© IJEDR 2019 | Volume 7, Issue 4 | ISSN: 2321-9939

Face Recognition using Deep Convolution Neural Networks

[1]Nishtha Jatana
[1]Assistant Professor
[1]MSIT

***Abstract* - In this paper, the evaluation of the performance of Deep CNN is being carried out in comparison to CNN for Face Recognition. This paper also includes Face Recognition of twins. The average accuracy for Face Recognition using Deep CNN and CNN has been shown. The experiments were carried out on ORL Database which contains 40 different classes of people with each class containing 10 images of the same person. The results show that Deep CNN performs better than CNN, reaching an average accuracy of 95 %.**

***keywords* - Deep Convolution Neural Networks, Face Recognition, Image Processing, Deep Learning**

I. INTRODUCTION

In recent years, deep convolution networks have achieved great success in the area of computer vision. Face recognition has been one of its extensively researched and most interesting applications. The significance of face recognition is due to its technical challenges and wide potential application in video surveillance, identity authentication, multimedia applications, home and office security, law enforcement and different human-computer interaction activities.

In recent years, the method of Convolution Neural Networks have achieved great results in the field of face recognition.

To get much higher accuracy for detecting people, we can add more layers to the Convolution Neural Network, so that it can learn more complex features. Not only different faces but also it can distinguish between twins having the quite similar faces. It could recognize men clean shaven or with beard with much higher accuracy

This addition of more layers to CNN to learn complex features is called as Deep CNN.

So, in the paper, in section II we briefly discuss three research areas related to our work: deep learning, face recognition and CNN. In section III we discuss about our proposed approach and CNN architecture used. In section IV we give the experimental analysis and results over ORL face database, and finally conclude in section V.

II. RELATED STUDIES

A. Deep Learning

Deep learning techniques are also seen as a part of machine learning methods based on learning multiple levels of representation and abstraction that helps to make sense of data such as images, sound, and text. Deep learning replaces handcrafted feature extraction with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction. [8]

The recent upsurge of deep learning techniques was set off in 2006, but deep learning has been here for quite a long time. In previous few years, deep learning has shown exceptional performance in natural languages, speech recognition and computer vision. Deep learning nowadays is generally centered on multilayered neural networks.

The deep neural architectures are: feedforward neural networks, recurrent neural networks and convolutional neural networks. Feedforward networks sends unstructured information from one end called input to the other end called output; so they are called feedforward. The feedforward network approximates some function f by defining a mapping of $y=f^*(x; \theta)$ and then learning the value of the parameters θ that best approximates f, so they are seen as universal function approximators. Recurrent neural networks model dynamics over time (and space) using self- replicated components. RNNs are specialized for processing sequential data. They preserve some amount of memory, and can save long-term dependencies. RNNs are powerful computational machines – they can approximate any program. A convolutional neural network has trainable filters and local neighborhood pooling operations applied alternatingly on the input images which results in a hierarchy of increasingly complex features. The pooling operations down-sample the input representation and enlarges the input patterns. CNNs take advantage of the repetitive local input patterns across time and space, so they are translation-invariant – the capability found in visual cortex of a human. Local input patterns are small data slices, of distinct size, e.g., a group of pixels in an image.

B. Face Recognition

In previous years, the performance of face recognition algorithms has increased a great deal. The significance of face recognition is due to its technical challenges and wide potential application. The first popular face recognition technique is Eigenface (Principal Component Analysis). It can be pictured as a single layer linear model. Fisher face (Linear Discriminant Analysis) is also a single layer linear model. Laplacian face (Locality Preserving Projection) also used linear features. Then, many handcrafted local nonlinear feature-based methods emerged, such as Local Phase Quantization (LPQ), Local Binary Patterns (LBP), and Fisher vectors. These hand-crafted features achieved excellent face recognition performance, however it decreased considerably in unconstrained environments where the face images cover intra-personal variations like, pose, illumination, expression and occlusion as shown in Labeled Faces in the Wild (LFW) benchmark[9]. In last few years, deep learning methods, especially CNN has achieved very impressive results on face recognition in unconstrained environment. The main benefit of CNNs is that all the processing layers, even the pixel level input have configurable parameters that can be learned from data. This averts the necessity for hand crafted feature design, and replaces it with supervised data driven learning of features. CNN learning based features are more resilient to complex intra-personal variations. CNN methods, have attained the best three face recognition rates on the FRUE benchmark database LFW (Labeled Faces in the Wild) [9].

In deep learning, a convolution neural network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks, most commonly applied to analyze visual imagery. CNNs use a variation of multilayer perceptrons designed to require minimal preprocessing. They are also known as shift invariant or space invariant artificial neural network (SIANN), based on their shared weights architecture and translation invariant characteristics. Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that they cover the entire visual field. CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage. They have applications in image and video recognition, recommender system and natural language processing.

III. Proposed Method

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

We have used Deep Convolutional Neural Networks for Face Recognition. Other algorithms such as Principal Component Analysis (PCA), Local Binary Patterns Histograms (LBPH) and K–Nearest Neighbour (KNN) have shown accuracies around 75%.

A CNN consists of generally four layers - a convolutional layer, a RELU layer, a fully connected layer and a pool layer.

An input image of size 32 x 32 has been considered.

A convolutional layer is used to solve the problem that if an image is not appearing in the centre of a window then the neural network will not recognize the image. Convolutional layer, due to its translational invariance characteristic helps in resolving this problem.

RELU layer is the activation function max(0,x).

A fully connected layer like a neural network has its each node being connected to every node in the previous layer. The fully connected layer outputs the class probabilities

The pool layer performs downsampling, i.e. it is used to reduce the size of an array reducing the width and the height by half considering a pool layer of size 2, stride 2 as shown in Fig. 1 . As shown in figure It considers the most important features by performing max pooling taking the maximum out of every 2 x 2 array.

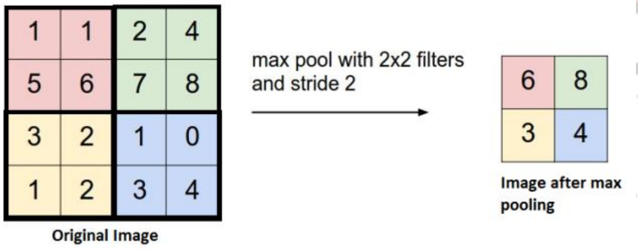

Fig. 1: Max Pooling
(Source: https://www.analyticsvidhya.com/blog/2017/05/25-must-know-terms-concepts-for-beginners-in-deep-learning/pooling/)

CNN Architecture

The architecture consists of three convolutional layers, RELU layers and three max pooling layers. Once the final pooling is done, the compressed array is fed into the fully connected layer which outputs the class probabilities.

First, an input image of size 32 x 32 is provided.

Second, the first convolutional layer, max pooling layer computes the edges of the faces.

The second convolutional layer, max pooling layer is used to compute features like nose, mouth or eyes.

Finally, the third convolutional layer, RELU layer, max pooling layer is used to detect the whole face. Then the detected face is fed into the fully connected layer generating the class probabilities and the class having the maximum probability, the face belongs to that class and hence to the person considered in that class.

IV. Experiment Analysis and results

We have used the ORL Face Database which contains 40 classes, with each class containing 10 images. 40 classes here means faces of 40 different persons. In this database the different angles of the faces such as beard or being beardless, closed and open eyes, laugh and without laugh, with and without glasses(where needed) are presented. All the face images were taken in a dark homogeneous background. In the image of a person the forehead and hair can be seen in the as shown in Fig. 2 which gives a glimpse of the face dataset. All the pictures are black and white with 112×92 pixels. The files are in PGM format. However, we have converted it into a JPG format for convenience.

Fig. 2: Glimpse of ORL Face Database
(Source: https://www.researchgate.net/figure/Part-of-the-face-images-from-the-ORL-database-for-testing_fig5_236105703)

Experiment

The dataset was divided into 3 parts - Training Set which is used to train the Deep Convolutional Neural Networks, Cross Validation Set which is used for our experimentation for bettering the results and the test set which acts as a new example and is finally used to check the results. The training set consists of 60% of the data, the cross-validation set contains 20 % of the data and the test set contains the left 20% data. Hence, 240 images belonged to the training set, 80 to the cross validation set and the left 80 to the test set.

The training data was trained using the CNN architecture above and was evaluated on the cross-validation layer. Initially, we started with one convolutional layer and max pooling layer but the results were not unto the mark, hence we added another convolutional layer and max pooling layer. The accuracy increased but it was similar to the other algorithms. Hence, we added another convolutional layer and max pooling layer which gave us fulfilling results reaching an average accuracy of 95% of recognizing faces.

The twin test was done considering the measurements of faces[2]. The twins were considered having slightly different measurements of faces[2] which makes face recognition of twins possible but not with a very good accuracy.

Results

The following results show an average accuracy of 95.1% while predicting faces of persons and an average accuracy of 83.4 % while predicting faces of twins. All these results were performed on the test set which contained 80 images of different persons. Every image had a size of 32 x 32.

Method Used for Face Recognition	Accuracy on ORL Database (%)
CNN	93.4
Deep CNN	95.1

Tab. 1: Table showing the accuracy of Deep CNN on ORL Database

Method Used for Face Recognition	Accuracy for predicting twins (%)
Deep CNN	83.4

Tab. 2: Table showing the accuracy of Deep CNN for predicting twins

A. Conclusion

Here, we proposed a way of recognizing faces with better accuracy, since Face Recognition has humungous applications today. It is being used in almost every major sector. Hence, it is important to perform Face Recognition correctly and accurately. It was found that using Deep CNN for predicting faces can have an average accuracy of 95% if a good training set has been provided. All the parameters and hyper parameters were trained on a computationally good hardware. For the future work we are working on improving the accuracy of Twin Face Recognition.

References

[1] U. Aiman and V. P. Vishwakarma, "Face recognition using modified deep learning neural network," in *IEEE*, Delhi, 2017.

[2] S. Y. Dhote, M. S. Nimbarte and A. D. Gotmare, "Differentiating Identical Twins by Using Conditional Face Recognition Algorithms," *International Journal of Innovative Research in Computer and Communication Engineering,* vol. 3, no. 9, pp. 8901-8909, 2015.

[3] S. S. Farfade, M. J. Saberian and L.-J. Li, Multi-view Face Detection Using Deep Convolutional Neural Networks, vol. 9, 2015.

[4] P. KAMENCAY, M. BENCO, T. MIZDOS and R. RADIL, "A New Method for Face Recognition Using Convolutional Neural Network," *DIGITAL IMAGE PROCESSING AND COMPUTER GRAPHICS,* vol. 15, no. 4, p. 10, 2017.

[5] M. Owayjan, A. Dergham, G. Haber, N. Fakih, A. Hamoush and E. Abdo, "Face Recognition Security System," in *Springer International Publishing*, Cham, 2013.

[6] P. J. Phillips, P. J. Flynn, K. W. Bowyer, R. W. V. Bruegge, P. J. Grother, G. W. Quinn and M. Pruitt, "Distinguishing Identical Twins by Face Recognition," *IEEE Transactions of Information Forensics and Security,* vol. 9, no. 2, pp. 285-295, 2014.

[7] D. V. Sang, N. V. Dat and D. P. Thuan, "Facial expression recognition using deep convolutional neural networks," in *IEEE*, Hue, Vietnam, 2017.

[8] Z. Wang, "The Applications of Deep Learning on Traffic Identification," [Online]. Available: https://www.blackhat.com/.

[9] G. Hu, X. Peng, Y. Yang, T. Hospedales and J. Verbeek, "Frankenstein: Learning Deep Face Representations using Small Data," in *IEEE Transactions on Image Processing*, 2017.