Proc. of the 1st International Conference on Electrical, Communication and Computer Engineering (ICECCE)
24-25 July 2019, Swat, Pakistan

Transfer Learning of a Neural Network Using Deep Learning to Perform Face Recognition

Suleman Khan
Dept. of Electrical Engineering
Bahria University Islamabad Campus
Islamabad, Pakistan
khansuleman444@gmail.com

Ehtasham Ahmed
Dept. of Electrical Engineering
Bahria University Islamabad Campus
Islamabad, Pakistan
ehtashamahmedea@gmail.com

M. Hammad Javed
Dept. of Electrical Engineering
Bahria University Islamabad Campus
Islamabad, Pakistan
hammadjavaid0000@gmail.com

Syed A A Shah
Dept. of Electrical Engineering
Bahria University Islamabad Campus
Islamabad, Pakistan
sasimalishah@gmail.com

Syed Umaid Ali
Dept. of Electrical Engineering
Bahria University Islamabad Campus
Islamabad, Pakistan
sumaid.buic@bahria.edu.pk

***Abstract*— Deep Learning can be used to solve different problems in image analysis and pattern recognition. Face recognition is it's one of the applications. The use of face recognition to solve different social problems like personal authentication and security is increasing rapidly. In past different face recognition techniques have been introduced like Fisherfaces, Eigenfaces, and LBPH. These methods have low accuracy, so neural networks are used. The deep learning based neural networks are the most accurate. The parameters of the network can be tuned to achieve high accuracy. Face recognition requires a data base of individuals to train a neural network using deep learning. The trained network is then capable of recognition. The framework of the facial recognition process using transfer learning of the pre-trained neural network is described in this paper. We have used AlexNet for recognition which is a pre-trained convolutional neural network. Transfer learning of this pre-trained network has given accuracy of 97.95%. It can be used to classify 1000 different people. For recognition purpose, it requires a large database at least one thousand images per class. The accuracy is although high as compared with techniques mentioned above.**

Keywords*— Deep Learning, Face Recognition, Convolutional Neural Networks, Image Processing, AlexNet, Personal Authentication, Security.

I. INTRODUCTION

Facial recognition's first signs were found in psychology in 1950s. In engineering literature, Woodrow Wilson Bledsoe first proposed it in 1960's. He is known as the father of face recognition. He performed the classification of facial images. In 1970's Takeo Kanade started research work of making automatic machine reorganization. After the popularity of this technology different review papers were published in 1995. Recognition from video was still a challenge at that time, but in the last 10 years, this technology has grown rapidly [1] now deep learning based methods are being used for face recognition and results are impressive.

Facial recognition is an application of pattern recognition in image processing. It can be used to solve different problems of society. In past different algorithms have been introduced to accomplish this task. These algorithms have the deficiency of accuracy. Due to which it was not possible to use this technology commercially. Deep learning based neural networks are the new approach [2]. They can be used for face recognition. The results of the face recognition through neural networks have high accuracy than other algorithms used in the past. We have used AlexNet which is a pre-trained convolutional neural network (CNN) [3]. Transfer learning of this CNN is done to perform face recognition. The accuracy achieved by AlexNet is 97.95% after training it for four different classes. Each class contains 1000 images. It can be used to classify up to 1000 different people.

Facial recognition is the process to identify a person in an image or a video. This identification is done through a stored database of more than one individual. It is different from face verification. Face verification requires the database of a single person [4]. Face recognition has an importance because of its various applications. It is used in biometric authentication of people and security. Facial recognition based authentication is more convenient than other methods. Other methods such as personal identification number (PIN), password and cards have many issues. They can be stolen or forgotten, and sometimes it is difficult to remember them. There is no such problem in biometric based technique especially face recognition. Face recognition is an easy and economical solution to these problems [5].

This paper elaborates about the whole face recognition process using convolutional neural network (CNN). CNN is a network having multiple layers. This network has features learning capacity. Feature can also be described as patterns. We can use CNN to get trained on our data by manipulation in its layers. This manipulation is known as Transfer Learning. Training of CNN falls in the category of deep learning. Deep learning is the transformation and manipulation of multiple layers. Through all this process the task is to perform face recognition. In past the techniques used for face recognition are as follows:

1. *Fisherfaces*

Use feature vectors to describe and discriminate different classes [6].

2. *Eigenfaces*

Use characteristics of an individual for face recognition [6].

3. *Local Binary Pattern Histogram (LPBH)*

Use of local binary pattern for feature description and making feature vector to recognize faces [7].

978-1-7281-3825-1/19/$31.00 ©2019 IEEE

These methods have problems like face expressions and light intensity. Also have low of efficiency and accuracy. These problems can be overcome using a CNN model for face recognition. There is only need of making a good database and training network with right options. Database creation and options to train the network are explained in section II. The CNN model used is AlexNet. AlexNet is a 25 layer deep CNN model trained on 1.2 million images of high resolution. It has an overall low error rate of 15.3%. There are five convolutional and three fully-connected layers in AlexNet. It can classify up to one thousand different objects. A large database is required for powerful learning and accurate results. AlexNet requires at least 1000 images of a person for recognition purpose. Using more images will help to improve the accuracy. The capacity of recognition can be controlled in AlexNet by manipulation in its layers. The transfer learning of this network to accomplish the task of face recognition is easier with the help software MATLAB. The implementation of a face recognition process is done and explained using MATLAB.

The features or face pattern of a person is matched with the trained network to perform face recognition. In a human face there are 80 nodal points, which makes face features or patterns. These are the basis of difference between individuals. CNN is trained to learn these features. Some of the features are:

- *Distance of both eyes*
- *Jaw line length*
- *Width of Nose*
- *Depth of eye sockets*
- *Cheek bones*

CNN recognizes these features because they are more prominent in a face. Face recognition with more features will give more accurate results.

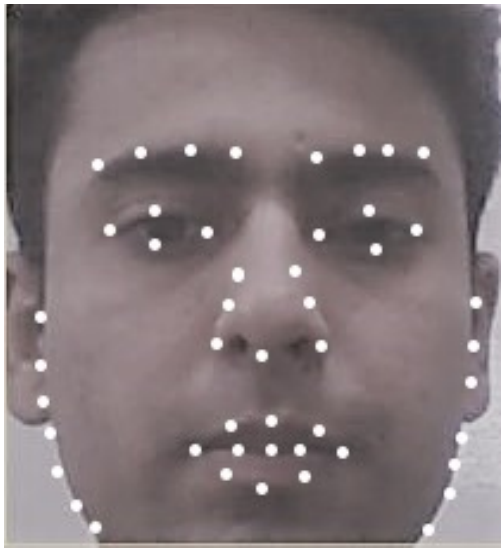

Figure 1. Human Face Nodal Points

Figure 1. Shows some of the nodal points of a human face. These can be verified through face pattern generated by Histogram of Oriented Gradient technique.

The face recognition process consists of different steps that are given in the following block diagram.

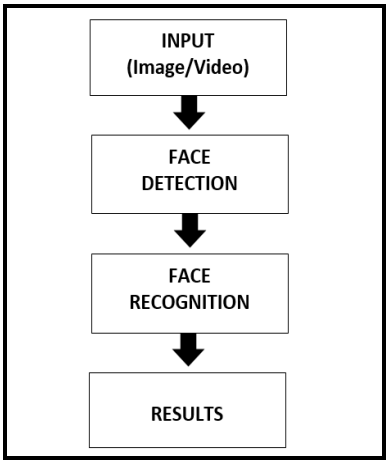

Figure 2. The Block diagram of Steps of the Face Recognition Process

The first step is to get the input in the form of image or video having the face of person to be recognized. Then, from this input the required thing is the face, so face detection is performed. The detected face is then compared through the database using trained CNN model. The face recognition system shows the results in the end.

Face detection is finding the face in an image. It can be done using the Haar-Cascade based face detector [8]. It is an efficient detector and have an accuracy of 98%. Haar-Cascade uses a series of Haar-like features [9]. Haar-like features are similar to a face image in grey scale as shown in Figure 3 and 4.

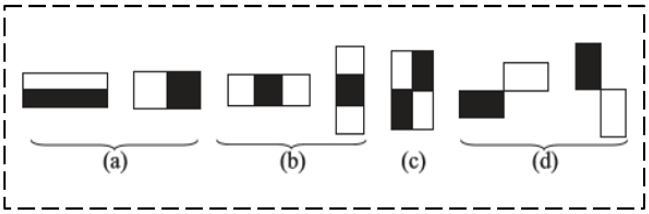

Figure 3. Rectangular Haar-like Features for Face Detection

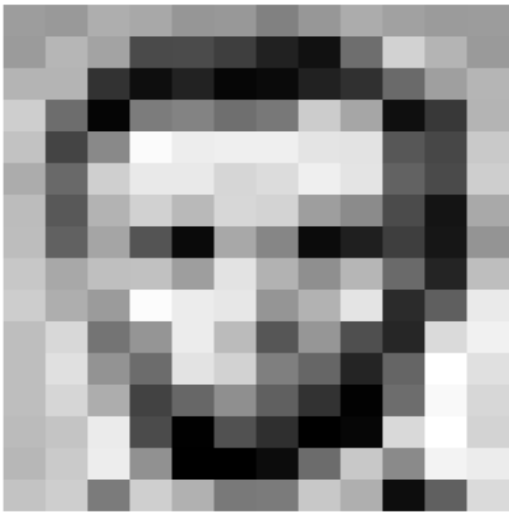

Figure 4. A Face Image in Grey Scale

These feature windows overlap the input image. Where the threshold for the maximum features is maintained the area is declared as the region of interest that means the face is detected. Threshold is decided while training of the

cascade. The training is done using two types of images. One is having face and other having non-face items.

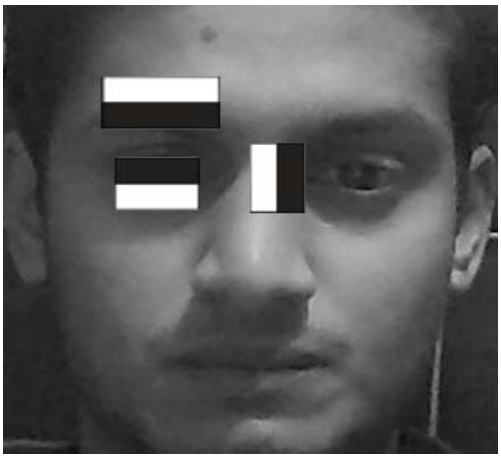

Figure 5. Applying Haar Features on Face Frontal View

If we apply a Haar-like feature on an eye. This feature has a certain threshold. After superimposing this feature over an eye, the sum of pixels is computed. The sum of brighter pixels is subtracted from the sum of darker pixels. The output is a threshold and if this value satisfies the threshold of eye feature, it declares that an eye is found. Similarly, all the other features are detected in the form of cascade and if all passes the result is detected face.

The detected face is then recognized. For recognition, a CNN AlexNet is used. AlexNet is a pre-trained model [10]. Transfer learning of this network is required to use it for facial recognition [11]. This CNN has 25 different layers that include 5 convolutional and 3 fully-connected layers. Convolutional layers have filters. These filters are used to learn features and then they are used for face recognition. Fully-connected layers are connected to all previous layers and they decide the result. The input images accepted by this network is of size 227x227x3. It is trained at the same resolution images.

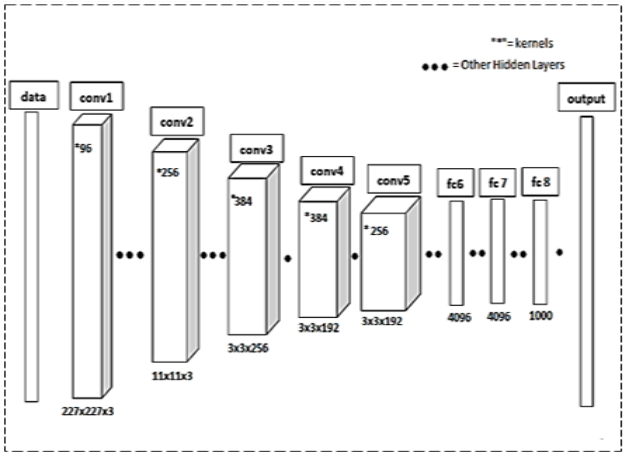

Figure 6. AlexNet Layers.

In Figure 6 convolution and fully-connected layers of AlexNet are shown. The size of inputs and kernels(filters) present in each convolution layer are also mentioned. The manipulation in these layers is performed to train the network for recognition of new data.

II. METHODOLOGY

Face recognition system is implemented according to the block diagram given in the section I (Figure 2). First a face detector is implemented on MATLAB using Haar-Cascade. The cascade used has 3099 different features. The detector can detect the frontal face view with an accuracy of 98%. This face detector helps in making the database and performing face recognition. The next step is to make a database of people. AlexNet is trained on this database to perform recognition. AlexNet requires a large database having at least 1000 variant face images. This is the reason the new database is generated instead of using the database available on the internet. All the images should be of size 227x227. We first gathered the images. Then using the Image Batch Processor tool of MATLAB got the face images of required size.

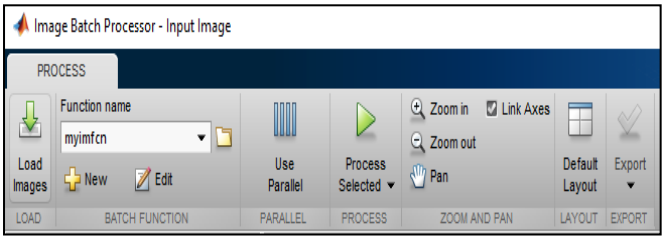

Figure 7. Image Batch Processor Toolbox

Using this toolbox of MATLAB, we created the database. All the collected images were saved in a folder and then loaded. On loaded images face detection and resizing has been implemented using a function. The results of loaded images are then saved. The images of a single person are known as a class.

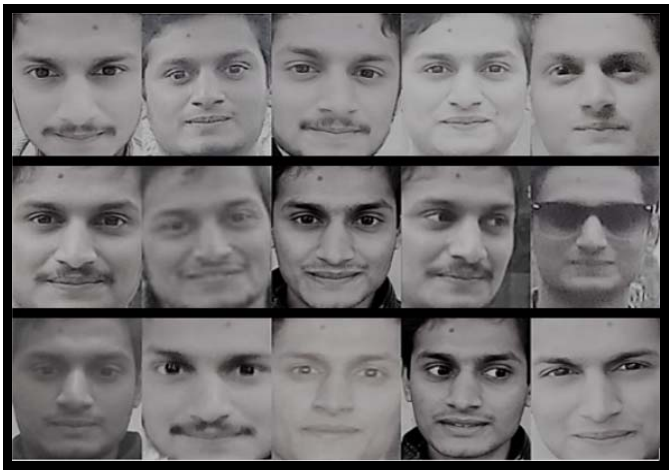

Figure 8. Data in a Class

Figure 8. Shows the images in a class. Images have different lighting conditions, face expression and posture. Variant images are used for training to increase accuracy of the network. AlexNet is then trained on the classes present in the database. The first step in training is labeling the data. Labeling means giving a name or tag to a specific group of images. It is differentiating the groups of images. In face recognition generally name of a person is considered as a label. After labeling images the images are used for the transfer learning of AlexNet. For transfer learning the last

two layers are manipulated. Layer number 23 and 25 are changed to recognize the required number of faces. After this, the training option like Epochs, Batch size and the Initial Learn Rate are given, and training is initialized. Epoch is the number of times the training data should move through a network. Data is large so it is trained in batches.

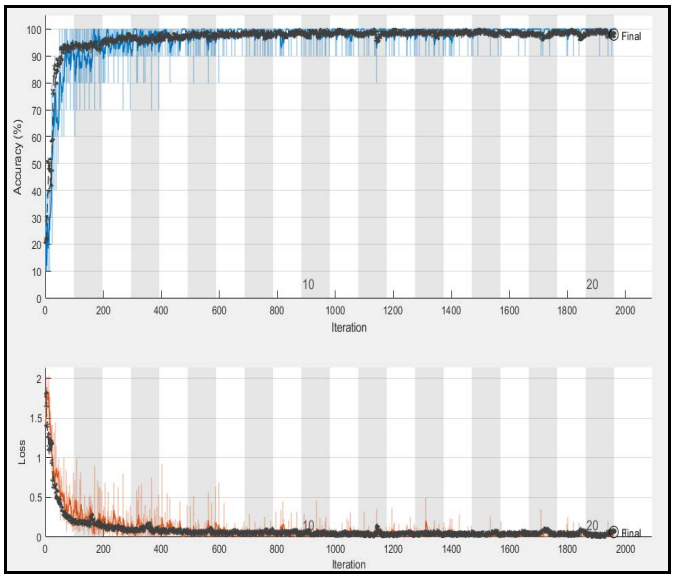

Figure 9. Graph of accuracy and loss of AlexNet after training it for a new database

The database is trained for epoch equal to 20 and batch size of 10. The accuracy achieved is 97.95%. The total number of iterations were 1960. After training a training file is generated and is saved to use for the task of recognition. For recognition first face is detected. Detected face is then recognized by loading the trained model and comparing the features with the input.

After training the layers in CNN recognize the faces. The initial layers recognize minor details like edges and circles. The next layers recognize larger details like parts of face (eyes, nose, lips). Last layers recognize full face. The filters present in the convolutional layer perform the recognition. Graph of accuracy and loss is shown in Fig.9.

III. Results

Figure 10 shows the result of face detection using Haar-Cascade.

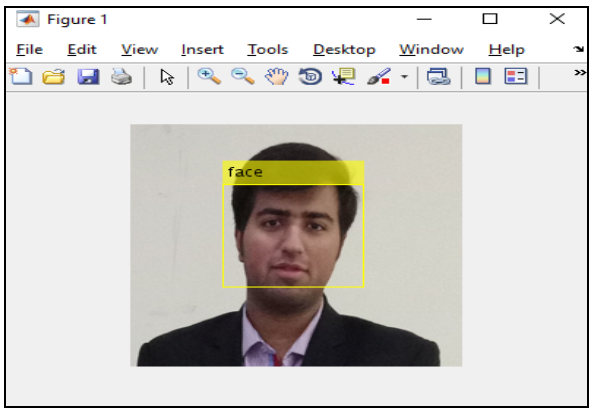

Figure 10. Face Detection

Figure 11 Shows the result of face recognition using new trained model.

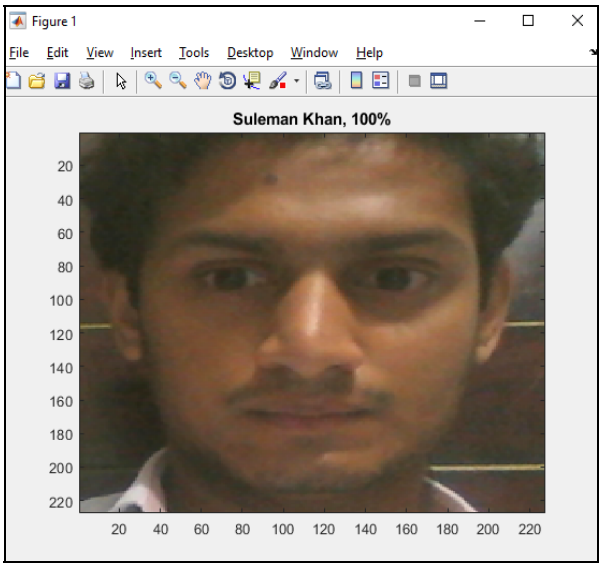

Figure 11. Face Recognition

These results are achieved using software MATLAB. For the recognition total four classes were trained. Each class having 1000 images. The output result is in the form of a label of the class and predicted score. Accuracy is 97.95% for recognition, but it can be increased by using a greater number of variant images.

IV. Conclusion

In this paper we have implemented the face recognition system using a convolutional neural network (CNN) that is AlexNet. It is a deep learned multiple layer model. We have performed the transfer learning of this network. Then trained the network on our database. In the end used this network for face recognition. A large database is required for training this network but the accuracy is high. We have used four different classes in the database for training. Each class containing 1000 images. The training options include Epochs that is equal to 20 and batch size equal to 10. The accuracy achieved by training the network using these options is 97.95%. We used MATLAB for implementing the system. It is a better approach than the techniques used in the past. A reliable face recognition system is proposed in this paper.

References

[1] R. Jafri and H. R. Arabnia, A survey of face recognition techniques, Journal of Information Processing Systems, vol. 5, no. 2, pp. 41-68, 2009.

[2] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. Nature, 521:436–444, 2015.

[3] Simard, D. Steinkraus, and J. Platt. Best practices for convolutional neural networks applied to visual document analysis. In Proceedings of the Seventh International Conference on Document Analysis and Recognition, volume 2, pages 958–962, 2003.

[4] Rupesh Sutara, Shahbaaz Rokadiab, Akshay Shahc, A survey on face recognition technologies and techniques, International Journal Of Technology And Computing (IJTC), ISSN-2455-099X, Volume 2, Issue 7 July 2016.

[5] Thomas Huang, Ziyou Xiong, and Zhenqiu Zhang, Face Recognition Applications. In: Li, S.Z., Jain A.K. (eds) Handbook of Face Recognition, Springer-Verlag, pp 617-683, 2011.

[6] P. Belhumeur, J. Hespanda, and D. Kriegman, "Eigenfaces versus Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, July 1997.

[7] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037-2041, Dec. 2006.

[8] P. Viola and M. Jones, Robust Real-Time Face Detection, Int'l J. Computer Vision, vol. 57, no. 2, pp. 137-154, May 2004.

[9] T. Mita, T. Kaneko, and O. Hori, Joint Haar-Like Features for Face Detection, Proc. 10th IEEE Int'l Conf. Computer Vision, 2005.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25, pages 1106–1114, 2012.

[11] Pan, S. J., & Yang, Q. A survey on transfer learning. The IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345–1359, 2010.