Face Recognition Challenge

Om Thakkar (201501109), Jay Mohta (201501036), Shivam Raval (201501088), Anushree Rankawat (201501007)
School of Engineering and Applied Science, Ahmedabad University

***Abstract*—Given a dataset of 588 images of all the members of the Machine Learning Course 2018 at School of Engineering and Applied Science, Ahmedabad University, we present a novel approach for Face Recognition. In this work, initially the background is removed by applying Haar Transform on all the images. The recognition is performed using the Keras Sequential Model of Neural Networks after using Principal Component Analysis for reducing dimensionality of the images. Extensive experimentation on the training and testing dataset of images yields an excellent testing accuracy proving that the approach used here is highly reliable.**

***Index Terms*—Face Recognition, Keras Sequential Model, Haar Cascades, Principal Component Analysis, Neural Networks**

I. INTRODUCTION

FACE recognition systems have been conducted now for almost 50 years. Face recognition is one of the finest research areas in pattern recognition and computer vision due to its numerous practical applications in the area of biometrics, Information security, access control, law enforcement, smart cards and surveillance system.

Technical evolution of algorithms used for Face Recognition Problem is highly significant. Face is one of the variables that are very easy to remember in real life. Generally, humans can remember and recognize a person based on his face. However, face is one of the most complex variables when viewed from the perspective of computer vision with varying features and characteristics of each face.

In this Face Recognition Challenge, a dataset of 588 images has been considered. Detailed description regarding the dataset has been provided in the *Dataset* section. Various observations such as the difference in accuracy of original images and pencil sketch images, testing accuracy for an image extracted from the original dataset (but used solely for testing), testing accuracy for images with cluttered background, etc are found interesting.

The rest of the report is organised as follows. Section II contains the details of the related work in the field of Face Detection and Recognition and the various innovative approaches in the literature that are proved to be very significant. Section III provides an overview of the dataset used for Face Recognition Section IV provides a detailed description about the approach by which the Face Recognition Challenge has been solved in this work. Section V lists the various experimental results and outcomes of the Face Recognition Challenge considering different varieties of problems in mind. Section VI has discussions regarding the future possibilities and improvements to be considered in future followed by References.

II. RELATED WORK

Until recently, face recognition technology was commonly viewed as something straight out of science fiction. But over the past decade, this groundbreaking technology has not just become viable, it has become widespread. Face Recognition System is a key technology and is being extensively used in various fields such as biometric security and safety for the convenience and ease of users.

Face recognition has gained significant importance in the research community which has led to the development of various algorithms for the same. Biometric usage is the highest security system compared to traditional systems (using password or ID Card for authentication) on the security system. One of the few applied examples of biometrics is the facial recognition used in the security system [5]. In general, Face recognition can be said to have two major sub-fields of research in itself, which are face detection and face recognition. To perform face detection, Cascade Classifier method or Viola-Jones method is used [6]. This method involves the use of Haar Cascades for feature selection and uses certain face related observations to localise faces in an image. This was one of the pioneer works in the field of Neural Networks which showed that face detection is a problem solvable in real time with great accuracy. To perform face recognition, there are many algorithm that can be applied with face recognition. In [4], face recognition was performed by Feed Forward Neural Networks algorithm (FFNW). In [1], [2], [3] face recognition was performed by Principal Component Analysis (PCA) algorithm. In 2000, a new method of feature selection for face detection had been proposed using Support Vector Machines (SVMs). After reducing the dimensionality of input space, the features have been selected from SVM feature set that have low contributions to the decision function of SVM.

III. DATASET

Face Recognition Challenge that has been considered here consists of a considerably smaller dataset of 588 images of 37 students enrolled in Machine Learning Course 2018. The dataset contains 16 images of each student with following varieties :

- 6 different basic facial expressions (Anger, Fear, Surprise, Sadness, Joy and Disgust)
- 1 image with goggles or glasses on eyes
- 1 image with hat on head
- Pencil Sketch of 8 images described above

The predefined file naming conventions for original images are as follows: studentID_expression_o.jpeg, studentID_goggles_o.jpeg, studentID_cap_o.jpeg

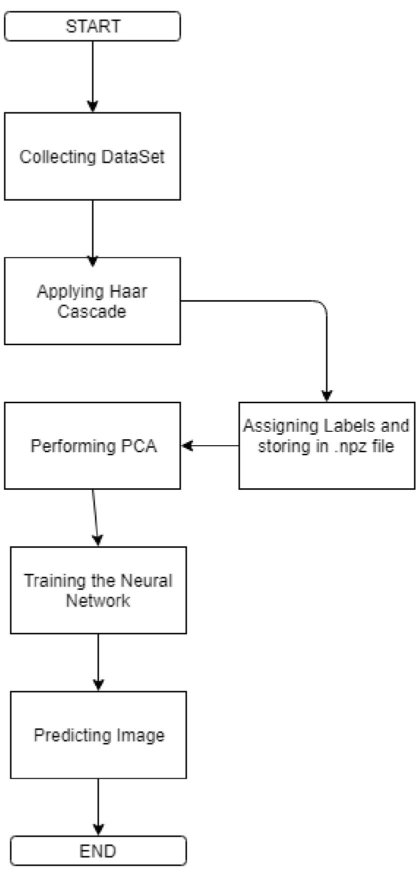
The predefined file naming conventions for pencil images are studentID_expression.jpeg, studentID_goggles.jpeg, studentID_cap.jpeg

Clearly, the dataset is enriched with variety of images and the small size of the dataset makes the Face Recognition Problem quite challenging. The dataset has lots of variations. The dataset has images with different resolutions, sizes, backgrounds, percentage of frame occupied by the face, different file formats, etc.

All of these variations result in the need for data-cleaning and pre-processing to be done before proceeding any further. The section below starts with the description of pre-processing followed by the algorithm/specifics of the Face Recognition model.

IV. Our Approach

Our approach to Face Recognition involves three major steps: Pre-processing, Dimensionality Reduction and Classification. We will address these three steps in the subsequent subsections. The flow of the approach used for solving the Face Recognition Challenge is as follows :



A. Pre-processing

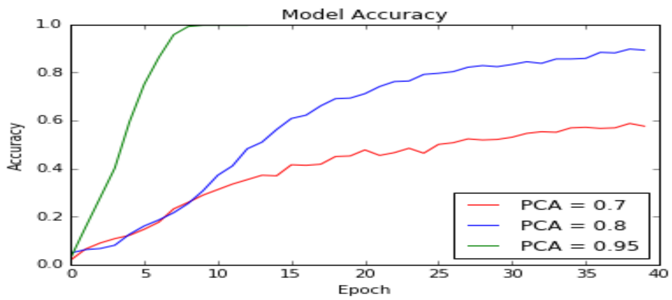
Initially, the collected dataset required a lot of preprocessing like removing the redundant images, renaming file names to maintain naming convention, assigning labels to images such that each label identifies a person, etc. Redundant images were removed manually since the dataset is quite small. However, if the dataset was large, we would have removed the redundant images programmatically after recognizing certain patterns in the same. The images were renamed in order to avoid errors in labelling the images. Here, since the problem only deals with identifying the person based on any image provided (instead of predicting the expression, etc.), we have assigned a unique number to each person and every image of that person has the same label number assigned to it.

Since, the backgrounds of all the images was quite different and diverse, there was a need to separate foreground (here, a face) from background to ensure that the model doesnt learn noisy background features. Using the approach of Haar Cascades, the majority of the background was removed from all the images. In Haar-cascade, the system is provided with lots of images, and the feature selection is done along with the classifier training using Adaboost and Integral images. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. Haar-cascade is an object detection algorithm used to locate faces, pedestrians, objects and facial expressions in an image and performs extremely well for face detection. The main advantage of using Haar-Cascades over any other algorithm for face localisation is that the decrease in time complexity for the algorithm is significantly higher than any other face localisation algorithm.

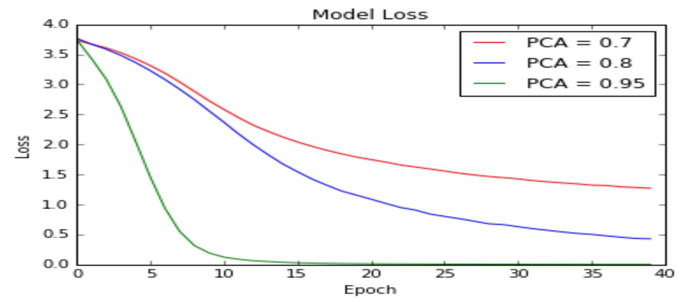
Also, the size of every image was reduced to a constant size of 100x100 for the purposes of uniformity. This enables the model to be trained on images having similar complexities and dimensions. This ensures that all images have equal contribution in training the model and no image is overpowering over the others in the model.

B. Dimensionality Reduction

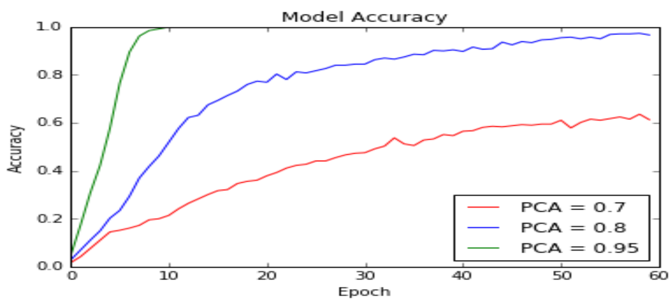
In order to make our Face Recognition Algorithm computationally less intensive and more robust, we have considered reducing the dimensionality of all the images in the dataset. Dimensionality reduction is the process of reducing the number of random variables under consideration by obtaining a set of principal values or basis which can represent or span the entire dataset. The linear technique of dimensionality reduction, namely Principal Component Analysis (PCA) has been considered here. PCA performs a linear mapping of the data to a lower-dimensional space in such a way that the variance of the data in the low-dimensional representation is maximized. In practice, the covariance (and sometimes the correlation) matrix of the data is constructed and the eigen vectors on this matrix are computed. The eigen vectors that correspond to the largest eigenvalues (the principal components) can now be used to reconstruct a large fraction of the variance of the original data. Moreover, the first few eigen vectors can often be interpreted in terms of the large-scale physical behavior of the system. The original space (with dimension of the number of points) has been reduced (with data loss, but hopefully retaining the most important variance) to the space spanned



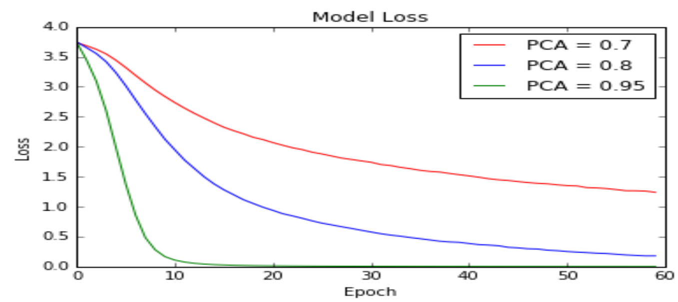
(a) Accuracy for Epochs = 40



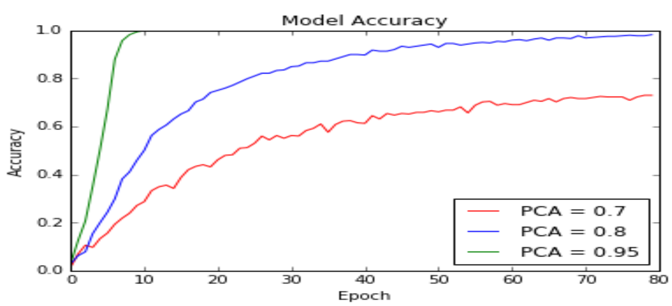
(b) Loss for Epochs = 40



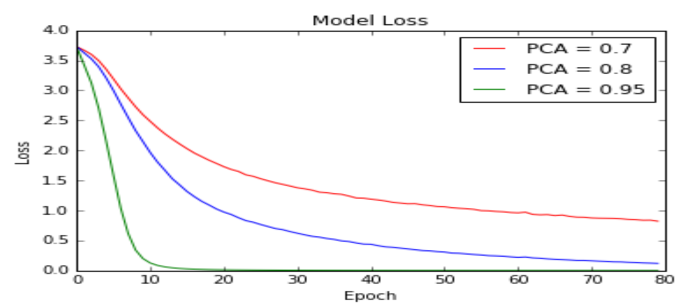
(a) Accuracy for Epochs = 60



(b) Loss for Epochs = 60



(a) Accuracy for Epochs = 80



(b) Loss for Epochs = 80

by a few eigenvectors. Its observed that there is insignificant loss in the energy function of the image as the basis vectors with low variance generally correspond to noise in the image.

Here, we have chosen the number of eigenvectors to be 90% of the original eigenvectors. The features are calculated by mapping the old points in this new vector space of reduced eigenvectors as the basis. This helps in reducing the dimensions by removing redundant information and without significant loss in the energy functions of the image.

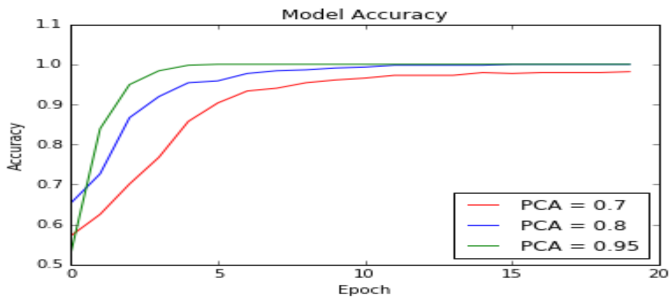
C. *Classification of Images*

For the final step of face detection, we have used a neural network based approach. Neural Networks are known to give great accuracies with most of the real world problems, though requiring large amounts of training data for the training phase.

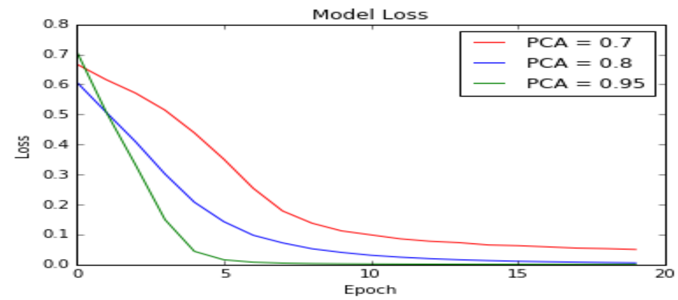
The ability to learn and model non-linear complex relationships is what gives an edge to neural network based approached over the traditional ones. Since it models the neuron structure of the brain without imposing any restrictions on the input variables, it learns hidden relationships even in data with high volatility and non-constant variance.

Here, we have created a neural network using the Keras Sequential model having 4 layers: 1 input layer, 2 hidden layers and 1 output layer. The number of neurons in the input layer is 232. We have used the RELU function as the non-linear activation function for the input and hidden layers. Its a prefered choice for modeling non-linearity as it accurately models the nature of many biological phenomenon with efficient computation, efficient and smooth gradient propagation, sparse activation nature and the useful property of scale-invariance. For the last output layer, we have used the softmax function as the non-linearity measure as it can help us map the values obtained in the range of [0,1], helping us give a probabilistic measure of the classification obtained for a given input image. The number of neurons of the output layer, i.e., 41 corresponds to the number of people in the Face Recognition Database.

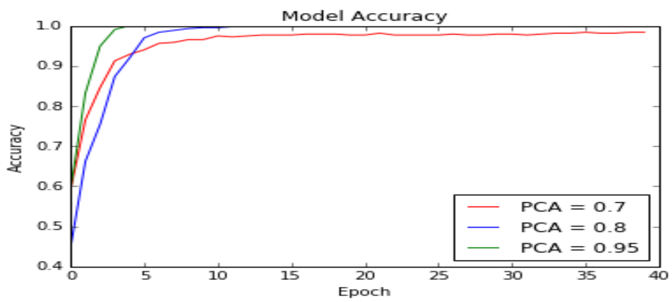
The Neural Network has a batch size of 16, with experiments conducted on the number of epochs used. We have used the categorical cross-entropy as our loss function and accuracy as a metric for evaluating the performance of our model. We have discussed the same in the next section.



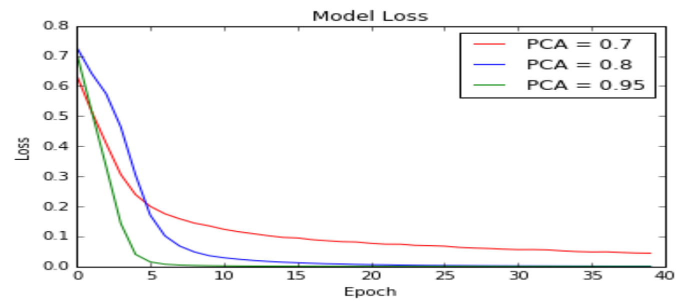
(a) Accuracy for Epochs = 20



(b) Loss for Epochs = 20



(a) Accuracy for Epochs = 40



(b) Loss for Epochs = 40

D. Classification of Pencil Sketch and Colored Images

The problem here is that of distinguishing between real image and pencil sketches. The first challenge here is to separate the dataset into pencil sketch and real images. Since the labelling convention was not unique nor was the ratio of pencil sketch and real images, a different approach had to be adopted for creating labels for the training set.

We have assumed here, in accordance to the dataset provided, that the real images are color images and the pencil sketches are grayscale images. Using the image matrix properties, color images are those with non-zero difference in each pixel value of three different image matrices. Grayscale images have zero difference of every pixel value from the 3 different image matrices.

Algorithm :

Extract pixel matrix of blue,green and red
rg = difference of red and green
rb = difference of red and blue
gb = difference of green and blue
diff = rg + rb + gb;
if diff == 0 (near to zero)
"Pencil Sketch"
else
"Coloured Image"

Now, for the task of classification, we have used a sequential neural network to distinguish between real image and pencil sketch. The loss function used was the binary cross entropy loss, as we are classifying the input into 2 main classes. The results obtained for the same are as shown above.

When tested on training set, the sequential neural network model for distinguishing between real and pencil sketch images provides an accuracy of 100%. When tested on wild images, the model was able to classify images correctly out of the images it was tested on.

V. SVM : A DIFFERENT APPROACH FOR CLASSIFICATION

Support Vector Machines is a discriminative classifier formally defined by a separating hyperplane. It performs the classification task by constructing hyperplanes in multidimensional planes that separates cases of different class labels. SVM supports both regression and classification tasks and can handle multiple continuous and categorical variables.

SVM based approach would be suited for our dataset because it works best with datasets of small size because of its quadratic and higher polynomial loss function. Support Vector Machines focus only on points which are the most difficult to tell apart from each other. The intuition behind this is that if the classifier is good at the most challenging comparisons, it would do great for other easy comparisons by translation.

Support vector machines focus on finding the best separating line/hyperplane. To do this, it searches for the closest points, which it calls support vectors. Once the closest points are found, SVM draws a line connecting both the points from the 2 separate classes. The best separating line is then the line that is perpendicular to the connecting line.

SVM has different models and kernels and we have chosen sklearns C-support vector classification model with the radial basis function kernel. The fit time complexity is a polynomial of degree 3, which makes it hard to scale to dataset with more than a couple of 10000 samples. Hence, it works well with our small dataset.

For C=2 and Gamma=10,

Obtained Training Accuracy = 100%

However, changing the value of the parameters, we obtained different and less accurate results.

VI. Experimental Results

We have conducted certain experiments by changing the number of epochs (training cycles of the training set) and by changing the number of features retained after PCA, with the cross-entropy loss magnitude.

Our metric for evaluating the performance of the model is Accuracy. As the name suggests, it simply figures out how many amongst the training data were classified correctly i.e., how many images amongst the training data had identified the correct person the face belongs to in the face recognition challenge. This performance metric isnt sufficient to validate that a model would perform well in the real world data set. We have also used the categorical cross-entropy as another metric for performance evaluation.

In our experiments, we observed that for 95% of eigenvectors being chosen as the basis of our new space, the model performs with very less variations if the epochs are varied in the range of 40-80. On the other hand, if 80% or 70% of the eigenvectors are chosen as basis of our new space, performance increases as number of epochs are increases from 40 to 60 and finally to 80. Our model predicts comparatively well with 80% of eigenvectors being retained as the basis of our new space and epochs being 60. These parameters give the best possible results with the most amount of dimensionality reduction. The accuracy obtained in this case is of 99.08%, which is extremely well for face recognition.

The model performs very well in images which are similar to the images present in the database. For the prediction of wild test images, our model predicts them satisfactorily well, though it does sometimes fail if the background is too cluttered or rich in variations. For the 7 wild images tested on the model, the model predicted 3 of them correctly.

VII. Discussions

Face Recognition is a problem which requires a large amount of data for correct recognition of wild test images. Since, our dataset was very limited in the number of images that the neural network could learn from, the model didnt perform so well for wild images. One of the techniques that could be applied for increasing the size of dataset is to perform certain image manipulations in the current dataset, like rotating the image, increasing brightness and so on for increasing diversity in the dataset to train the model on. The major issues in the problem statement were large variations in the image matrix sizes and the drastic variation in the background images. These were handled by removing the background using Haar Cascades and keeping the image matrix sizes same. The problem of recognition in wild images in still unsolved and would be addressed by us in our future work.

References

[1] A. Singh, and S. Kumar, Face recognition using PCA and Eigenface approach, Thesis: National Institute of Technology Rourkela, 2012.

[2] D. Kumar, Rajni, An efficient method of pca based face recognition using simulink, International journal of advanced in computer science and technology, vol. 3, pp. 364-368, May 2014.

[3] M. Slavkovic, and D. Jevtic, Face recognition using Eigenface approach, Serbian journal of electrical engineering, vol. 9, pp. 121-130, February 2012.

[4] M.S.R.S. Prasad, et al, Face recognition using pca and feed forward neural networks, International journal of computer science and telecommunications, vol. 2, pp. 79-82, November 2011.

[5] P. Musa, M. Lamsani, and N. Nuryuliani, Rancang Bangun Pengendali Pintu Automatis Berdasarkan Ciri Wajah Menggunakan Metode Euclidean Distance Dan Fuzzy C-mean, STMIK MDP Palembang, Indonesia 2012

[6] P. Viola, and M. Jones, Robust Real-time Object Detection, IJCV, Vancouver, Canada, 2001.