IDC 409, Project 2: Computer Vision

Digital Attendance System Based on a Group Photo

Amlan Nayak, MS18197
Dhananjay Joshi, Ms18065 : Equal Contribution

12/12/2021

1 Ideas, Approach & Pipeline

1.1 Basic Approach

Our project is divided into two parts. First, we tried to deploy a basic **Principal Component Analysis** combined with a **support vector machine framework** to create a setup that recognizes human faces from a photograph and can be used as an attendance marking system. We work on the AT&T dataset which consists of 40 individuals, each having 10 fully frontal images. The size of each image is 112×92 and all images are in grayscale.

The second part involves detection of faces from an input images once the PCA + SVM framework is complete. For this we had to use transfer learning methods in **opencv** and **dlib** libraries. Both have simple pre trained detectors called **harr cascade** and **frontal face detector**. We used this to detect faces from an given input image and store it in a folder.

1.2 Intial Trials & Improvements

We make a matrix out of all the 400 images, with the rows being the total numbers of pixels and columns being the number of images. As done in class, we provide with the standard PCA method. After doing PCA and finding the **top PCs** which given 90 percent of cumulative variances. We then proceed with calculation of euclidean distance of a PCA projected test images. However, the predictions were quite poor with barely 55-65 percent accuracy even on the test dataset.

We then decided to use another method taught in class to improve our test accuracy. We used the **Support Vector Machine** approach on the PC projected images to obtain a 98% accuracy on test set.

1.3 Final Machine Learning Pipeline

Our submission consists of two codes. The harr cascade/dlib code finds the faces in a random image and saves them in a folder. The PCA + SVM code then finds PCs of the original dataset and projects our input images onto them. It then processes them through an svm classifier and predicts the classes of the image. We summarise all the predictions in a dataframe as our attendance and save it in excel spreadsheet.

2 A PCA based Attendance Marking System

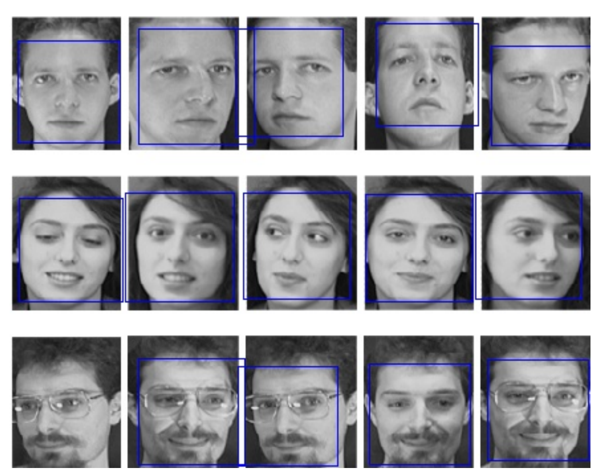
The various steps involved in executing the above mentioned ideas are as follows:

2.1 Procuring Data

We used the openly sourced **ATT face dataset** which contains a set of face images taken between April 1992 and April 1994. This database was used in the context of a face recognition project carried out Cambridge University Engineering Department. The database contains 10 images each of 40 human subjects taken in varying lighting, angles and facial expressions.

2.2 Prepossessing Data

- The raw image data was suitably pre-processed by first converting the images to gray-scale and then cropping and scaling to obtain a workable image containing the requisite facial features. We used an open source algorithm: **"Haar feature based cascade classifier"** to achieve the same. The algorithm detected most of the faces but there were some false positives. In the hope of improving detection accuracy we also used a **face detection algorithm by dlib**.



The Haar feature based cascade classifier crops the image as shown above.

2.3 Building Principal Components

Principal Components for the data set were built using the the **PCA** library in the **Sklearn** module which implements the same linear algebra operations as described above. The scores were calculated by using a reduced eigenbasis matrix of reduced dimension r, which was optimized as explained later.

2.4 Query Recognition

The query image is also similarly cropped and scaled as the original data, and then projected on the same eigenbasis as the dataset. After the similarity between the query image and each distinct image value in the database is calculated through the following ways:

2.4.1 Euclidean Distance

In the first approach, we simply consider the Euclidean distance between the query image and database images and declare two images two be similar if the same is less than a pre decided threshold value. However, the results of this approach were poor, forcing us to abandon it for an alternative.

2.4.2 Support Vector Machine Algorithm (SVM)

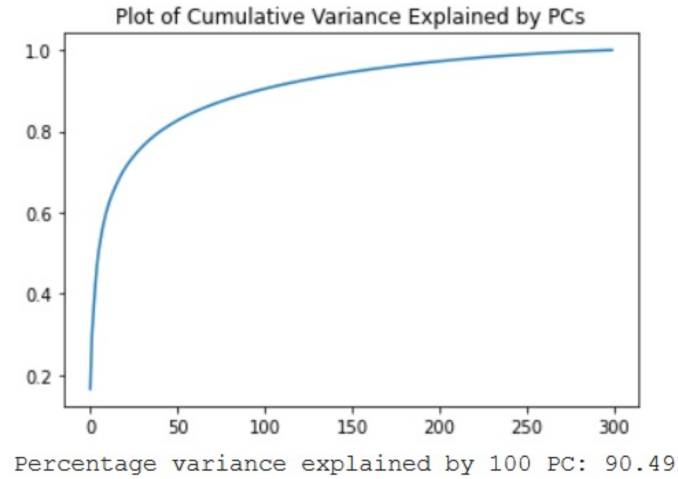
We have used randomized search cross validation module from the sklearn package which is a heuristic method for hyper parameter tuning. Our best parameters for SVM in sklearn of python were a linear kernel with gamma value of 0.001 and C value of 1000. SVM showed a higher accuracy (98% accuracy on test data) as compared to comparatively poor Euclidean distance $55-65\%$ accuracy on test data. Comparison shown below:

Euclidean vs SVM: Compared

2.5 Optimization and Error Calculation

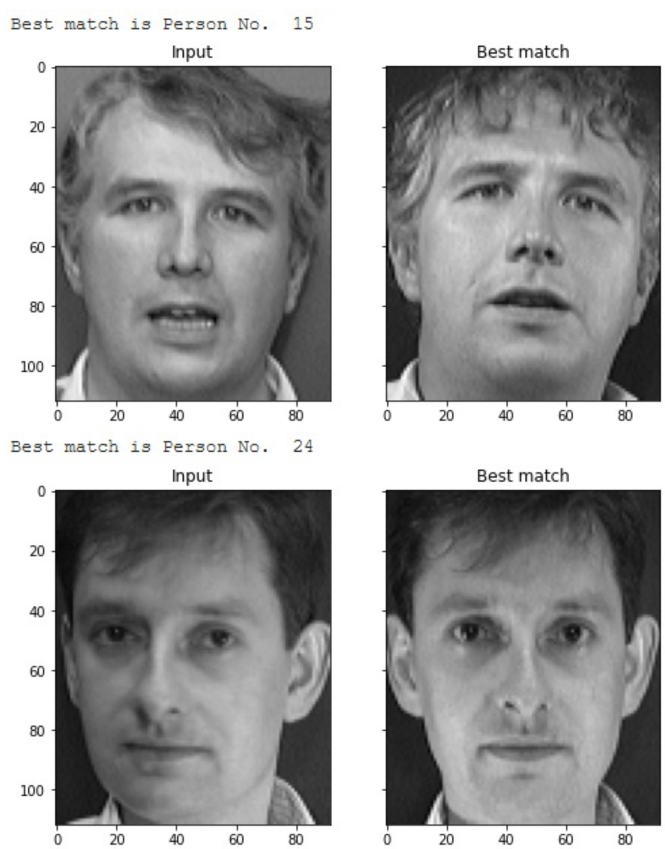
2.5.1 Optimizing R

For optimizing R, or the reduced dimension of the principal component matrix, we plot a curve between the cumulative variance explained vs R and obtain the following curve:



The saturating curve shows that 90% of the variance in data is explained by R=100.

3 Results



Dunno what to put.

Person	Attendance
1	1
2	NaN
3	NaN
4	NaN
5	NaN
6	NaN
7	NaN
8	NaN
9	NaN
10	NaN
11	NaN
12	NaN
13	1
14	NaN
15	NaN
16	NaN
17	NaN
18	1
19	1
20	NaN
21	NaN
22	NaN
23	NaN
24	NaN
25	NaN
26	NaN
27	NaN
28	NaN
29	NaN
30	NaN
31	NaN
32	NaN
33	NaN
34	NaN
35	NaN
36	NaN
37	NaN
38	NaN
39	NaN
40	NaN

Final attendance data frame

4 Sources of Error

- The system's accuracy is largely limited by the ability of the **Haar feature-based cascade classifier** to accurately detect faces from the supplied group photograph. This method roughly detects 90% of the faces in a photograph.
- Another possible source of error is in how the **Haar feature based cascade classifier** pre-processes images. In order to get globally comparable features, some secondary features like the hairline, ears and shoulders of the subject are cropped. While this ensures the most essential primary features are used in PCA, but perhaps effective inclusion of the secondary features would further augment the recognition efficiency.
- Image quality: Image quality parameters like low resolution and low colour gradient can affect the efficiency of the system.

5 Scope of Improvement

- Instead of using the Harr classifier we can use a custom built and detection algorithm employing CNNs
- Improvement in the resolution of images and using a bigger dataset can improve the accuracy significantly.
- In future we would want to construct an API which would unite the front end and back end operation for a smooth user experience.

6 Interpretation & Final Thoughts

A combination of dimension reduction and simple SVM classifier allows us to generate a decent yet transparent and open source machine learning pipeline for a class attendance system. With the deep learning based methods being taught in IDC410, we will come back to this with a neural network based approach which has the ability to dramatically increase the scope as well as the accuracy of the project. For now, a very simple framework with the help of transfer learning achieves the purpose. Through this project we could appreciate the importance of transfer learning in solving problems. Without transfer learning, we could not have detected our faces given the lack of suitable hardware to train custom made models.

7 References

- ATT database
- Scikit Learn package
- Frontal Face detector dlib