A Computer Aided System to Identify the Face shape of a Person using Machine Learning and Image Processing Techniques

H. N. Gunasinghe and T. G. I. Fernando

Department of Computer Science, University of Sri Jayewardenepura, Sri Lanka; hngunasinghe@gmail.com

Abstract

Beauty cultural activities, such as hairstyles and makeup, require the knowledge of face shape of a person which, is not always accurate or efficient and requires the expert knowledge. In this paper, we present a computer aided system based on image processing and machine learning to identify face shape automatically without having to seek an expert.

The system gets face images as input and then sends them to a pre-trained neural network model. The network handles the classification tasks and produces the predictions according to the face shape and gives results for the best matching categories. The network, designed using convolutional neural network model and a dataset was prepared after pre-processing Google image search results. The prototype system managed to perform at 98.5 % accuracy level on classifying the shapes of seven basic face shapes. The experimental results indicate that the proposed approach is a valuable approach, which can significantly support an accurate detection of face shapes with a little computational effort.

Introduction

Diamond, triangle, inverted triangle, oval, oblong, square and round face shapes are commonly referred to as basic face shapes because most of people have one of them and hairstyles, makeup and eyeglasses are based on them. For instance, a person with round shape should wear eye glasses to make the face look thin and long. In addition, a person seeking for haircut needs to know the shape of his face. Sometimes, the process of manual identification of the face shape itself is a stressful and time wasting task. Identification of face shape is not always accurate and requires expertise[1].

The purpose of this research is to design and develop a neural network to identify a person's face shape using facial images and display the results on the screen. The system also produces alerts whenever it detects an unusual shape. This system is most suitable for identifying the face shape at opticians, beauty salons and even at home.

Many researchers have carried out researches to identify various shapes using computer based techniques. Digital images are being used in them for object recognition, gender classification and facial expression detection. Recent researches indicate that the machine learning techniques can be used as a classifier of images.

An integrated approach to extract features in images from both locally and globally, using Histogram of Oriented Gradient for face classification has been proposed[2]. A linear Support Vector Machine (SVM) is used for the local feature classification using one-versus-all SVM classifiers for multiclass classification[2]. An Active Appearance Model to extract facial features, which includes face geometry information in the form of shape parameters using an SVM-based classification method to classify the main face shapes; melon seed (triangle), round and square has also been proposed[3]. A back-propagation Neural Net approach to classify irregular shapes by their brightness and convex hulls with Ten fold cross-validations being used to verify the effectiveness of the established approach has been reported[4]. An approach that is based on the shape and work to recognition in various circumstances, variable lighting conditions for the affine transformation has been described A feed-forward network with back propagation error and a self-organizing map is used to deal with interpretation of objects in the image[5].

An automatic analysis of facial appearance based problems with machine learning techniques based on Convolutional Neural Network (CNN) has been introduced[6]. CNNs can overcome the problem of feature alignment and its strong performance in difficult test images is experimentally demonstrated. A system that uses a hierarchical procedure in which it first locates roughly the eyes, nose and mouth and then refines the result by the detection of 10 different facial feature points was used. Finally, a face recognition system having a specific CNN architecture which can learn a nonlinear mapping of the area of the image into a lower dimensional subspace where it is easier to separate into different classes is proposed. The method was applied to several public face databases and could achieve better recognition rate than conventional face recognition logics[6]. Furthermore a large, deep CNN has been formed to classify 1,200,000 high-resolution images into 1,000 different classes in the IMAGEnet LSVRC-2010 contest. Neural networks, with 60000000 parameters, 650000 neurons and consists of five convolutional layers, some of which are followed by layers of max-pooling and three layers are fully connected to the end of 1000 weights of Softmax. To reduce overfitting in the layers that are fully connected, they use a new regularization method called "dropout" which has been very effective[7].

Our system facilitates the automatic classification of facial images according to the seven basic face shapes. We created a new dataset with more than 2500 frontal facial images, designed CNN and MLP networks and used high resolution images without feature extractions. The next section will further describe the methodology we used for preparing the datasets and networks. As described in the results section, we monitored the validation accuracy and computational time for both networks. It could be seen that CNN achieved the higher accuracy with lesser amount of time. We suggest applications of our work in discussion section.

Methodology

As in Fig.1 the system is divided into four sections, namely, the image acquisition, image processing, design, training and validation of neural network. The image acquisition consists of collecting suitable facial images and preprocessing includes resizing, cropping, etc. The application allows the user to input a facial photograph and provides best predictions for the face shape.

Google image search results as the main image source for the system. Images were searched based on each fundamental face shape. Human frontal, straight, full face images were chosen. Images are cropped to isolate faces, and then hair, ears, neck and background are removed from images. Each image was resized to 256 x 256 pixels without distortions and the background color set as white. Those images were in *.jpg* format.

Data augmentations were done by obtaining edge detected images in greyscale. Additionally we flipped each image horizontally, rotated 5°, 10°, 15° clockwise, rotated 5°, 10°, 15° counter clockwise, shrinked the shape into 180 and 220 pixels using Photoshop. All the images were labelled by human labellers.The face dataset consists of 4200 frontal face images and divided into 3150 (75%) images as the training set, 420 (10%) as the testing set and 630 (15%) as the validation dataset.

Our development environment consists of a PC having 4GB NVIDIA GPU. We developed our system in Ubuntu 14.04 operating system. In our method, we designed multi layer perceptron and convolutional neural network to identify the best model. Networks were trained in the DIGITS[8]software which is built on Caffe[9].

The architecture for the Multi Layer Perceptron (MLP) was chosen with three fully connected hidden layers. Each layer has rectified linear units with (1048-256-7). Softmax function was used as the activation function in the output layer with 7 neurons. Network was trained using the stochastic gradient descent (SGD) algorithm with batch size of 16. The momentum with value of 0.9 was used as the learning rule of the training algorithm. The learning rate was set to 0.0001 with step down after each 10 epochs and 0.6 step size. Network trained for 50 epochs.

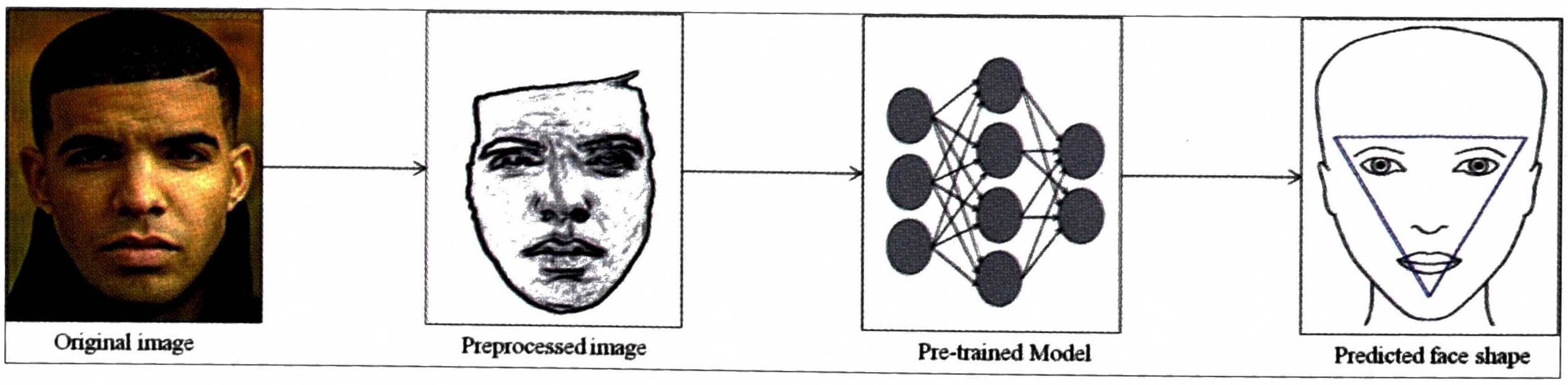

Fig. 1: Overview of the system

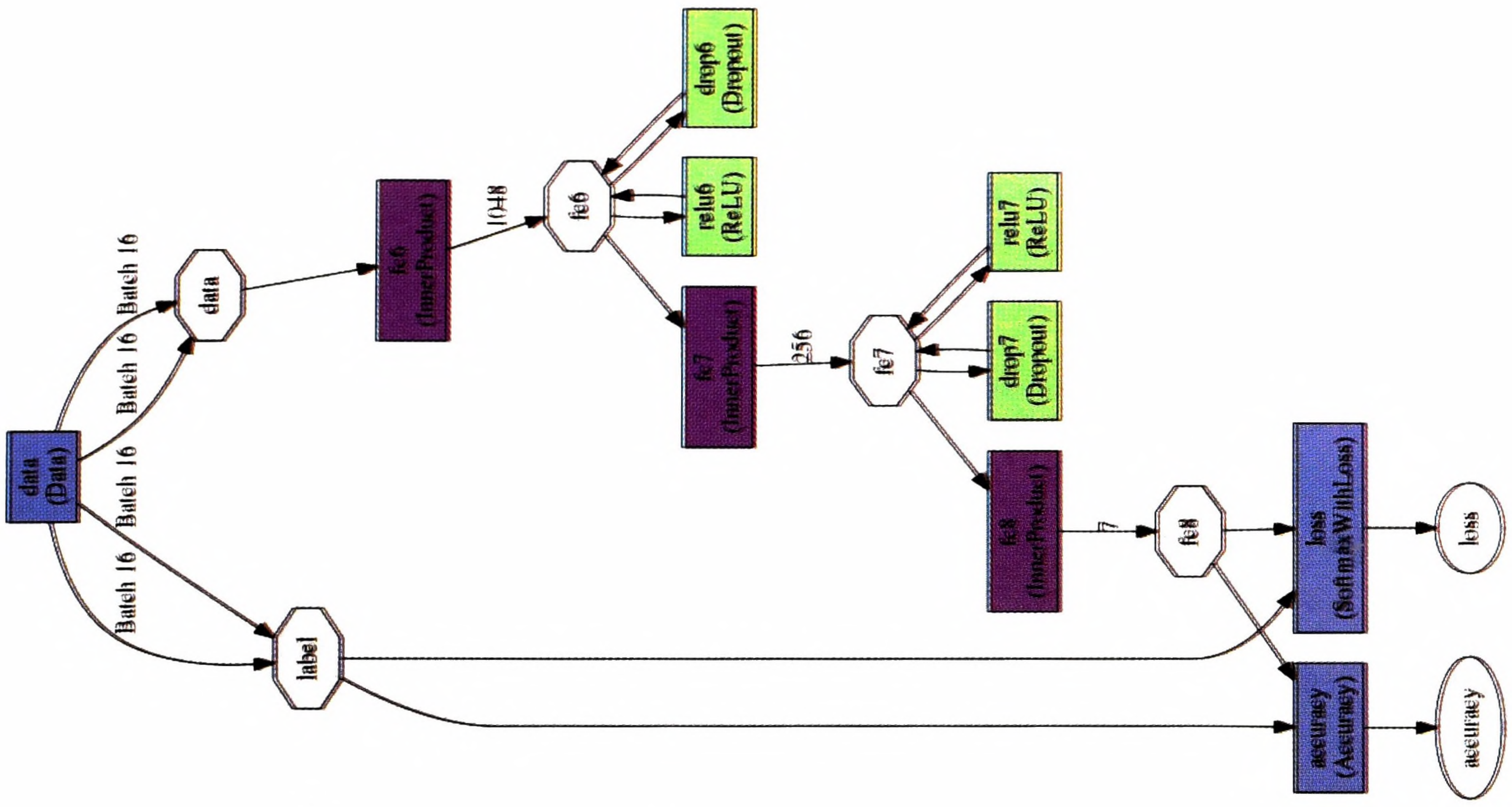

Fig. 2: Multi-layer perceptron architecture

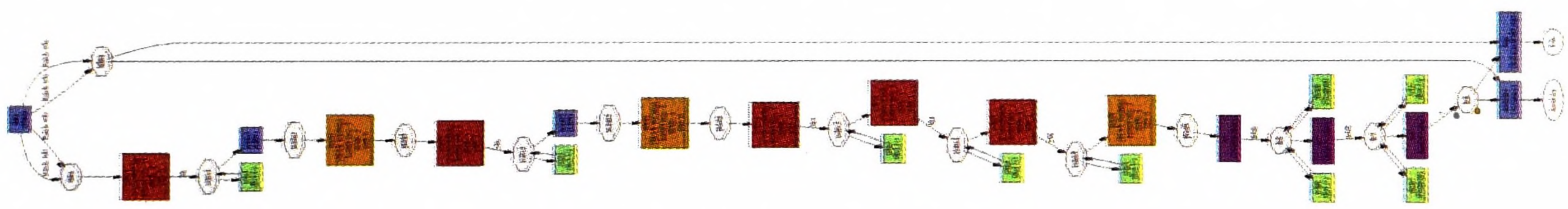

Fig. 3: Implemented CNN architecture

Convolutional Neural Network (CNN) is a regular Multi Layer Perceptron (MLP) yet with deep architecture and advanced algorithms[10-12] that is used in computer vision. The network can have many hidden layers. Here it has used five convolution layers with rectified linear function (96-256-384-384-256) and max pooling, three fully connected layers (4096-4096-7) and output layer with softmax function. The network was trained by choosing the SGD algorithm. The learning rate was set to 0.01. We trained the network for 500 epochs.

Results and Discussion

Table 1 shows the results obtained for each model for validation images The results show that the CNN is better in facial image classification according to the shape. CNN has several advantages compared to MLP. CNN processes images in two dimensional space which leads to better approximations using 2D kernels. The other network use vector space to input data. The processing of data of size 65536 increases the training time of the network.

Furthermore it could be seen that an overfitting problem in MLP when running for more than 50 epochs (100, 500). This happened due to the convergence to a local minimum for the training data. This leads the accuracy of 20% and loss of 10 for validation data.

The size of the dataset is crucial to determine the accuracy and robustness of system. Most practical larger CNN systems in the literature were trained with millions of data in multi GPUs with parallel processing. Large dataset can produce more accurate classification because of the number of large amount of training examples. It could be seen that the performance of CNN is higher than that of other model.

Conclusion

The goal of the study was to develop a prototype system to identify face shapes by using image processing and machine learning. Multi-layer perceptron and convolutional neural network

Model	Validation accuracy	Epochs	Training time
Convolutional NN	98.5%	500	1 hour, 3 minutes
MLP	58.3%	50	18 mins, 46 seconds

architectures were considered in advance. Each system was trained, validated and tested. To select the best model, the accuracy of validation dataset was compared. The results indicate that the convolutional neural network is more useful in identification shapes than multilayer perceptron.

Table 1: Results of the study

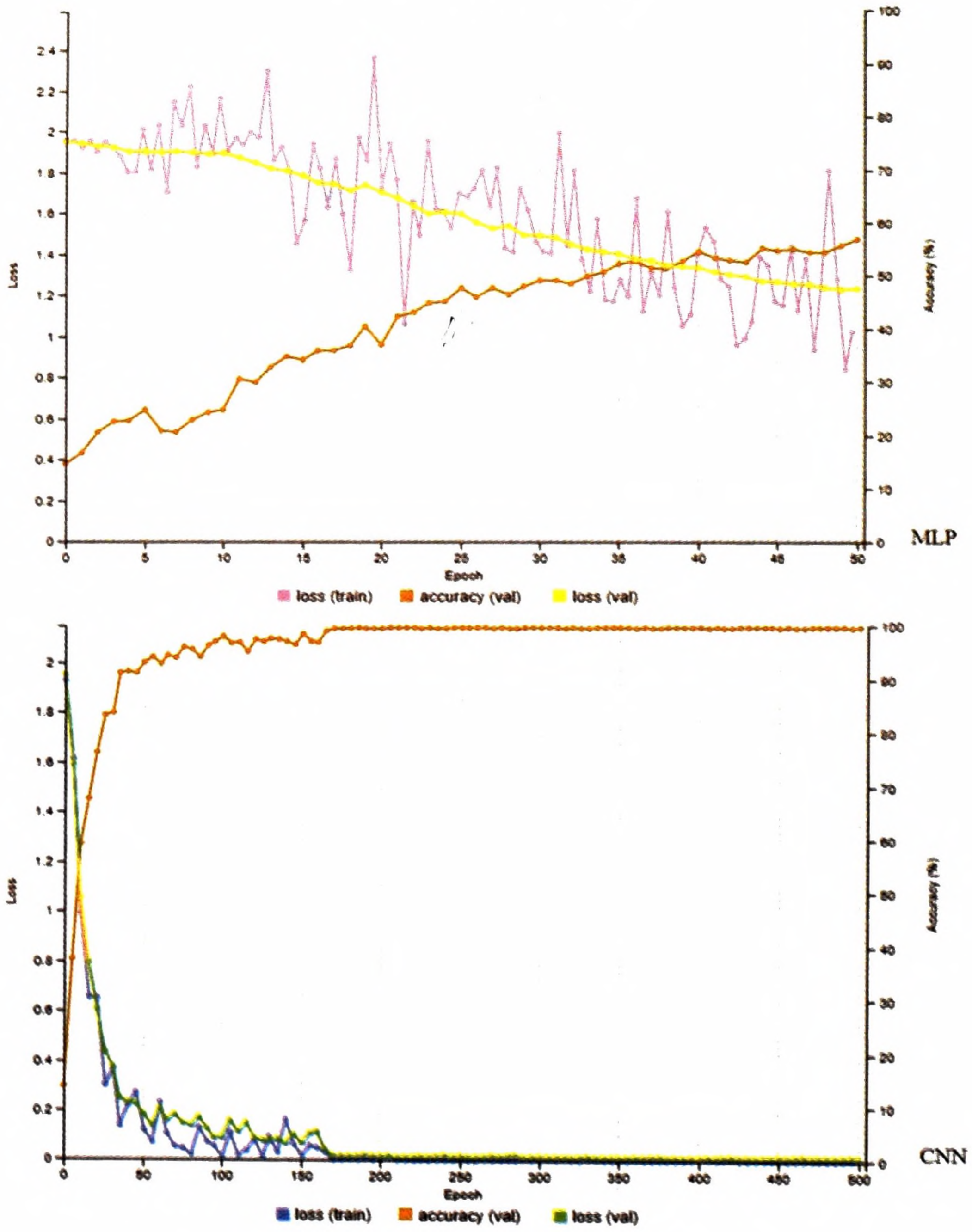

Fig. 3: The graph of validation accuracy vs. training and validaion loss, MLP (top), CNN (Bottom)

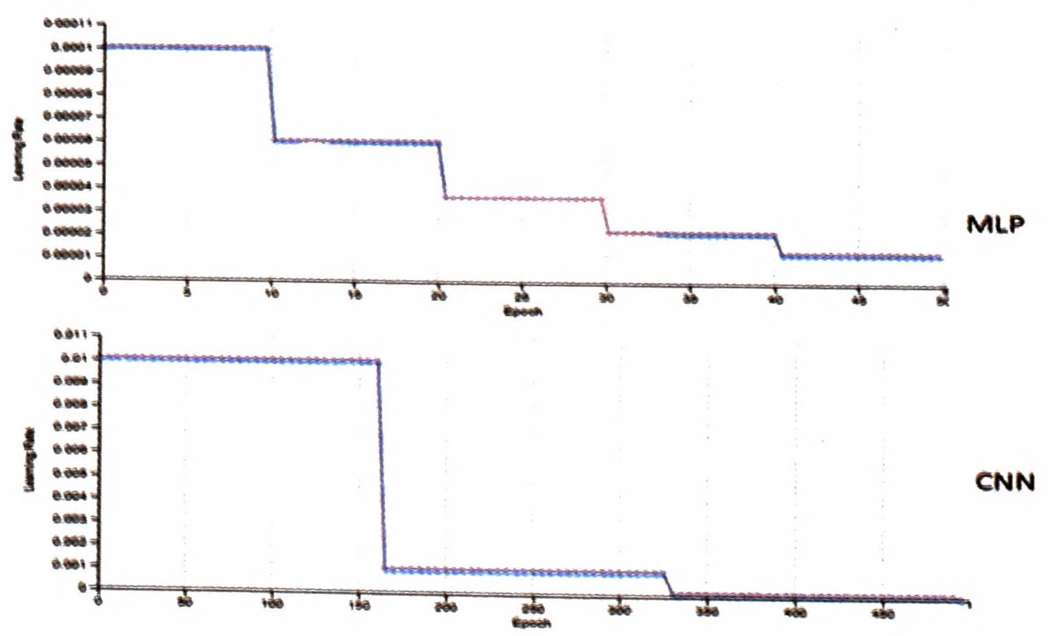

Fig. 4: Graph of Learning rate vs epoch for each network, MLP(Top), CNN(Bottom)

The implemented system can identify face shape with good accuracy rate. However there are some potential enhancements. This system only trained with 3150 faces of ages of 20-40 which conform more to a certain shape. Training with age different images may enhance the scope of the system. Training with larger images can improve the generalization of the system. Efficiency of the training of the system can be increased by using better computer resources such as multi GPUs. Applications can be implemented using this system to include eyeglasses, hair styles, haircuts etc. Improving the GUI of the system can improve its usability. It can be further expanded into web based or mobile friendly applications.

References

1. M. B. Gantt, How to wear your hair, Tate Publishing and Enterprises LLC, USA, 2010.
2. K. L. Lim and K. H. Galoogahi, Shape Classification Using Local and Global Features, *Fourth Pacific-Rim Symposium on Image and Video Technology*, 2010.
3. J. So, S. Han, H. C. Shin and Y. Han, An AAM-based Face Shape Classification Method Used for Facial Expression Recognition, *Intl. J. Res. in Eng. Technol. (IJRET)*, 2013, 2-4, 164-168.
4. S. W. Lin, S. Y. Chou and S. C. Chen, Irregular shapes classification by back-propagation, *Springer*, 2006.
5. J. D. Vries, Object Recognition:A Shape-Based Approach using Artificial Neural Networks, *Dissertation.*
6. S. Duffner, Face Image Analysis With Convolutional Neural Networks, *Dissertation,* 2007.
7. A. Krizhevsky, I. Sutskever, and G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, *Neural Information Processing Systems,* 2012.
8. NVIDIA DIGITS [Computer Software], Retrieved: https://developer.nvidia.com/digits.
9. A.Y. Jia, E.Shelhamer, J. Donahue, S.Karayev, J. Long, R.Girshick, S.Guadarrama, and T. Darrell, Caffe: Convolutional architecture for fast feature embedding, *Proceedings of the ACM Intnl. Conf. on Multimedia,* 2014, 675-678.
10. B. E. Boser, I. M. Guyon and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classiers, *Proc. Fifth Ann. Wkshp on Computational Learning*, 1992,144-152.
11. K. Jarrett, K. Kavukcuoglu, M. A. Ranzato and Y. LeCun, What is the Best Multi-Stage Architecture for Object Recognition?,*The Courant Institute of Mathematical Sciences, New York University.*
12. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *J. Machine Learning Res.*, 2014, 15, 1929-1958.