Comparative Analysis of Face Recognition using Deep Learning

Ripal Patel, Tanuja Patgar, Kavya S
Electronics & Communication Engineering Department,
Dr. Ambedkar Institute of Technology
Banglore- 560056, India.
Email: ripalpatel315@gmail.com

***Abstract*—Most of the work in the machine learning field inclined towards deep learning. We present the comparison of convolution neural network architecture on the various dataset for face recognition. The deep learning phenomenon is used in this work to achieve higher accuracy. The CNN architecture consists of two convolution layers, two Rectified linear unit (RELU) layers, two pooling layers and finally fully connected layer. We apply this deep learning algorithm to various face dataset to check the robustness of the architecture under various challenges. The face datasets have been used in this work are Yale, ORL and Extended Yale Face database B. We have also shown extensive result analysis using different parameters.**

Keywords— Deep Learning, Convolutional neural network, Face Recognition.

I. INTRODUCTION

This Comprehension is a strategy for recognizing or checking the personality of an individual utilizing their face. Face acknowledgment frameworks can be utilized to recognize individuals in photographs, video, or continuously. Law requirement may likewise utilize cell phones to distinguish individuals amid police stops. It compares the information with a database of known faces to find a match. Facial recognition can help verify personal identity, but it also raises privacy issues.

A lot of people and organizations use facial recognition and in a lot of different places: Airport, Mobile phone makers in products like Apple, Colleges in classrooms, Social media companies on website like Facebook, Businesses at entrances and restricted areas, Retailers in stores, Marketers often consider things like gender, age, and ethnicity when targeting groups for a product or idea. Facial recognition can be used to define those audiences even at something like a concert.

II. REVIEW WORK

First, Lots of research has been published for face recognition. Much of the research in face recognition has been done using feature detection following with feature classification. In [2], the author represents the face in terms of linear combination of Eigen vectors, which is called Eigen faces. Eigen faces are feature extracted from faces and feed to the unsupervised classifier. One step ahead, rather than eigen, face is represented using Laplacian in [2][3]. Author presented the method for feature extraction. Local information of face is preserved using Laplacian. Laplacian method is linear and computationally efficient and it is build using locality preserving projection.

Cai et.al [4] present the usage of subspace learning for face recognition. Smooth subspace learning method has been proposed which is better compared to other techniques for image representation. However, this technique requires decomposition into Eigen vector matrices, which is expensive in terms of memory and time consuming too. To reduce the dimensionality, the framework has been proposed called spectral regression in [5].

Recently from past decade, progress on the development of deep convolutional neural network [6] has significantly improved the accuracy in face recognition field. Deep neural network architecture called DeepID3 is proposed. This is a combination of two popular deep neural architectures called stacked convolution and inception layers. An ensemble of the proposed two architecture is achieving good result. DeepID3 is extension of DeepID2 architecture [7]. It consists of three convolutional layers, max pooling and finally fully connected layers. Moreover, the capability of CNN has been extended to check whether it can work for very large scale face dataset in [8]. The networks are very deep comprised of long sequences of convolution layers. Most of the changes are done with size of convolution layers but in [9], author does the changes in learning function. To enhance the discrimination power of deeply learned features this approach uses joint supervision of softmax loss and center loss.

III. PROPOSED WORK

The framework of the proposed work has been shown in figure 1. In generic form of face recognition using deep learning, the task is three folded. First is the collected sample of faces have been converted into uniform form to provide into deep networks. In second stage, they fed to deep learning network to build theory own set of feature and classify them. Finally, in third stage the outcome is label to which class the input face image belongs.

Deep learning Network: In this proposed approach, we have used convolutional neural network. The first layer in a CNN is always a Convolutional Layer. First thing to make sure you remember is what the input to this convolution layer is. Like we mentioned before, the input is a 32 x 32 x 3 array of pixel values. Now, the best way to explain a conv layer is to imagine a flashlight that is shining over the top left of the image. The second layer in the structure is the Rectified Linear Unit (ReLU). The rectified linear activation function is a piecewise linear function that will output the input directly if is positive, otherwise, it will output zero. It has become the default activation function for many types of neural networks because a model that uses it is easier to train and often achieves better performance.

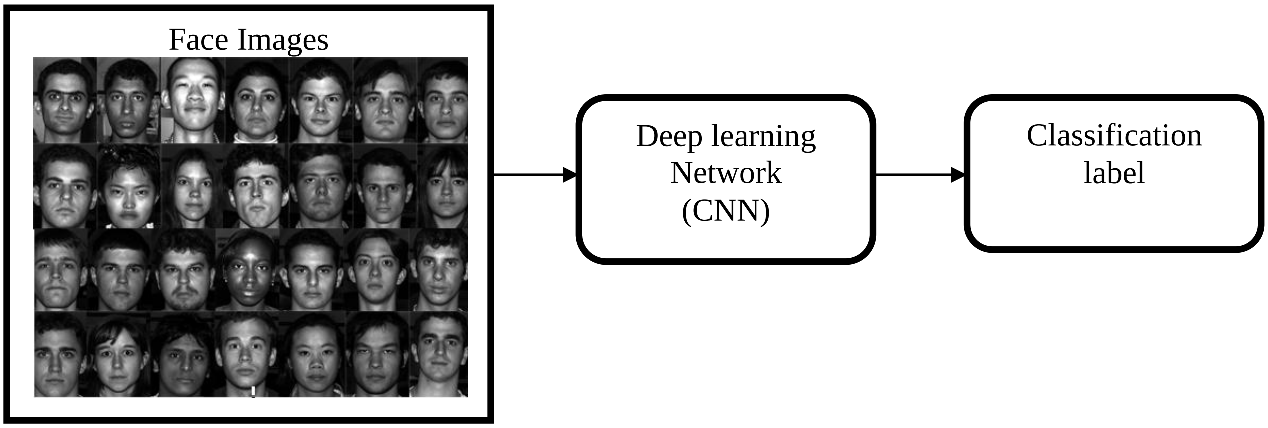

Fig. 1 General Framework for Proposed approach

Convolutional layers are often interweaved with pooling layers. In particular, there is a kind of layer called a max-pooling layer that is extremely popular. A max-pooling layer takes the maximum of features over small blocks of a previous layer. The output tells us if a feature was present in a region of the previous layer, but not precisely where. Finally, after detecting these high level features, the icing on the cake is attaching a fully connected layer to the end of the network.

For this approach, we have used the structure as shown in figure 2.

IV. Results & Discussions

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

A. *Datasets*

The Yale database

Yale dataset [10] contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink.

Fig.3 Snapshot from the Yale Dataset

ORL database

ORL dataset [11] contains ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).

Fig.4 Snapshot from the ORL Dataset

Extended Yale Face Database B

The extended Yale Face Database B [12] contains 16128 images of 28 human subjects under 9 poses and 64 illumination conditions. The data format of this database is the same as the Yale Face Database B.

Fig.5 Snapshot from the Extended B Yale Dataset

B. *Training Strategy*

The proposed approach uses 70% of whole dataset to train the CNN and 30% of images used for training for all dataset. To check the performance of the system with different number of training images we have check with other strategy like 60-40 and 80-20.

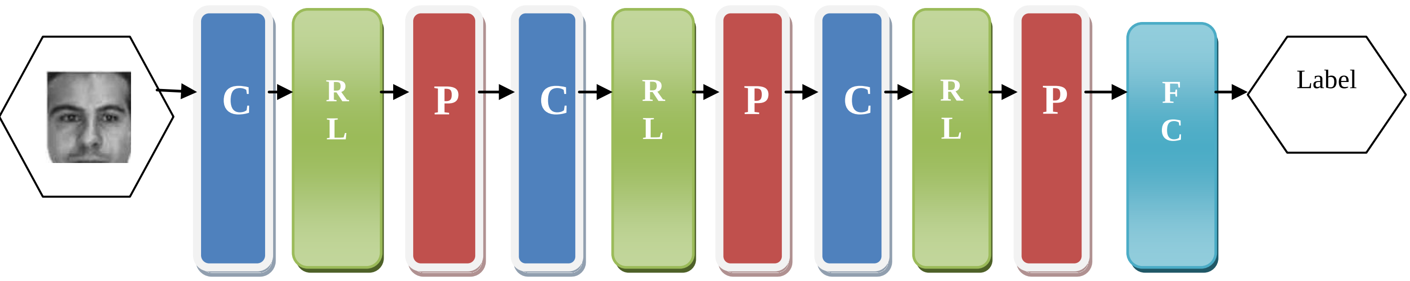

Fig. 2 The CNN architecture using for face recognition experiments. C: The convolution layer,P: The max-pooling layer, RL: The ReLu activation layer, FC: The fully connected layer

C. Results

In the following we first test several variations of the proposed models and training data using the LFW dataset to test performance.

The proposed approach was performed on 2.50Ghz Intel core i5 64 bit processor of a 64 bit operating system having RAM of 8 GB, MATLAB 2018a with 2GB of graphic card.

Table 1 shows the classification accuracy on various face dataset used in this paper. For two different size of input images has been fed to the CNN and results have been shown for both the cases.

In this paper, we have also tried the various pooling technique to analyze the effect on accuracy. Table 2 shows the classification accuracy for various pooling technique on Yale extended B face dataset. Table 3 shows the classification accuracy for the various training strategies.

TABLE I. PROPOSED APPROACH RESULTS ON VARIOUS DATASETS.

Input image size	ORL	Yale	Extended Yale B
32x32	92.50	88.89	47.24
64x64	93.75	90	53.31

TABLE II. CLASSIFICATION ACCURACY WITH VARIOUS POOLING TECHNIQUE

Pooling Techniques	Extended Yale B
Max pooling	57.31
Average Pooling	49.31

TABLE III. CLASSIFICATION ACCURACY WITH VARIOUS TRAINING STRATEGY

Training strategy	Extended Yale B
60:40	54.66
70:30	57.31
80:20	70.39

V. CONCLUSION

Due to high-security concerns the recent technologies are evaluating various types of biometrics system. The face recognition system is one of the up roaring biometrics technology which has gained maximum interest. The proposed technique is using deep learning algorithm, which is prominent technology used face recognition. The proposed face recognition system gives appropriate 93.75% accuracy. In future, accuracy can be enhanced by incorporating various features and by integrating another classifier in the system.

REFERENCES

[1] Turk, M. A., & Pentland, A. P. (1991, June). Face recognition using eigenfaces. In Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 586-591). IEEE.

[2] He, X., Yan, S., Hu, Y., Niyogi, P., & Zhang, H. J. (2005). Face recognition using Laplacianfaces. IEEE Transactions on Pattern Analysis & Machine Intelligence, (3), 328-340.

[3] Cai, D., He, X., Hu, Y., Han, J., & Huang, T. (2007, June). Learning a spatially smooth subspace for face recognition. In 2007 IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-7). IEEE.

[4] Cai, D., He, X., & Han, J. (2007, October). Spectral regression for efficient regularized subspace learning. In 2007 IEEE 11th international conference on computer vision (pp. 1-8). IEEE.

[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[6] Sun, Y., Liang, D., Wang, X., & Tang, X. (2015). Deepid3: Face recognition with very deep neural networks. arXiv preprint arXiv:1502.00873.

[7] Sun, Y., Wang, X., & Tang, X. (2015). Deeply learned face representations are sparse, selective, and robust. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2892-2900).

[8] Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015, September). Deep face recognition. In bmvc (Vol. 1, No. 3, p. 6).

[9] Wen, Y., Zhang, K., Li, Z., & Qiao, Y. (2016, October). A discriminative feature learning approach for deep face recognition. In European conference on computer vision (pp. 499-515). Springer, Cham.

[10] Tan, X., & Triggs, B. (2007, October). Enhanced local texture feature sets for face recognition under difficult lighting conditions. In International workshop on analysis and modeling of faces and gestures (pp. 168-182). Springer, Berlin, Heidelberg.

[11] Liu, C., & Wechsler, H. (2003). Independent component analysis of Gabor features for face recognition. IEEE transactions on Neural Networks, 14(4), 919-928.

[12] Tan, X., & Triggs, B. (2007, October). Enhanced local texture feature sets for face recognition under difficult lighting conditions. In International workshop on analysis and modeling of faces and gestures (pp. 168-182). Springer, Berlin, Heidelberg.