride 2), LeakyReLU
2	Conv-(Channel 64, Kernel 3, Stride 2), LeakyReLU	Deconv-(Channel 32, Kernel 3, Stride 2), LeakyReLU
3	Conv-(Channel 128, Kernel 3, Stride 2), LeakyReLU	Deconv-(Channel 16, Kernel 3, Stride 2), LeakyReLU
4	Fc-(Channel 50), LeakyReLU	Deconv-(Channel 3, Kernel 5, Stride 2), Sigmoid

Table 2: Suggested structure for landmark auto-encoder.

	Encoder	Decoder
0	Fc-(Channel 100), LeakyReLU	Fc-(Channel 100), LeakyReLU
1	Fc-(Channel 10), LeakyReLU	Fc-(Channel 68*2), Sigmoid

Reperforming the experiments in the PCA analysis, landmarks can be reconstructed and generated by an auto-encoder with a fully-connected architecture, while two-dimensional face images can be reconstructed and generated by an auto-encoder with a convolutional architecture.

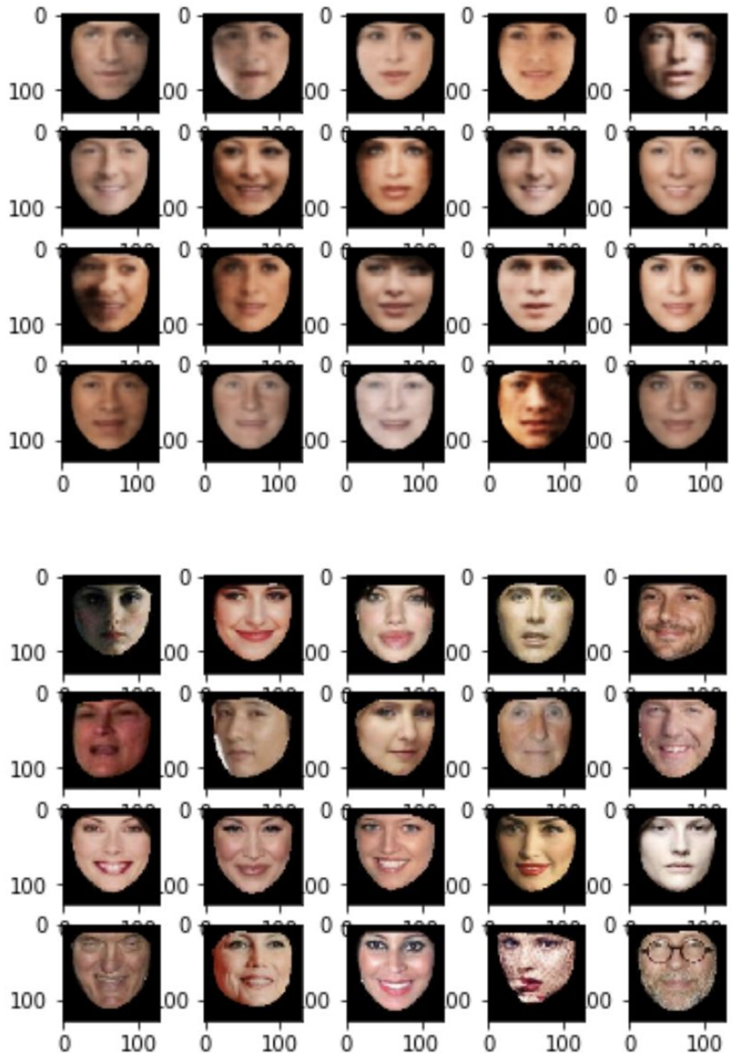

Figure 1: 20 reconstructed faces based on landmark and appearance autoencoders and corresponding original faces

Losses:
Last Epoch Loss for Training Appearance Model (with warped images): 0.01874306309036910
Loss for Reconstructing Appearance of Warped Test Images: 0.006929922848939896
Loss for Reconstructing Geometry of Test Landmarks: 0.0008123503648675978

Figure 2: Interpolation results for one face for each of the 4 dimensions of the latent variables of **appearance** that have the maximal variance, while keeping other dimensions fixed

Figure 3: Interpolation results for one face each of the 2 dimensions of the latent variables of **landmarks** that have the maximal variance, while keeping other dimensions fixed

Fisher Linear Discriminants (FLD)

Figure 1: Error rate on 200 test faces for the Fisher face, which considers both geometry and appearance differences between females and males

Figure 2: All faces projected to the 2-D feature space learned by the fisher-faces