

# The Effect of Augmented training dataset on performance of Convolutional Neural Network in face Recognition

Tharun Kumar Reddy Kunduru

Electronics And Communication Engineering  
R.V.R & J.C College of Engineering  
Guntur, India

Thrinayani Konakanchi

Electronics And Communication Engineering  
R.V.R & J.C College of Engineering  
Guntur, India

Ganga Satya Sai Siva Matta

Electronics And Communication Engineering  
R.V.R & J.C College of Engineering  
Guntur, India

**Abstract**—To deal with the issue of human face recognition on small original dataset, a new approach combining convolutional neural network (CNN) with augmented dataset is developed in this paper. The original small dataset is augmented to be a large dataset via several transformations of the face images. Based on the augmented face image dataset, the feature of the faces can be effectively extracted and higher face recognition accuracy can be achieved by using the ingenious CNN. The effectiveness and superiority of the proposed approach can be verified by several experiments and comparisons with some frequently used face recognition methods.

**Keywords**—Transformations, augmented dataset, superiority, accuracy, performance

## I. INTRODUCTION

With rapid development in technology, Artificial intelligence (AI) is widely used for authorization, authentication and surveillance. In most of the cases we need to find the identity of the person or criminal. We have many face recognition models to achieve the objective. In particular Deep Neural Network (DNN) has been under intensive investigation in the past few years. Many researches reveals the dominance of DNN over other models of face recognition with Support Vector Machine (SVM) and Principle Component Analysis (PCA). Nevertheless, it is worth noting that most of DNN-based face recognition methods are usually developed based on a large original dataset.

Obviously, a large original dataset can provide more features of the face images, but it is usually difficult to be obtained in comparison with a small original dataset. As a result, some of the existing successful methods could lead to poor performance on the small original dataset. Besides, though a larger original dataset can bring higher accuracy of the model and stronger generalization ability of the network, the data labelling of a large original dataset is really a tedious and time-consuming work. Therefore, from a practical point of view, it is a promising topic to develop the DNN-based face recognition methods on the small original dataset.

## II. CNN MODEL DETAILS

The Convolutional Neural Network is a type of neural network. In general CNN contains convolution layer, batch normalization layer, pooling layer, flatten, dense and many other. Due to weight transfer between layers in a CNN model which does not only reduces the connection weights but also simplifies the complexity of neural network. In a CNN model there is no need of complicated works such a feature extraction and data reconstruction. Ahead of all these CNN can perform well with several image transformation methods like scaling, rotation and shifting. Basic details about some of the important layers are discussed later in this section.

### A. Convolution Layer

Convolution layer is the basic building block of a CNN model. It is kind of mathematical operation used in most image processing operations. A filter or a kernel in a conv2D layer “slides” over the 2D input data, performing an element wise multiplication. As a result, it will be summing up the results into a single output pixel. The kernel will perform the same operation for every location it slides over, transforming a 2D matrix of features into a different 2D matrix of features. Convolution layer helps in common feature extraction.

### B. Pooling Layer

Pooling Layer helps in reducing the size of the resultant feature map and reduces the probability of over fitting in the neural network. There are many types of pooling techniques like average pooling, max pooling and overlapping pooling, stochastic-pooling, and global average pooling. For instance lets discuss about max pooling, in this we will extract maximum values of the feature points and achieves better texture extraction. If a feature map of 16 x 16 is sampled by using a kernel with size of 2 x 2, the output is a feature map with size of 8 x 8. while in average pooling can extract the average value of the feature points and has the effect of maintaining the relative background.

C. Batch Normalization

Normalization is a pre-processing tool used to bring the data to a common scale, generally 0 - 1 but can be varied according to the requirement.

Batch normalization is a technique for training very deep neural networks that standardizes the inputs to a layer for each mini-batch. Batch normalization helps in removing the mean with out effecting the information in it. According to information theory, mean that is present in a data does not contain any information. Batch normalization helps in increasing the training speed, handling internal covariant shift.

D. Activation Function

The performance of a neural network depends not only on the structure, kernel values and size. It depends on adopted activation function. There are many types of activation functions, some of them are sigmoid, tanh, Rectified Linear Unit (ReLU) e.t.c. Formulae for some of the activation functions :

$$\text{Sigmoid: } f(x) = \frac{1}{1+e^{-x}} \tag{1}$$

$$\text{Tanh: } f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

$$\text{ReLU: } f(x) = \max(0,x) \tag{3}$$

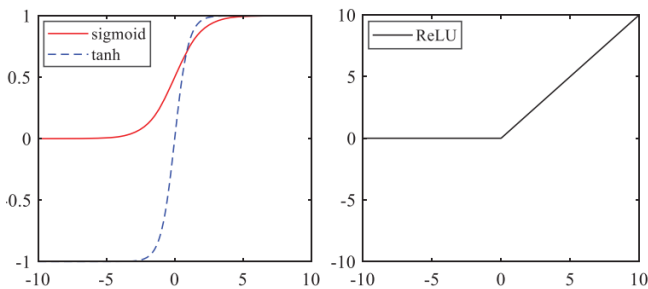


Fig. 1. Activation functions

III. FACE RECOGNITION MODEL

A. CNN model for face recognition

In this paper a CNN model is composed with 3 consecutive blocks each having Convolution Layer(C), Pooling Layer(P) and Batch Normalization Layer(N). Arranged in the form of C-N-P. Input array in the form of an numpy array passed to C1 layer. C1 is the first convolution layer which includes 32 feature maps. Each neuron is convoluted with a randomly generated kernel with a size of 3x3. P1 be the first pooling layer with kernel size of 2x2 that generates an output of 16 feature maps. N1 be the first batch normalization layer.Each element in the feature map is connected with the mean convolution kernel of the corresponding feature map in C1 layer, and the receptive fields of the elements will not overlapped with each other. C2, N2 and P2 are, respectively, the second convolutional layer, pooling layer and normalization layer, similar calculation steps with their counter parts. Two dense layers and one flatten layers is placed in between P3 and output layer. Flatten layer helps in converting feature maps into single one dimensional array and dense layer helps in building a fully

connected layers, in which every input depends on every output.

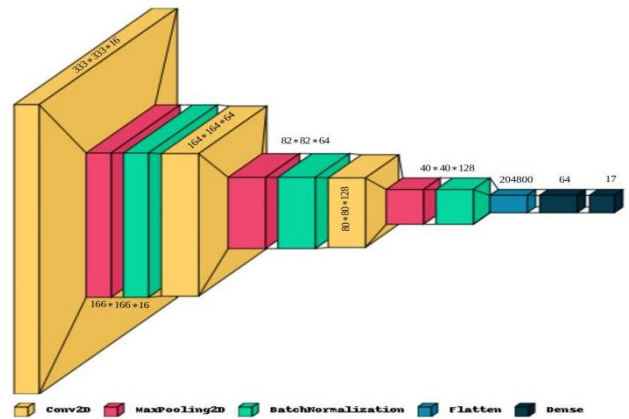


Fig. 2. CNN model

B. Dataset and Augmentation

Our dataset is a collection of 250 human face images of 17 individuals. Each individual devotes 15 face images for training and 4 images for testing. The size of each image is 336 X 336 pixels and each image is saved as the file type png and jpg. The amount of images is not abundant to train a deep neural network for accurate face recognition.

To deal with this problem, the image amount of the dataset is augmented by using data augmentation. Augmentation is the process of synthesizing new data by modifying existing data. Some of the methods of data augmentation, including horizontal flip, shift, scaling, rotation, noise addition, random erasing, and random brightness and contrast. The dataset can be tremendously by tuning the parameters of the augmentation methods. Then, the images are scaled, normalized and labelled before they are put into the face recognition system. It can be predictable that the augmented dataset can not only reduce the probability of over-fitting but also improve the robustness of the system.



Fig. 3. Augmented images from dataset

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we will see the results of our model and some analysis based on the results. The experiment is carried out using kaggle notebook environment and code is written in python. To test the effect of dataset, the model is trained using different amounts of data i.e., 250 (original dataset), 1000, 4000, 8000, and 16000. where all epochs and batch sizes are set to 1 and 16 respectively. In figure 3 we can see the few augmented images. We can observe that model accuracy is boosted with the increase in dataset size.

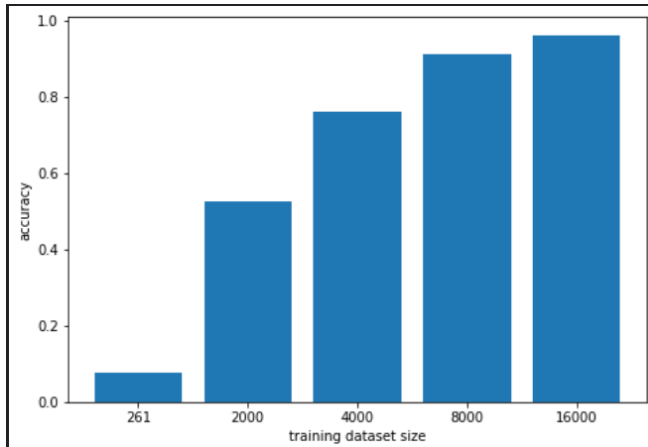


Fig. 4. Accuracy of model with different training dataset

Fig. 5.

It’s clear that performance of face recognition model will increase with increase of training samples and increase in number of epochs will increase accuracy prominently. But with the increase in epochs and training sample size the time taken to train the model will increase, so we should balance both time and accuracy by varying hyper parameter like batch size, steps per epoch and others.

To further verify the superiority of the proposed approach, the results of the approach are compared with some other face recognition methods in the literature based on ORL face dataset, including ANN (artificial neural network), PCA+ANN, PCA+SVM

TABLE I. COMPARISON OF FACE RECOGNITION ACCURACY.

Methods	Accuracy (%)
CNN	98.7
ANN	80.3
PCA+ANN	91.0
PCA+SVM	71.4

V. CONCLUSION

In this we had observed the effect of neural network by building a face recognition model using python and augmenting the dataset using different augmentation methods like flip, shift, rotate, erasing, varying brightness and contrast. Several experiments are carried out to verify the effectiveness of the augmented dataset, and the superiority of the new approach can also be confirmed in comparison with some of the frequently used face recognition methods. As dataset increase the time taken to train the model will also increase. We should select the methods of augmentation according to the condition where model will be used and type of input model will encounter.

REFERENCES

- [1] Peng Lu , Baoye Song & Lin Xu (2020): Human face recognition based on convolutional neural network and augmented dataset, Systems Science & Control Engineering, DOI: 10.1080/21642583.2020.1836526.
- [2] Gumus, E., Kilic, N., Sertbas, A., & Ucan, O. N. (2010). Evaluation of face recognition techniques using PCA, wavelets and SVM. Expert Systems with Applications, 37(9), 6404–6408. <https://doi.org/10.1016/j.eswa.2010.02.079>
- [3] Guo, L., Leng, J., Mei, W., Kong, X., Liao, Y., & Liao, H (2015). Research on human face recognition technology based on PCA and SVM. Journal of Hubei University for Nationalities (Natural Science Edition), 33(2), 193–196+214. <https://doi.org/10.13501/j.cnki.42-1569/n.2015.06.020>
- [4] Jurgen, S. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.00>