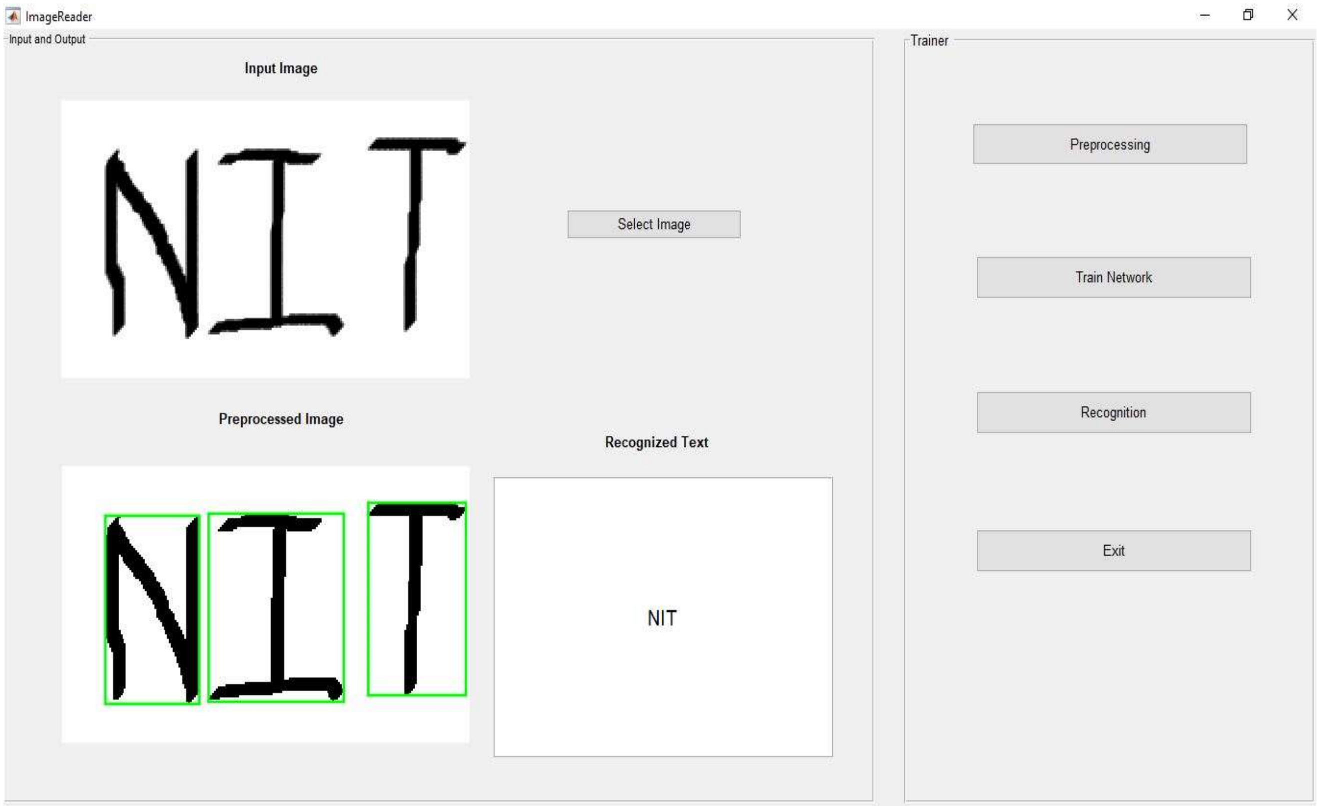
Handwriting Recognition Using Neural Network



Ashirbad Meher
1504017
Department of ECE
NIT Patna

Akshay Kumar Gupta
1504074
Department of ECE
NIT Patna

Aayush Malviya
1504076
Department of ECE
NIT Patna

ACKNOWLEDGEMENT

This project is a work of extensive research and adherence. This is our first minor project, which we have done during the 6th semester of our B. Tech course in Electronics and Communication. We have successfully completed this project in a span of 3 months. It would not have been possible if we did not have the support of our esteemed professors.

We are thankful to our mentor Dr. Gayadhar Pradhan, HOD, ECE Department for his continuous guidance and motivation.

We would also like to thank Dr. Syed Shahnawazuddin without whom this project would have hit a wall.

Special thanks to the good people at NIST who have worked hard for a large and vivid database such as EMNIST.

--------------------- --------------------------- ----------------------

Aayush Malviya Akshay Kumar Gupta Ashirbad Meher

CERTIFICATE

This is to certify that the minor project entitled "Handwriting Recognition using Neural Networks" is submitted by

Ashirbad Meher	1504017
Akshay Kumar Gupta	1504074
Aayush Malviya	1504076

has been completed successfully for the degree of Bachelor of Technology in Electronics and Communication Engineering.

……………………………

Prof. Gayadhar Pradhan

(Project Supervisor)

ABSTRACT

A handwritten character recognition system using Convolutional neural network and Feed Forward Neural Network are proposed in this paper. The character data set suitable for recognizing postal addresses contains 47 elements, which include 26 uppercase alphabets, 10 numerals and others were lowercase alphabets that look different from their uppercase counterparts. EMNIST dataset was used for training the neural network for classification and recognition of the characters. The trained neural recognition system is tested for various inputs and found to perform well. The proposed systems could aid conversion of any handwritten document into structural text form.

Keywords

Handwritten character recognition, Image processing, Convolutional neural networks

1. INTRODUCTION

Handwriting Recognition has been one of the most fascinating and challenging research areas in field of image processing and pattern recognition in the recent years. It contributes immensely to the advancement of an automation process and can improve the interface between man and machine in numerous applications. Several research works have been focusing toward evolving newer techniques and methods that would reduce the processing time while providing higher recognition accuracy[1] [2].

In general, handwriting recognition is classified into two types as off-line and on-line handwriting recognition methods. In the off-line recognition, the writing is usually captured optically by a scanner and the completed writing is available as an image. But, in the on-line system the two dimensional coordinates of successive points are represented as a function of time and the order of strokes made by the writer are also available. The online methods have been shown to be superior to their off-line counterparts in recognizing handwritten characters due to the temporal information available with the former [3]. However, in the off-line systems, the neural networks have been successfully used to yield comparably high recognition accuracy levels [4].

Several applications including mail sorting, bank processing, document reading and postal address recognition require off-line handwriting recognition systems. As a result, the off-line handwriting recognition continues to be an active area for research towards exploring the newer techniques that would improve recognition accuracy [5] [6].

The first important step in any handwritten recognition system is pre-processing followed by segmentation and feature extraction. Pre-processing includes the steps that are required to shape the input image into a form suitable for segmentation [7]. In the segmentation, the input image is segmented into individual characters and then, each character is resized into m x n pixels towards the extracting the features.

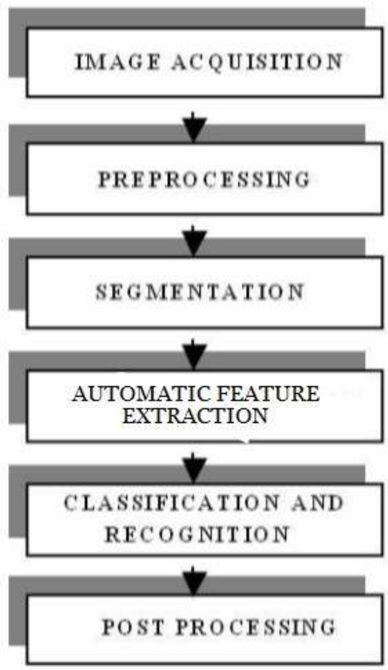
The Selection of appropriate feature extraction method is probably the single most important factor in achieving high recognition performance. Several methods of feature extraction for character recognition have been reported in the literature [8]. The widely used feature extraction methods are Template matching, Deformable templates, Unitary Image transforms, Graph description, Projection Histograms, Contour profiles, Zoning, Geometric moment invariants, Zernike Moments, Spline curve approximation, Fourier descriptors, Gradient feature and Gabor features.

2. LATEST ADVANCEMENTS RELATED TO THIS PROJECT

Recent deep learning based methods have achieved the state-of-the-art performance for handwritten Chinese character recognition (HCCR) by learning discriminative representations directly from raw data. Nevertheless, the long-and-well investigated domain-specific knowledge should still help to boost the performance of HCCR. By integrating the traditional normalization-cooperated direction-decomposed feature map (directMap) with the deep convolutional neural network (convNet), we are able to obtain new highest accuracies for both online and offline HCCR on the ICDAR-2013 competition database. With this new framework, it can eliminate the needs for data augmentation and model ensemble, which are widely used in other systems to achieve their best results. This makes our framework to be efficient and effective for both training and testing. Furthermore, although directMap+convNet can achieve the best results and surpass human-level performance, writer adaptation in this case is still effective.[9]

3. THE PROPOSED RECOGNITION SYSTEMS

A typical handwriting recognition system consists of pre-processing, segmentation, feature extraction, classification and recognition, and post processing stages. The schematic diagram of the proposed recognition system is shown in Figure.1[9]



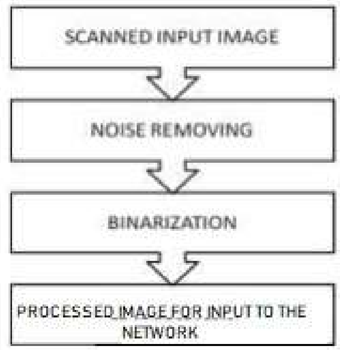
A. IMAGE ACQUISITION

In Image acquisition, the recognition system acquires a scanned image as an input image. The image should have a specific format such as JPEG, PNG etc. This image is acquired through a scanner, digital camera or any other suitable digital input device.

B. PRE-PROCESSING

The pre-processing is a series of operations performed on the scanned input image. It essentially enhances the image rendering it suitable for segmentation. The various tasks performed on the image in pre-processing stage are shown in Fig.2.
Binarization process converts a grayscale image into a binary image using global

thresholding technique. Detection of edges in the binarized image using sobel technique, dilation of the image and filling the holes present in it are the operations performed in the last two stages to produce the pre-processed image [1] suitable for segmentation.



C. SEGMENTATION

In the segmentation stage, an image of sequence of characters is decomposed into sub-images of individual character [1]. In the proposed system, the pre-processed input image is segmented into isolated characters by assigning a number to each character using a labelling process. This labelling provides information about number of characters in the image. Each individual character is uniformly resized into 28x28 pixels for extracting its features.

D. USING TRAINED NEURAL NETWORKS

In this stage, the features of the characters that are crucial for classifying them at recognition stage are extracted. This is an important stage as its effective functioning improves the recognition rate and reduces the misclassification. In our training dataset, all images are centred. If the images in the test set are off-centre, then the MLP network approach fails miserably. We want the network to be Translation-Invariant. The first part consists of Convolutional and max-pooling layers which act as the feature extractor. The second part consists of the fully connected layer which performs non-linear transformations of the extracted features and acts as the classifier. The convolutional layer can be thought of as the eyes of the CNN. The neurons in this layer look for specific features. If they find the features they are looking for, they produce a high activation. Convolution can be thought of as a weighted sum between two signals (in terms of signal processing jargon) or functions (in terms of mathematics). In image processing, to calculate convolution at a particular location (x, y), we extract k x k sized chunk from the image centred at location (x,y). We then multiply the values in this chunk element-by-element with the convolution filter (also sized k x k) and then add them all to obtain a single output. That's it! Note that k is termed as the kernel size.

E. CLASSIFICATION AND RECOGNITION

In classification and recognition, two architectures were used one using Feed Forward neural Network and another using Convolutional Neural Network.

The architecture for Convolutional Neural Network is as follows:

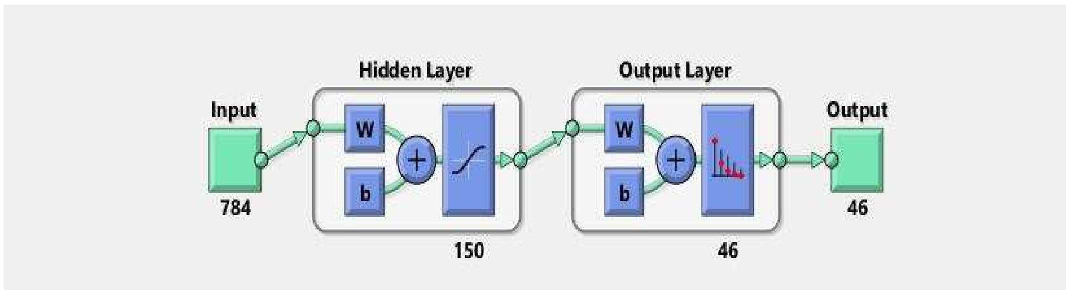
- Input nodes : 784
- Output nodes :37
- Training algorithm : Adam Optimizer
- Performance function : Categorical Cross Entropy

Layer (type)	Output Shape	Number of Parameters
conv2d_9 (Conv2D)	(None, 32, 28, 28)	320
activation_13 (Activation)	(None, 32, 28, 28)	0
conv2d_10 (Conv2D)	(None, 32, 26, 26)	9248
activation_14 (Activation)	(None, 32, 26, 26)	0
max_pooling2d_5(MaxPooling2)	(None, 32, 13, 13)	0
dropout_7 (Dropout)	(None, 32, 13, 13)	0
conv2d_11 (Conv2D)	(None, 16, 13, 13)	4624
activation_15 (Activation)	(None, 16, 13, 13)	0
conv2d_12 (Conv2D)	(None, 16, 11, 11)	2320
activation_16 (Activation)	(None, 16, 11, 11)	0
max_pooling2d_6 (MaxPooling2)	(None, 16, 5, 5)	0
dropout_8 (Dropout)	(None, 16, 5, 5)	0
flatten_3 (Flatten)	(None, 400)	0
dense_5 (Dense)	(None, 400)	160400
activation_17 (Activation)	(None, 400)	0
dropout_9 (Dropout)	(None, 400)	0
dense_6 (Dense)	(None, 10)	4010
activation_18 (Activation)	(None, 10)	0

CNN Architecture

While for Feed Forward Neural Network, it was as follows:

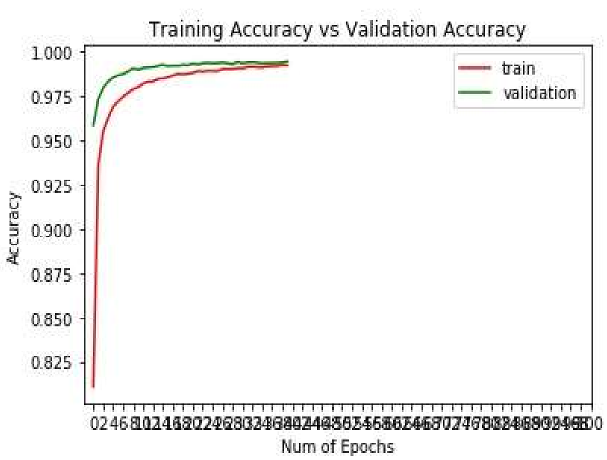
- Input Nodes: 784
- Output Nodes: 46
- Training Algorithm: Scaled Conjugated Gradient backpropagation
- Performance function: Cross Entropy

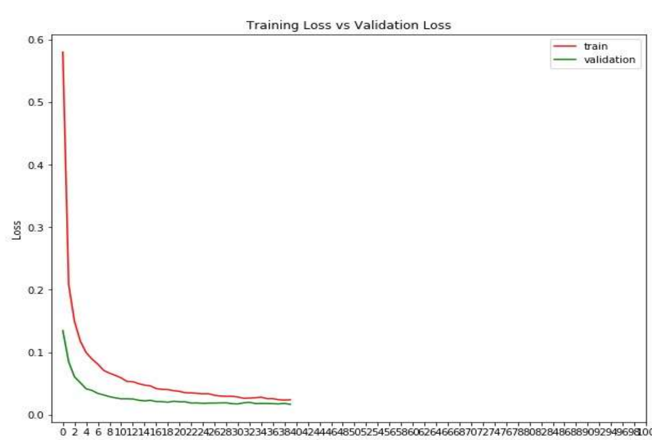


Feed Forward Neural Network Architecture

4. EXPERIMENTAL RESULTS

The recognition system has been implemented using Matlab and on Python. The scanned image is taken as dataset/ input and feed forward architecture is used. The structure of neural network includes an input layer with 784 inputs, one hidden layers each with 150 neurons and an output layer with 46 neurons. The gradient descent back propagation method with momentum and adaptive learning rate and RELU transfer functions is used for neural network training. Neural network has been trained using e-mnist dataset. After training the network, the recognition system was tested using several unknown dataset and the results obtained are presented in this section.





CONCLUSION

A simple recognition system for recognizing handwritten English alphabet characters is proposed. Experimental results show that the feed forward propagation neural network yields good recognition accuracy of 81% with 46 classes and the convolutional counterpart yields 99.45% accuracy with 37 classes. The proposed off-line hand written character recognition system with superior recognition rates will be eminently suitable for several applications including handwritten name recognition and conversion of any handwritten document into structural text form.

REFERENCES

[1] S. Mori, C.Y. Suen and K. Kamamoto, "Historical review of OCR research and development," Proc. of IEEE, vol. 80, pp. 1029-1058, July 1992.

[2] V.K. Govindan and A.P. Shivaprasad, "Character Recognition – A review," Pattern Recognition, vol. 23, no. 7, pp. 671- 683, 1990.

[3] N. Arica and F. Yarman-Vural, "An Overview of Character Recognition Focused on Off-line Handwriting", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol.31 no.2,pp. 216 - 233. 2001.

[4] R. Plamondon and S. N. Srihari, "On-line and off- line handwritten character recognition: A comprehensive survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 6384, 2000.

[5] U. Bhattacharya, and B. B. Chaudhuri, "Handwritten numeral databases of Indian scripts and multistage recognition of mixed numerals," IEEE Transaction on Pattern analysis and machine intelligence, vol.31, No.3, pp.444-457, 2009.

[6] Pal, U. and B.B. Chaudhuri, "Indian script character recognition: A survey," Pattern Recognition, vol. 37, no.9, pp. 1887-1899, 2004.

[7] Mantas, J, "An overview of character recognition methodologies," Pattern Recognition, vol.19, Issue 6, pp. 425-430, 1986.

[8] Anil.K.Jain and Torfinn Taxt, "Feature extraction methods for character recognition-A Survey," Pattern Recognition, vol. 29, no. 4, pp. 641-662, 1996

[9] Xu-Yao Zhang, Yoshua Bengio, Cheng-Lin Liu, arXiv:1606.05763, Cornell University Library

[10] Rafael C. Gonzalez, Richard E. woods and Steven L.Eddins, Digital Image Processing using MATLAB, Pearson Education, Dorling Kindersley, South Asia, 2004.

[11] R.G. Casey and E.Lecolinet, "A Survey of Methods and Strategies in Character Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No.7, July 1996, pp. 690-706