ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

Handwritten Digit Recognition: Convolutional Neural Network as a Classifier

Jisha Mol L, Femy John

M Tech Student, Dept. of ECE, VJCET, Kerala Technological University, Vazhakulam, Ernakulam, India

Assistant Professor, Dept. of ECE, VJCET, Kerala Technological University, Vazhakulam, Ernakulam, India

ABSTRACT: Convolution Neural Networks (CNNs) consist of multiple layers. It is a powerful technique for classification of visual inputs like handwritten digits and faces recognition. The classification task is performed using a Convolution Neural Network (CNN). The main purpose of using multilayer neural network is to minimize the mean square error, between the arrived output and the desired output. Here each subnet between the input and the hidden layer are initialized with random weights and also trained with different feature maps. All the connections here are adaptive in nature and are generally trained using back propagation algorithm. A dataset of 1000 samples were obtained from college and school students. The proposed system predicts the handwritten digits with an overall accuracy of 95-100%.

KEYWORDS: Handwritten digit recognition, convolutional neural networks, MATLAB

I. INTRODUCTION

Handwritten digit recognition plays an important role in pattern recognition and Optical Character Recognition (OCR). It has a wide range of practical applications in real life, such as zip code recognition in postal mail sorting[1], writer identification and verification, form processing, and handwritten digit recognition on bank check etc. Over the past decades, lots of machine learning methods have been employed for effective handwritten digit recognition , such as Linear and Non-Linear Classifier, Support Vector Machines (SVMs) , Neural Networks (NNs), Boosted Stumps, CNN-SVM Classifier, etc.

The task of recognizing the handwriting of an individual from another is difficult as each personal possess a unique handwriting style. This is one reason why handwriting is considered as one of the main challenging studies. The need for handwritten digit recognition came about the time when combinations of digits were included in records of an individual [2]. Although there are several image processing techniques designed, the fact that the handwritten digits do not follow any fixed image recognition pattern in each of its digits makes it a challenging task to design an optimal recognition system. This study concentrates on the offline recognition of digits. Using an MLP neural network.many methods have been proposed till date to recognize and predict the handwritten digits. Digit recognition was carried out using different algorithms like neural network algorithm and FDA algorithm. The FDA algorithm proved less efficient with an overall accuracy of 77.67%, whereas the back-propagation algorithm with PCA for its feature extraction gave an accuracy of 91.2%. in 2014, a novel approach using SVM binary classifiers and unbalanced decision trees was presented. The proposed system has been trained on samples of 700 images and tested on samples of 100 images written by 50 persons and could achieve a recognition rate of 95-100%

In most cases, a hand-crafted feature extractor is often designed, specifically adapted to the problem. This is a hard task, which must be redone for each new problem. Therefore, in this work a type of multilayer perception called Convolution Neural Network (CNN) is specifically designed to recognize two-dimensional shapes with a high degree of invariance to translation, scaling, skewing, and other forms of distortion [5]. This network includes in its structure some forms of constraints: feature extraction, feature mapping and subsampling. In these experiments, the network architecture is composed by an input layer, six hidden layers and an output layer. The topology of a typical CNN contains two types of hidden layers. Some of them perform convolution, i.e., each layer is composed by some feature maps which are used for local feature extraction. This is achieved by an operation equivalent to convolution, followed by an additive bias and squashing function.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

II. RELATED WORKS

In 1990, Y.LeCun, et.al[1], proposed a Back-propagation network for recognizing handwritten digits. The preprocessing stage here includes Acquisition of the data, binarization, location of the zipcode, preliminary segmentation and finally the normalization of the digits using linear transformation method. The recognition task is performed by the multi-layer network, as well as the connections within the network are adaptive and the network was trained with back-propagation network. With their approach the error rate of the training set was 3.4% and the MSE was 0.024. The results show that the method can be extended for larger applications.

In 1998, Daniel Cruces Alvarez, et.al[4], proposed a neural network method for recognizing printed and handwritten digits. The recognition is performed using a multilayer and clustered back propagation algorithm. For extracting the features kirsch masks are adopted and for classifying the numerals, they used five independent subnetworks. The main purpose of using multilayer neural network is to minimize the mean square error, between the arrived output and the desired output. Here each subnet between the input and the hidden layer are initialized with random weights and also trained with different feature maps.The refinement network here is made up of 45 neurons. Each one is trained to distinguish between the numerals. Thus with their method, the rejection rate is 9% and the error rate is reduced to 1%.

In 2009, CalinEnachescu and Cristian- DumitruMiron[6], proposed a neural computing method for recognizing handwritten digits. A framework was presented by them to classify handwritten digits, and the classification was performed using Convolutional Neural network. It was mainly designed to recognize patterns from pixel images directly with minimal preprocessing. CNN is a feed forward neural network. Features are extracted from the input image, in the first hidden layer and the patterns are classified in the final hidden layer. The number of hidden layer neurons can be varied to control the learning capacity as well as the generalization capacity of the classifier. With CNN it is possible to extract information from images with minimal preprocessing. Apart from this, it is also necessary to remove unwanted pixels from the background, which improves the performance not only during the learning process, but also when using CNN in normal. In this case they encoded the image, that are used in the learning process, in which the white pixels are set to 1 and the background pixels are set to 0. The parameter which is affecting the learning process is the learning rate. The higher the learning rate, learning performance is low, whereas with small learning rate, the learning performance is better. After the learning process is over, they tested the network with different images. With original NIST dataset, it provided 96.74% accuracy, and with the images that are without background it provided 96.56% accuracy.

II. DATA COLLECTION

In this study, a subset of 1000 samples were collected from the college and school students. Each sample was a gray-scale image of size (32×32) pixels. Figure 1 shows some sample images. The input dataset was obtained from the database with 100 samples of each digit from 0 to 9. As the common rule, a random 600 samples of the input dataset were used for training and the remaining 100 were used for validation of the overall accuracy of the system.

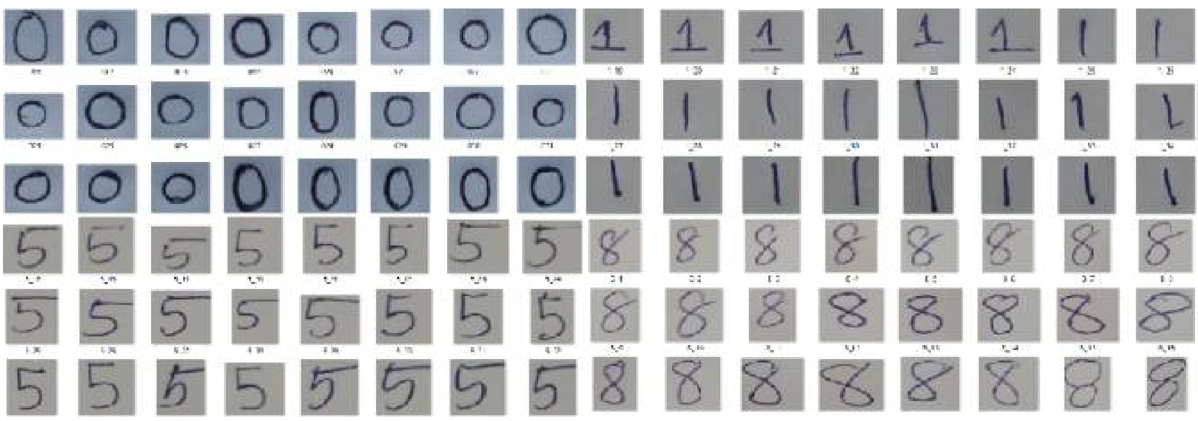

Fig.1 Collected datasets

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

III. PROPOSED ALGORITHM

Convolutional Neural Networks are designed to process two-dimensional (2-D) image. A CNN consists of three main types of layers: (i) convolution layers, (ii) sub-sampling layers, and (iii) an output layer. Network layers are arranged in a feed-forward structure: each convolution layer is followed by a sub-sampling layer and the last convolution layer is followed by the output layer.The convolution and sub-sampling layers are considered as 2-D layers, whereas the output layer is considered as a 1-D layer. In CNN, each 2-D layer has several planes. A plane consists of neurons that are arranged in a 2-D array. The output of a plane is called a feature map [4].

In a *convolutional layer*, each plane is connected to one or more feature maps of the preceding layer. A connection is associated with a convolution mask, which is a 2-D matrix of adjustable entries called *weights*. Each plane first computes the convolution between its 2-D inputs and its convolution masks. The convolution outputs are summed together and then added with an adjustable scalar, known as a *bias* term. Finally, an activation function is applied on the result to obtain the plane's output. The plane output is a 2-D matrix called a *feature map*; this name arises because each convolution output indicates the presence of a visual feature at a given pixel location.

A convolution layer produces one or more feature maps. Each feature map is then connected to exactly one plane in the next sub-sampling layer. *Sub-sampling* layer has the same number of planes as the preceding convolution layer. A subsampling plane divides its 2-D input into non-overlapping blocks of size 2 × 2 pixels. For each block, the sum of four pixels is calculated; this sum is multiplied by an adjustable weight before being added to a bias term. The result is passed through an activation function to produce an output for the 2×2 block. Clearly, each sub-sampling plane reduces its input size by half, along each dimension. A feature map[7] in a sub-sampling layer is connected to one or more planes in the next convolution layer.

In the *last convolution layer*, each plane is connected to exactly one preceding feature map. This layer uses convolution masks that have exactly the same size as its input feature maps**.**

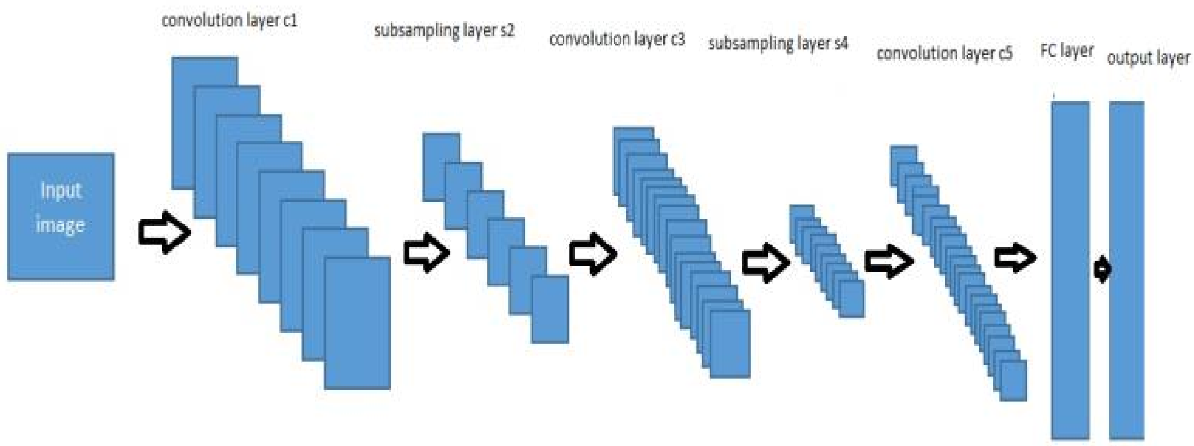

Fig.2 CNN architecture

IV. LAYERS OF CNN

By stacking multiple and different layers in a CNN, complex architectures are built for classification problems. Four types of layers are most common: convolutionlayers pooling/subsampling layers, non-linear layers, and fully connected layers. INPUT [32x32x3] will hold the raw pixel values of the image with width 32, height 32, and three colour channels R,G,B[8].

CONV layer will compute the output of neurons that are connected to local regions in the input, each computing a dot product between their weights and a small region they areconnected to the input volume[6]. This may

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

result in volume such as [32x32x12]. This system use 12 filters. RELU layer will apply an elementwise activation function, such as max(0,x). This leaves the size of the volume unchanged ([32x32x12]). Non-linear layers Neural networks in general and CNNs in particular rely on a non-linear “trigger” function to signal distinct identificationof likely features on each hidden layer. CNNs may use a variety of specific functions such as rectified linearunits (ReLUs) and continuous trigger (non-linear) functions to efficiently implement this non-linear triggering. A ReLU implements the function y = max(x,0), so the input and output sizes of this layer are the same. It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields

The pooling/subsampling layer reduces the resolution of the features. It makes the features robust against noiseand distortion. There are two ways to do pooling; max pooling and average pooling. In both cases, the input isdivided into non-overlapping two-dimensional spaces. For average pooling, the average of the fourvalues in the region are calculated. For max pooling, the maximum value of the four values is selected.

FC (i.e. fully-connected) layer will compute the class scores, resulting in volume of size [1x1x10], where each of the 10 numbers correspond to a class score. Fully connected layers are often used as the final layers of a CNN. These layers mathematically sum the weights of the previous layer of features, indicating the precise mix of “ingredients” to determine a specific target outputresult. In case of a fully connected layer, all the elements of all the featuresof the previous layer get used in thecalculation of each element of each output feature[11].

V. TRAINING OF NETWORK

The simplest among the associative memory model is the feed-forward type of neural network. A multi-layer feed-forward back propagation neural network is used for recognizing handwritten numerals. The training network consists of three layers; the input layer, hidden layer and the output layer. An input vector X is applied to the network through the input layer and the output Y is produced across the output layer. In the case of multilayer network, the **NET** value is calculated for each neuron layer by layer using,

NET=XW eq(1)

where X is the input and W is the weight value. For processing the NET signal to further layers, an activation function is applied to the NET value, and the OUT value is calculated by,

OUT=F(XW). eq(2)

Here for example the target value is set to 1. If the calculated OUT value is not greater than or equal to 1, the desired target value, the backpropagation algorithm is applied and the associated weights are adjusted using the Delta rule. The Delta value is calculated using the formula,

δ=**OUT(1-OUT)(Target-OUT)**eq(3)

where (Target-OUT) gives the error signal, It is then multiplied by the squashing function OUT(1-OUT). After the Delta value is calculated, it is multiplied with the OUT value of the desired layers neuron, which is further multiplied with the training rate coefficient μas,

Δ**W=** $\mu.\delta$**.OUT** eq(4)

The result thus obtained is added to the weight value of the corresponding neuron in the hidden layer to the output layer by,

W(n+1)=W(n) + Δ**W** eq(5)

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

where W(n) is the value of weight from neuron in the hidden layer to the neuron in the output layer, and W(n+1) is the weight value after adjustment. After the weight value is adjusted, it is multiplied with the input vector X in the hidden layer and the OUT value is calculated[10]

VI. SIMULATION RESULTS

The performance of the proposed handwritten digit recognition technique is evaluated experimentally using 100 random test samples. The experiments are implemented in MATLAB 2012 under a Windows7 environment on an Intel Core2 Duo 2.4 GHz processor with 4GB of RAM[14].

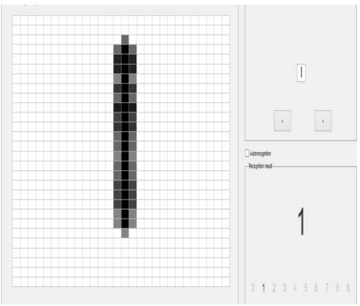

Fig.3.Predicted as 1

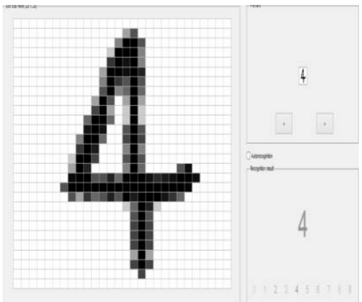

Fig. 4.Predicted as 4

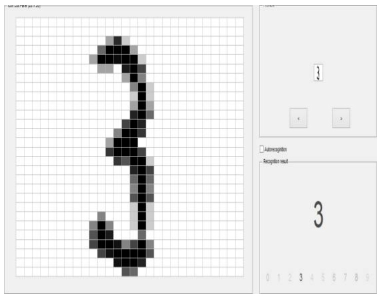

Fig5. Predicted as 3

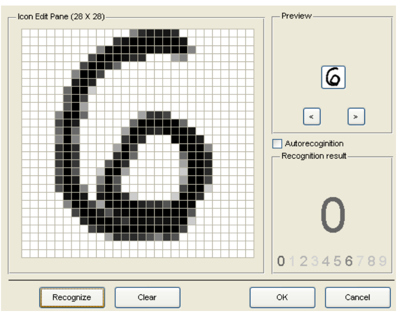

Fig.6.Predicted as 0

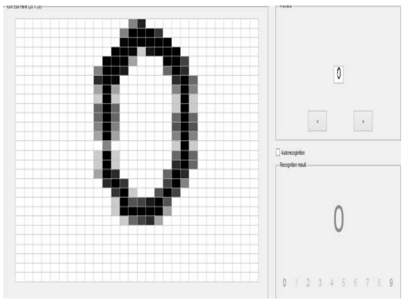

Fig7.Predicted as 0

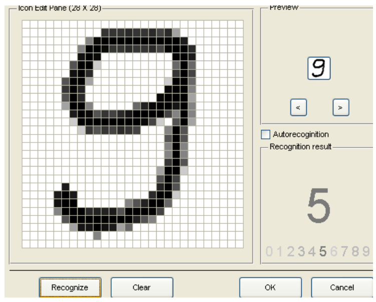

Fig 8. Predicted as 5

In this study, a subset of 1000 samples was collected from the college and school students. Each sample was a grayscale image of size (32×32) pixels. here classification is done through CNN algorithm. fig.3 is predicted as '1',correct classification. Fig.4 is predicted as '4', correctly predicted. But fig.6 predicted as '0', wrong classification. Digitized handwritten images are tested with the proposed method and the recognition is found to be 95-100% accurate

VII. CONCLUSION AND FUTURE WORK

In this paper, the authors implemented a multi-layer feed-forward backpropagation network that recognizes handwritten numerals. Given input image[21] of numerals undergoes three stages namely preprocessing, training and recognition. Preprocessing stage plays a vital role and influences the accuracy of recognition. Images of various sizes and shapes undergo preprocessing stage first. Then the multi-layered feed-forward back propagation network is trained with preprocessed images. The final recognition stage recognizes images of any size and shape as input to the method. The proposed method is implemented withMATLAB coding. Digitized handwritten images

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

are tested with the proposed method and the recognition is found to be 95-100% accurate. Similar strokes of two numerals are recognized as one for the other, was the drawback of this method. Further research is in progress to eradicate this negative recognition[18].

REFERENCES

1. Y.LeCuu, B.Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel," Handwritten Digit Recognition with a Back-Propagation Network," AT&T Bell Laboratories, Holmdel, N.J.07733 1990. JagadeeshBabu V, Prasanth L, Raghunath Sharma R, PrabhakaraRaoG.V., Bharath A., HMM-based Online Handwriting Recognition System for Telugu Symbols, ICDAR'2007, 23-26 September2007.
2. L.R.Rabinar,"A tutorial on Hidden Markov Model and selected applications in speech recognition", Proceeding of the IEEE,vol.77,no 2,feb 1989
3. K.Tyagi,V.Rastogi,"Implementation of character recognition using Hidden Markov Model",Department of CS ,IET,vol 3,issue 2,feb 2014.
4. Daniel Cruces Alvarez, Fernando Martin Rodriguez, Xulio Fernandez Hermida,"Printed and Handwritten Digits Recognition Using Neural Networks", E.T.S.I.T. Ciudad Universitaria S/N.36200 Vigo.SPAIN, 1998. A.Choudharya, R.Rishib, S.Ahlawatc," Off-Line Handwritten Character Recognition using Features Extracted from Binarization Technique", 2013 AASRI Conference on Intelligent Systems and Control.
5. Manojkumarp,S.chandran," Handwritten Malayalam Word RecognitionSystem using Neural Networks", International Journal of Engineering Research & Technology (IJERT) SSN: 2278-018 Vol. 4 Issue 04, April-2015.
6. CalinEnachescu, Cristian-DumitruMiron,"Handwritten Digits Recognition Using Neural Computing", Scientific Bulletin of the PetruMaior University of TirguMures, Vol.6(XXIII), 2009, ISSN 1841-9267. Rodolfo Luna-Pérez, Pilar Gómez-Gil," Unconstrained Handwritten Word Recognition Using a Combination of Neural Networks", WCECS 2010, October 20-22, 2010, San Francisco, USA.
7. DewiNasien, Siti S. Yuhaniz, HabibollahHaron," Recognition of Isolated Handwritten Latin Characters Using One Continuous Route of Freeman Chain Code Representation and Feedforward Neural Network Classifier", World Academy of Science, Engineering and Technology 67 2010.
8. Mahesh R. Barkul1, S. S. Lokhande, "A Review: Handwritten Character Recognition System ", International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 2, February 2016
9. Sonalkhare, Jaiveersingh,"handwritten devanagari character recognition system: a review" , international journal of computer applications (0975 – 8887) volume 121 – no.9, july 2015 .
10. S.M.Krishna Ganesh, P.Saranya,"An Evolutionary Neural Network Architecture Optimization Algorithm for Handwritten Digits Recognition", IOSR Journal of Engineering, Apr.2012, Vol.2(4) pp:726-730.
11. J.V.S. Srinivas, P.Premchand," Handwritten Digit Recognition Using Elman Neural Network", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 5, May 2012. Aiming, He and Xiangyu, Zhang and Shaoqing, Ren and Jian Sun "Spatial pyramid pooling in deep convolutional networks for visual recognition" European Conference on Computer Vision, 2014. 22.
12. Xiao-Xiao Niu and Ching Y. Suen, A novel hybrid CNN–SVM classifier for recognizing handwritten digits, ELSEVIER, *The Journal of the Pattern Recognition Society, Vol. 45*, 2012, 1318– 1325.

BIOGRAPHY

Jishamol L is a M Tech student in the Electronics and communication Department, ViswajyothiCollege of Engineering and Technology, Kerala Technical University. She completed her Bachelor of Technology from ASIET, MG university. Her research interests are Neural Networks, Verilog, Image processing etc.

Femy John is an Assistant Professor in the Electronics and communication Department, Viswajyothi College of Engineering and Technology, Kerala Technical University. She received her Master of Degree from Model Engineering College in 2010. Her research areas are VLSI, Microprocessor, Solid State Devices etc.