Face Recognition Using Deep Learning

A.Alagulaskhmi
B.Tech., M.E., Arumugham Palaniguru Arts and Science College for Women, Rajapalayam

Abstract - For real world applications like video surveillance, human machine interaction, and security systems, face recognition is of great importance. Deep learning based methods have shown better performance in terms of accuracy and speed of processing in image recognition compared to traditional machine learning methods. This paper presents a modified architecture of the Convolution Neural Network (CNN) by adding two operations of normalization to two of the layers. The operation of normalization that is normalization of the batch provided acceleration of the network. CNN architecture was used to extract distinctive facial characteristics and Softmax classifier was used to classify faces within CNN's fully connected layer. Our Face Database has shown in the experiment part that the proposed approach has improved the performance of face recognition with better results of recognition. Deep learning is an approach to perform the face recognition and seems to be an adequate method to carry out face recognition due to its high accuracy. Experimental results are provided to demonstrate the accuracy of the proposed face recognition system.

Key Words: face recognition, convolutional neural network, softmax classifier, deep learning.

1. INTRODUCTION

In recent years, artificial intelligence is developing in a rapid manner. Nowadays, it can be seen that the invention of self-driving car or self-service supermarket have been introduced. Artificial intelligence is closely linked to computer vision. Humans, by utilizing vision to adapt and to understand the environments where they are surrounded with, whilst computer vision is working on duplication of human vision but in electronically to perceive and interpret an image. Computer vision is not only working as an eye to see but it needs to react.

Within the last numerous years, numerous algorithms have been proposed for face detection. While lots of development has been made in the direction of spotting faces below small versions in lighting fixtures, facial expression and pose, dependable techniques for popularity beneath greater excessive variations have been established elusive. Face detection is essential to many face applications, together with face popularity and facial expression evaluation. But, the huge visible versions of faces, together with occlusions, massive pose variations, and excessive lighting, impose splendid demanding situations for those obligations in actual-world applications.

The illumination variation and variation in facial pose have a greater impact on the accuracy of these systems. The orientation of illumination leads to shifts in the position & shape of the shadow, changes in highlights, and also the reversal of contrast gradients. The proper orientation and alignment of the face pose lead to the extraction of facial features that are Robust to geometric variations.

Face recognition is the process of recognizing a person's face through a vision system. Because of its use in security systems, video surveillance, commercial areas, it has been an important human - computer interaction tool and is also used in social networks like Facebook. After the fast development of artificial intelligence, face recognition has attracted attention due to its meddlesome nature and since it is the main method of human identification when compared with other types of biometric methods. Face recognition can be easily checked in an uncontrolled environment without the knowledge of the subject person.

2. INTRODUCTION TO DEEP LEARNING

Deep learning is the branch of machine learning which is based on artificial neural network architecture. An artificial neural network or ANN uses layers of interconnected nodes called neurons that work together to process and learn from the input data.

Deep learning is a branch of machine learning which is based on artificial neural networks. It is capable of learning complex patterns and relationships within data. In deep learning, we don’t need to explicitly program everything. It has become increasingly popular in recent years due to the advances in processing power and the availability of large datasets. Because it is based on artificial neural networks (ANNs) also known as deep neural networks (DNNs). These neural networks are inspired by the structure and function of the human brain's biological neurons, and they are designed to learn from large amounts of data.

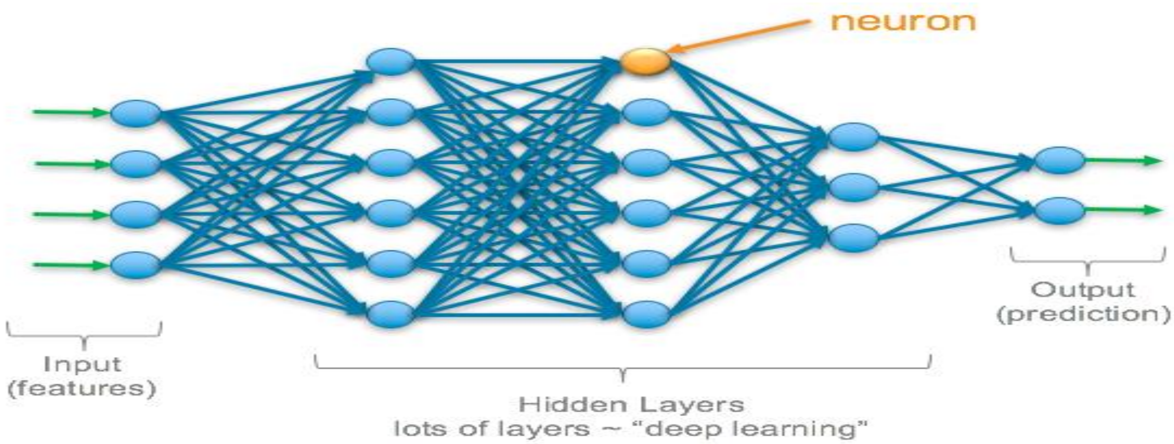

Figure 1: Deep Learning Architecture

Neuron:
Neurons are the building blocks of the nervous system. They receive and transmit signals to different parts of the body. This is carried out in both physical and electrical forms. There are several different types of neurons that facilitate the transmission of information.

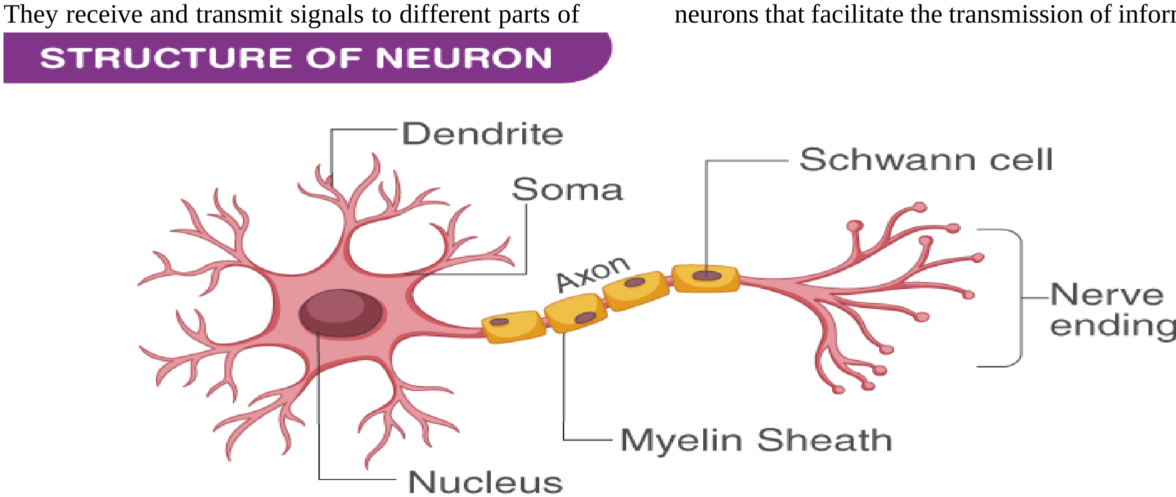

Figure 2: Structure of Neuron

3. CONVOLUTION NEURAL NETWORK

Convolutional Neural Network (CNN) is the extended version of artificial neural networks (ANN) which is predominantly used to extract the feature from the grid-like matrix dataset. For example, visual datasets like images or videos where data patterns play an extensive role.

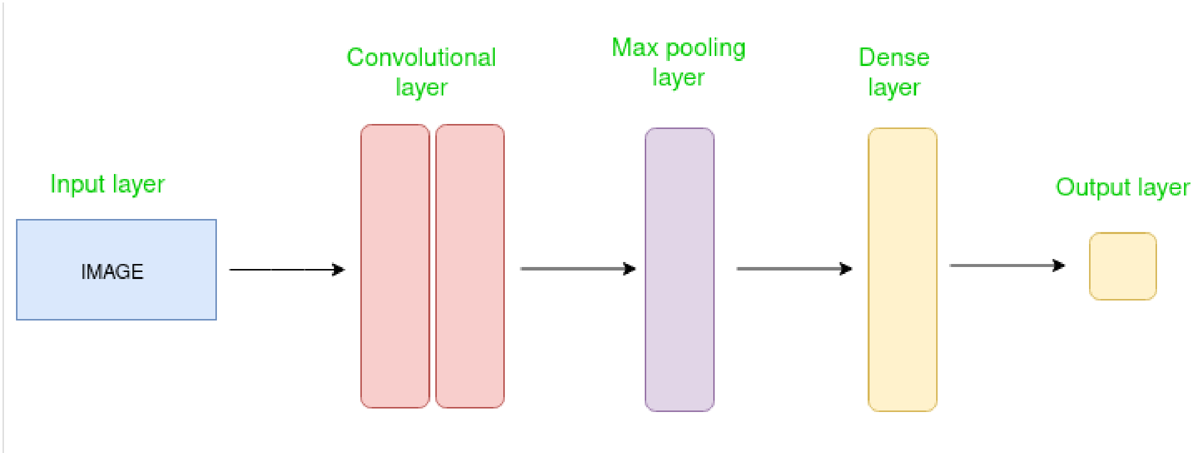

Figure 3: Architecture of CNN

3.1 Convolutional Layer

This layer is the first layer that is used to extract the various features from the input images. In this layer, the mathematical operation of convolution is performed between the input image and a filter of a particular size MxM. By sliding the filter over the input image, the dot product is taken between the filter and the parts of the input image with respect to the size of the filter (MxM).

The output is termed as the Feature map which gives us information about the image such as the corners and edges. Later, this feature map is fed to other layers to learn several other features of the input image.

The convolution layer in CNN passes the result to the next layer once applying the convolution operation in the input. Convolutional layers in CNN benefit a lot as they ensure the spatial relationship between the pixels is intact.

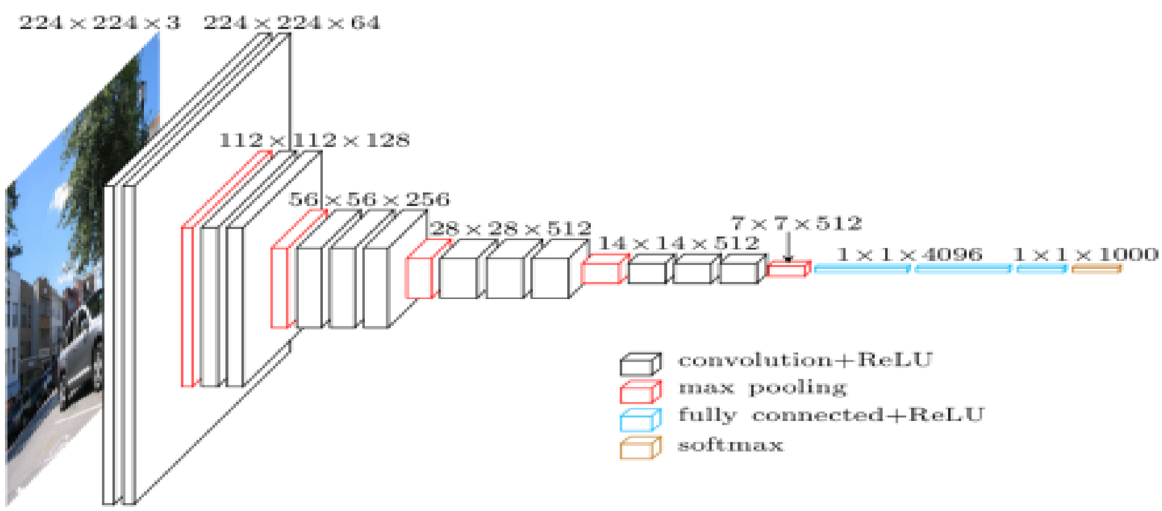

Figure 3: Convolutional Neutral Network Layer

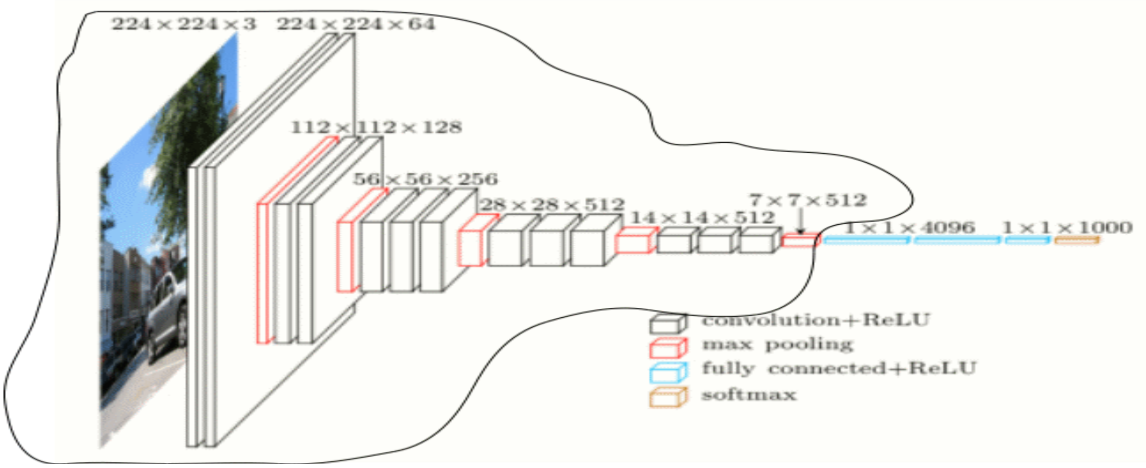

Figure 3.1: The first block is encircled in black

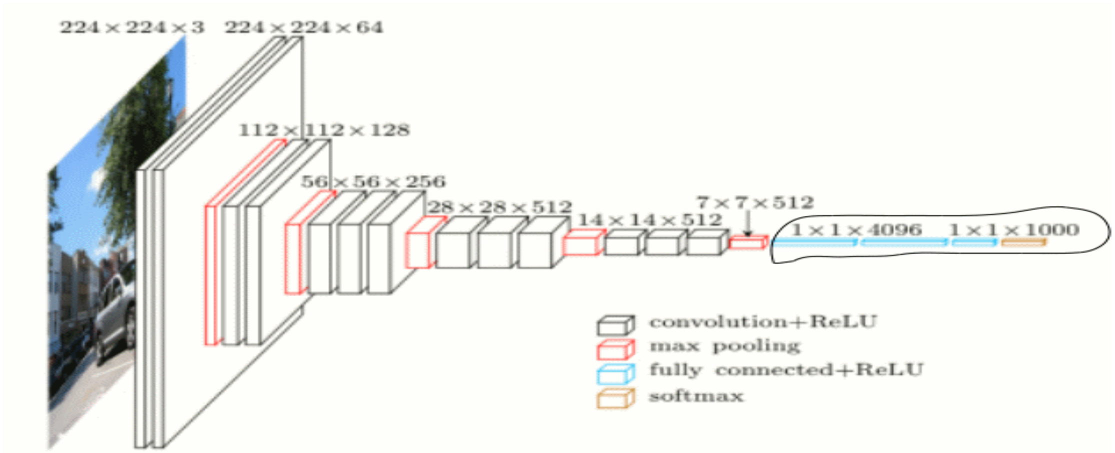

Figure 3.2 : The Second block is encircled in black

3.1.1 key component of convolutional neural network Its purpose is to detect the presence of a set of features in the images received as input. This is done by convolution filtering: the principle is to “drag” a window representing the feature on the image, and to calculate the convolution product between the feature and each portion of the scanned image. A feature is then seen as a filter**:** the two terms are equivalent in this context. The convolutional layer thus receives several images as input, and calculates the convolution of each of them with each filter. The filters correspond exactly to the features we want to find in the images.

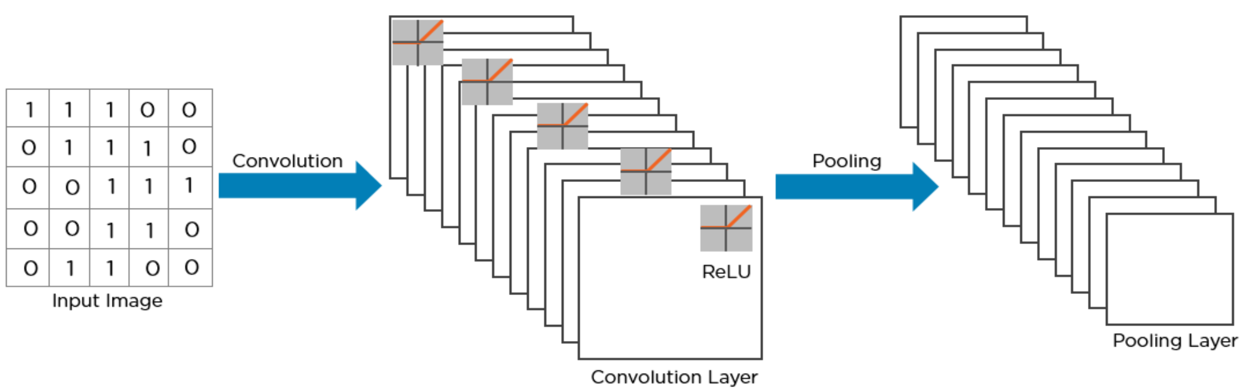

Figure 3.1.1: CNN Embedded System Vision

Fig 3.1.1, a typical CNN architecture is shown. CNN's structure contains layers of Convolution, pooling, Rectified Linear Unit (ReLU), and Fully Connected.

3.2 Pooling Layer

In most cases, a Convolutional Layer is followed by a Pooling Layer. The primary aim of this layer is to decrease the size of the convolved feature map to reduce the computational costs. This is performed by decreasing the connections between layers and independently operates on each feature map. Depending upon method used, there are several types of Pooling operations. It basically summarises the features generated by a convolution layer.

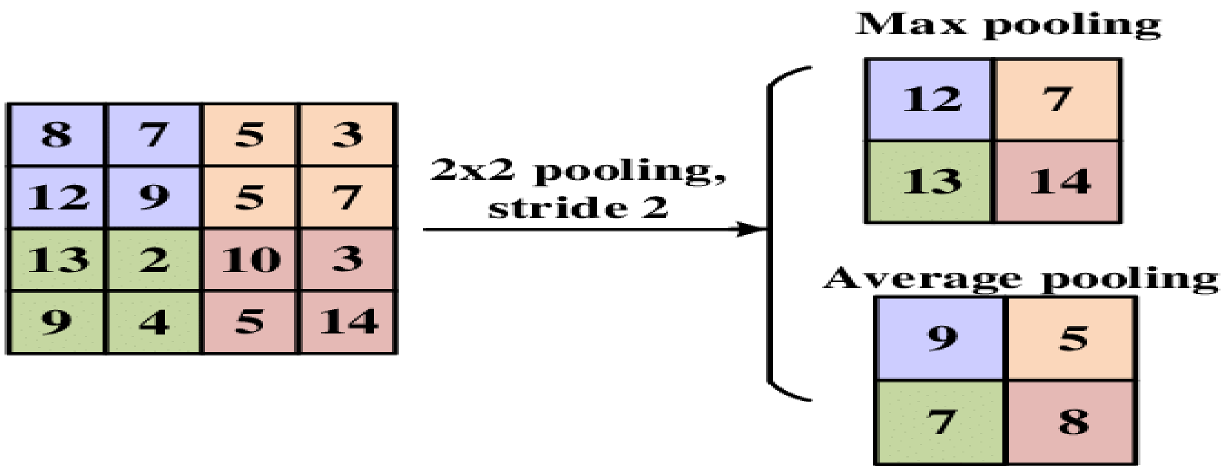

Figure 3.2: Pooling Layer

3.3 Fully Connected Layer

The Fully Connected (FC) layer consists of the weights and biases along with the neurons and is used to connect the neurons between two different layers. These layers are usually placed before the output layer and form the last few layers of a CNN Architecture.

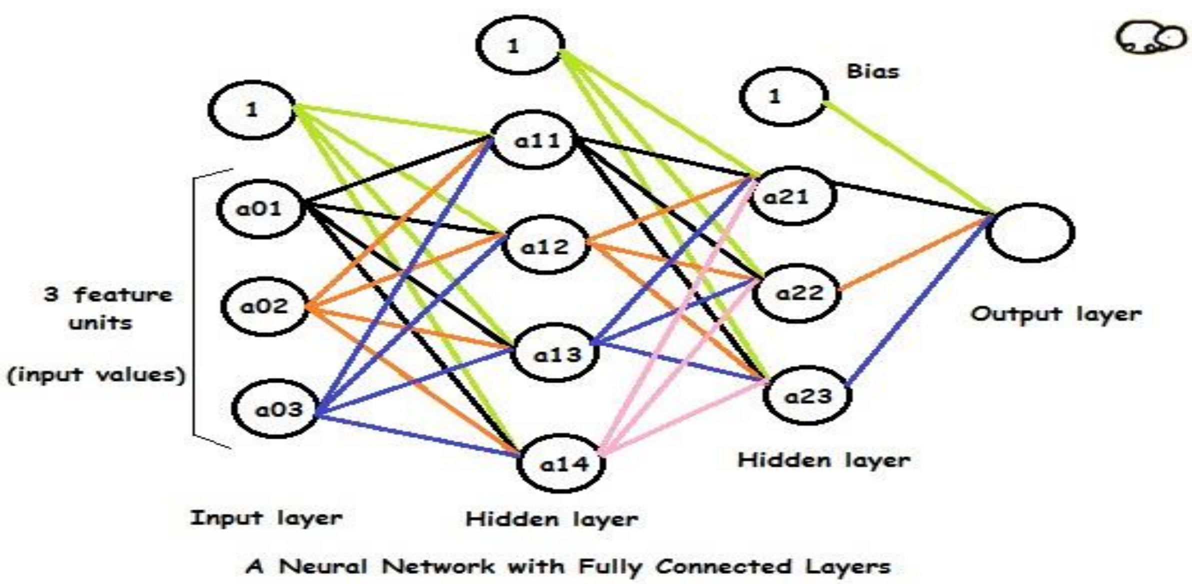

Figure 3.3: Fully Connected Layer

3.4 ReLU Layer

ReLU is a non-linear operation that involves units that use the rectifier. It is an element-wise operation which means that it is applied per pixel, reconstituting all the negative values by zero in the feature map. To understand how ReLU operates, we assume that in the literature for neural networks there is a neuron input given as x and from that the rectifier is defined as f(x)= max(0,x).

4. THE PROPOSED ALGORITHM

The block scheme of the proposed algorithm for CNN recognition is shown in Fig. The algorithm is performed in the following three steps:

1. Resize the input images as 16x16x1, 16x16x3,32x32x1, 32x32x3, 64x64x1, and 64x64x1.
2. Make a CNN structure with eight layers made up of convolutional, max pooling, convolutional, max pooling, convolutional, max pooling, convolutional, and convolutional layers respectively.
3. Use Softmax classifier for classification after extracting all the features.

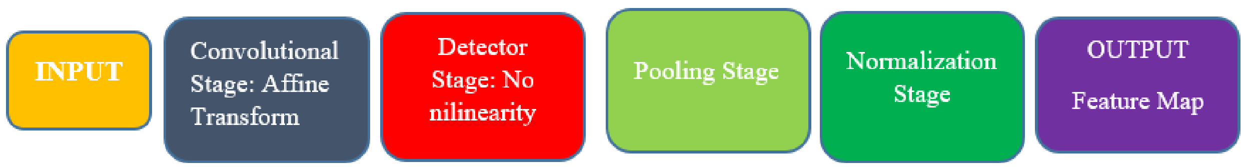

Fig 4: CNN Block Diagram

4.1 CNN Extraction Structure

In picture. 3, The structure of the proposed CNN extraction block feature is shown.

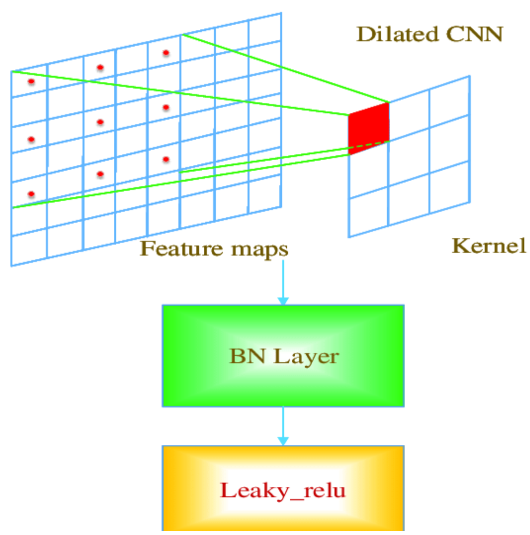

Fig 4.1: The structure of the proposed CNN extraction block Diagram

5.DATABASE

Our face database contains images of 4 people taken at different times with different lighting conditions taken between 2/03/2019 to 8/03/19. Each individual in the database is represented by hundred coloured JPEG images with cluttered background taken at 640 pixels' resolution. In these images, the average size of the faces is 150bis150 pixels. The pictures show frontal and/or inclined faces with different conditions of lighting and scale. Photograph. Fig. 4 presents some face images from our face database of different subjects.

6. EXPERIMENTAL RESULTS

We designed our CNN with the MatConvNet software tool Beta23 version. The size of each image was changed after the pre-processing stage as 16x16x1, 16x16x3, 32x32x1, 32x32x3, 64x64x1, and 64x64x3. 70% of the pictures were assigned to the training set, 30% to the test set. By making changes in image size, learning rate, batch size, and so on, we implemented various tests. For 35 epochs, CNN was trained. The performance of the proposed CNN was assessed based on top-1 and top-5 errors. Top-1 error rate checks whether the top class is identical to the target label and top-5 error rate checks whether the target label is one of your top five predictions.

7. CONCLUSION

This paper presents an empirical assessment of the CNN architecture-based face recognition system. The prominent features of the proposed algorithm are that it uses batch normalization for the outputs of the first and final convolution layers and higher accuracy rates are achieved by the network. Softmax Classifier is used to classify the faces in a fully connected layer step. Our Face Database tested the performance of the proposed algorithm.

REFERENCE

[1] S. G. Bhele and V. H. Mankar, “A Review Paper on Face Recognition Techniques,” Int. J. Adv. Res. Comput. Eng. Technol., vol. 1, no. 8, pp. 2278–1323, 2012.

[2] V. Bruce and A. Young, “Understanding face recognition,” Br. J. Psychol., vol. 77, no. 3, pp. 305–327, 1986.

[3] D. N. Parmar and B. B. Mehta, “Face Recognition Methods& Applications,” Int. J. Comput. Technol. Appl., vol. 4, no. 1, pp. 84–86, 2013.

[4] W. Zhao et al., “Face Recognition: A Literature Survey,”ACM Comput. Surv., vol. 35, no. 4, pp. 399–458, 2003.

[5] Dan Cireúan, “Deep Neural Networks for Pattern Recognition.”

[6] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, “Natural Language Processing (Almost) from Scratch,” J. Mach. Learn. Res., vol. 12, pp. 2493–2537, 2011.

[7] R. Collobert and J. Weston, “A unified architecture for natural language processing: Deep neural networks with multitask learning,” Proc. 25th Int. Conf. Mach. Learn., pp. 160–167, 2008.

[8] E. Shelhamer, J. Long, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 4, pp. 640–651, 2017.

[9] R. Xia, J. Deng, B. Schuller, and Y. Liu, “Modeling gender information for emotion recognition using Denoising autoencoder,” in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings, pp. 990–994, 2014.

[10] G. E. Hinton, S. Osindero, and Y. W. Teh, “A fast learning algorithm for deep belief nets,” Neural Comput., vol. 18, no. 7, pp. 1527–1554, 2006.

[11] Y. Bengio, “Learning Deep Architectures for AI,” vol. 2, no. 1, 2009.

[12] Y. LeCun, “Backpropagation Applied to Handwritten Zip Code Recognition,” Neural Comput., vol. 1, no. 4, pp. 541– 551, Dec. 1989.

[13] A. Krizhevsky, I. Sutskever, and H. E. Geoffrey, “ImageNet Classification with Deep Convolutional Neural Networks,” Adv. Neural Inf. Process. Syst. 25, pp. 1–9, 2012.

[14] A. Uçar, Y. Demir, and C. Guzelis, “Object Recognition and Detection with Deep Learning for Autonomous Driving Applications,” Simulation, pp. 1-11, 2017.

[15] Z. Lei, D. Yi and S. Z. Li, “Learning Stacked Image Descriptor for Face Recognition,” IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 9, pp. 1685–1696, Sep. 2016.