

Text Detection and Pattern Recognition Based on Deep Learning

Pallab Banerjee¹ Biresh Kumar² Govind Dwivedy³ Piyush Raj⁴

^{1,2}Assistant Professor ^{3,4}Student

^{1,2,3,4}Department of Computer Science and Engineering

^{1,2,3,4}Amity University Jharkhand, India

Abstract— In this paper we proposed a fast text detection method in which it is based on deep learning technique of neural network and also clustering method is used in classification or recognition of text. This project is aimed at summarizing and analyzing the major changes and significant progresses of scene text detection and recognition in the deep learning era. The extraction of text from a natural image is a challenging task. Text detection and character recognition in images is a research area which attempts to develop a computer system with the ability to read the text and pattern from images. At first a plain paper containing text is scanned and saved as image (.JPG). This image then pre-processed, segmented and detection of text done successfully using pixel scanning system. Text can be recognized with and without segmentation of characters. Segmentation can be line, word or character. While without segmentation character is recognized from whole text image. Here we use a new technique to detect text from the image called diagonal based feature extraction in last layer of convolutional neural network with the help of genetic algorithm. Once the extraction of its feature done we provide the training to learning machine. Along with this feature we use feed forward network as a classifier and convolution neural network for feature extractor. We basically use deep learning technique for training and testing. Character recognition convolutional neural network CRConvNet has more layers working of all layer shown in flowchart. One dataset which contain 360 training set data that are all in capital(A-Z) and small(a-z) alphabet, digit(0-9) and some special character are also used.

Keywords: Text Detection, Pattern Recognition

I. INTRODUCTION

Text recognition has innumerable applications in the field of image processing and pattern recognition like authentication of signatures in banks, processing of archived institutional records, recognizing ZIP code addresses on letters, forensic evidence, etc. This study is focused on increasing the recognition rate of the English characters. Feature of English character is extracted by a new method (diagonal method), and ANN is used to classify those features. This study is also focusing on evaluating the recognition rate of the suggested method to the current promising techniques, which is used in English numeral recognition, and English character recognition.

It has the capability to advance automation process and can reduce human intervention in numerous applications. Numerous techniques and methods are being developed to acquire higher accuracy with reduced processing time.

In this paper we will study Pattern recognition using supervised learning where we show how machine can observe the environment and able to distinguish pattern of interest.

The method of machine learning resembles with the data mining. Machine learning use for detection of pattern. If we enter in to machine learning we will find that machine learning is of two type First one is supervised learning and another one is Unsupervised learning.

Supervised learning is a type of learning where every input sample used for network is having desired output. Based on the error due to generated output the parameter of the network is adjusted.

In Unsupervised learning desired output for the sample is not provided the network learn on the basis of discovery and adjustment to the features on the basis of input pattern.

In text detection and pattern recognition from natural scene images the first important step is image acquisition or taking images from various resources, preprocessing the image followed by segmentation and feature extraction. Preprocessing include the step that remove noise or resize the image for proper segmentation. Various method are there for character segmentation.

In this paper diagonal based feature learning method is used with the help of genetic algorithm its provide accuracy in recognition. Various method are there for feature extraction like Template matching, graph description fourier description, projection histogram, zoning etc

A convolutional neural network is used for feature extraction and feed forward network is perform classification and recognition task.

The organization of the paper is as follows. Section I having the proposed recognition system. section II, the new feature extraction method i.e diagonal based feature extraction is explain. III contains classification and text recognition using deep neural network. Section VI represent Experimental result. Section V having conclusion of this paper.

II. PROPOSED SYSTEM

In this part of our proposed paper contain a deep learning based recognition system is explained. Text recognition from multimedia document having the following step i.e videos and image acquisition, pre-processing, segmentation, ROI detection and extraction, feature extraction using diagonal base learning technique, classification and recognition and result.

A. Image Acquisition

The image is acquired through a suitable input device like scanner and digital camera we scanned image as an input. Image should be in .jpeg formate. All image should be in machine edited form.



Sample image

B. Preprocessing

In pre-processing we are going to perform some basic operations of image processing steps are as follows:

- 1) noise reduction
- 2) binarization
- 3) edge detection
- 4) dilation
- 5) filling.

1) Noise Reduction:

With the intension to improve the quality of image noise reduction is one of the most important step. Noise may be impulse noise, Poisson noise and gausssian noise can be removed by filtering this method contain linear filter, adeptive filter etc and morphological operations.

2) Binarization:

he Binarization Method converts the grey scale image (0 up to 256 gray levels) in to black and white image (0 or 1). The result of OCR highly depends upon the binarization. The high quality binarized image can give more accuracy in character recognition as compared original image because noise is present in the original. A technique which is taken in to account for binarization is called thresholding. In thresholding, pixels that are alike in grayscale are grouped together. A thresholded image is defined as

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) \geq T \\ 0, & \text{Otherwise} \end{cases}$$

Where T shows global threshold value, $f(x, y)$ is the gray level at the point (x, y) . The pixels value is 1 correspond to objects, and pixels value is 0 correspond to the background.



Fig. 1. Gray Scale (Multi Tone) Image



Fig. 2. Binarized Image (Two Tone Image)

3) Edge Detection:

Edge detection is an image processing technique for finding the boundaries of objects within images .It also locating sharp discontinuities in an image. Some methods that perform edge detection are Robinson method canny method, sobel method, prewitt and roberts method etc.



b) Edged Image

4) Dilation and Filling:

Dilation is one of the two basic operators in the area of mathematical morphology, the other being erosion. It is typically applied to binary images, but there are versions that work on grayscale images. The basic effect of the operator on a binary image is to gradually enlarge the boundaries of regions of foreground pixels (i.e. white pixels, typically). Thus areas of foreground pixels grow in size while holes within those regions become smaller. Simply dilation is the process of stretching or shrinking the original image. It can also transforms an image in different size. The description of dilation includes the scale factor or ratio and the center of the dilation.

$$A \square B = \{ Z | (B)_z \wedge A \neq \emptyset \}$$

Where A set of foreground pixel

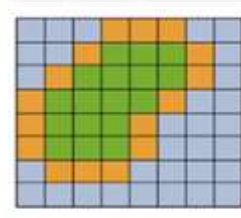
B is structuring element

Z foreground value

Filling is a process of filling the holes present in the image.



Original



after dilation

Filling is a process of filling the holes present in the image.

Preprocessing is the way of enhancing the image quality. Preprocessing is done in digital image. All type of noise can be removed from images. Various OCR method rely on preprocessing specially on binarization step. Introduce a automatic binarization based on a ConvNet for complex background and low resolution. Next we show a flow chart that will show our way of our proposed deep learning based recognition system.

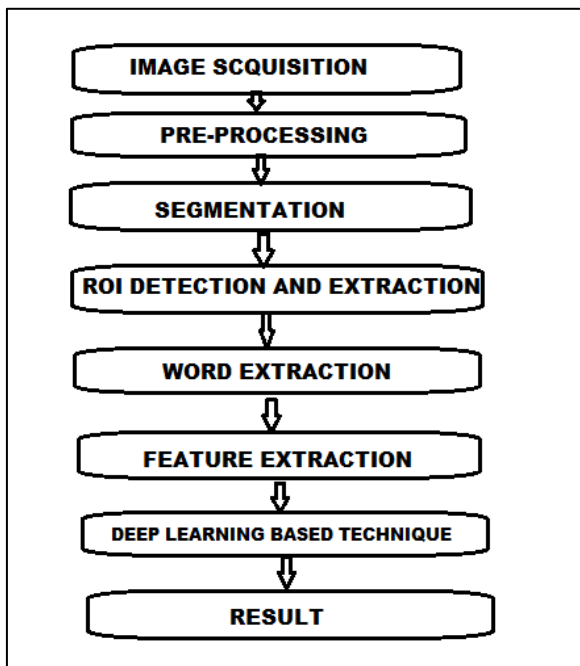


Fig. 2: Flow chart of deep learning base system

III. SEGMENTATION

There are two approach of segmentation through which we can detect the text and pattern

- 1) segmentation based
- 2) segmentation free.

In the segmentation based approach text is segmented into individual character before recognition.

In the segmentation free approaches the character are recognized as whole text without any segmentation.

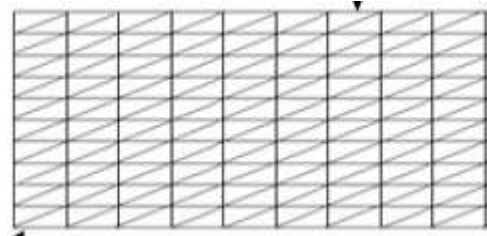
Segmentation means an image having text and then each character in this word is individual segmented with different segmentation method. we have lots of characters segmentation method. In this paper we are use global threshold and watershed segmentation method and Proposed system are use preprocessed image that is segmented into isolated character. Global thresholding means a single value of pixel intensity threshold is used for all pixels in the images for converting it to binary image. In watershed segmentation an image is regarded as a topographic landscape with ridges and valleys. The elevation values of the landscape are typically defined by the gray values of the respective pixels or their magnitude. Each individual image is uniformly resized into 81×81 pixels for classification and recognition.

IV. PROPOSED FEATURE EXTRACTION TECHNIQUE

There are various techniques which are showing different feature extraction technique. This paper is based on supervised learning text detection and character recognition which is concept of pattern recognition. In pattern recognition the most important thing is that the concept being apply for feature learning. In the proposed system we are use diagonal base feature learning technique. In that particular method we take each segmented character and resize it into 81×81 pixels and every 81×81 pixel is sub divided into 9×9 pixels. The feature are extracted from each zone of pixels by moving along the diagonals of its respective 9×9 pixels. Each zones

has 19 diagonals lines and the foreground pixel present long and each diagonal line is summed so that we get a single sub feature, thus each zone provide 19 sub feature. All zone repeat that procedure as same. There could be some zones whose diagonal are empty so that the value of zone is zero. Finally 1539 feature are extracted from each character. Feature is selected with using genetic Algorithm.

Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9
Z10	Z11	Z12	Z13	Z14	Z15	Z16	Z17	Z18
Z19	Z20	Z21	Z22	Z23	Z24	Z25	Z26	Z27
Z28	Z29	Z30	Z31	Z32	Z33	Z34	Z35	Z36
Z37	Z38	Z39	Z40	Z41	Z42	Z43	Z44	Z45
Z46	Z47	Z48	Z49	Z50	Z51	Z52	Z53	Z54
Z55	Z56	Z57	Z58	Z59	Z60	Z61	Z62	Z63
Z64	Z65	Z66	Z67	Z68	Z69	Z70	Z71	Z72
Z73	Z74	Z75	Z76	Z77	Z78	Z79	Z80	Z81



Procedure for extracting feature from character (a) Normalize character image (b) Image divided into 81 zones (c) Diagonal based feature extraction in the zone each zone having 9×9 pixel

V. CLASSIFICATION AND RECOGNITION

This step will provide us accurate result. A good classifier help us to detect best performance and result. otherwise a loose classifier can harm our performance. So it is good and important to give our project a good classifier with best feature extraction technique. fortunately we going to have both in our project. Various statical method is present like Artificial neural networks (ANNs), Multiple classifier combination and kernel method, Character classifier are Nearest neighbor classifier, Baye's classifier, Radial basis function and Support Vector Machine. Now here is the classifier part of our paper. In various method for classification and recognition are there. In this paper supervised learning technique is use which is convolution neural network for feature extraction and feed forward network for classification or recognition of character from multimedia document. There we are use pretrained CRConvNet and in fc7 layer we extracted feature with using diagonal based feature learning technique. After extracted the feature network is trained with the Back Propagation feed forward algorithm and With the help of this algorithm we find accurate result. CRConvNet has more layers: input layer, preprocessing layer, segmentation layer, feature extraction layer and output layer. This network have many number of layer therefore it is called supervised learning system. Basically CNN is used for feature extraction. After feature is extracted then it is send to the Feed Forward network for training and testing.

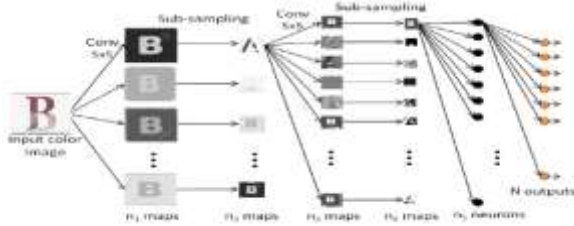


Fig 4. Architecture of Character Recognition using Convolutional Neural Network

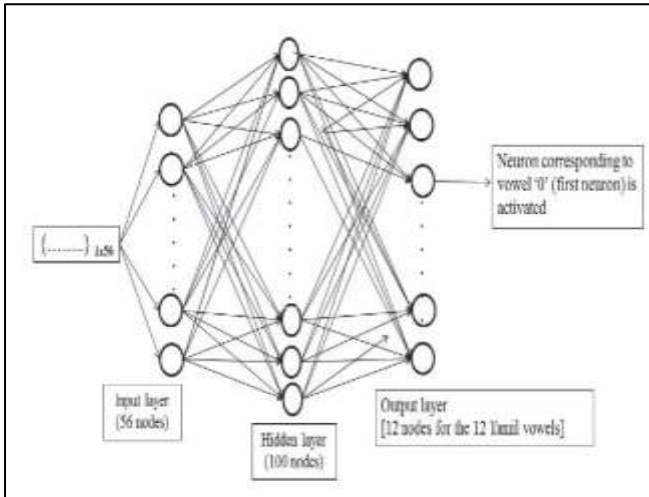


Fig.4 shows the level of convolutional followed by sampling. In the First layer preprocessing is performed after then stretch the image then segmentation take place. Last but one layer apply diagonal based feature learning and feature is selected using genetic algorithm and recognition of character is take place by using feed forward network

VI. RESULT DISCUSSION

We have implemented the proposed system using MATLAB. Testing dataset is taken as input to classifier in which we provide training using feed forward network and according to it we find result. Various graphs are shown below at running time condition result. We all know that all the diagonal value provide accuracy. First diagonal value is for character, second diagonal value is for digit and the last one is for special character and non-diagonal value shows misclassify data. Performance of our system

A. Confusion matrix

when we receive data after pre processing the first step we go for is to feed it to an outstanding model and get output as probabilities. Confusion Matrix is a performance measurement for machine learning classification. It allows the visualization of the performance of an system

O u t p u t C l a s s	1	31 60.2%	0 0.0%	1 3.2%	94.3% 5.3%
	2	2 4.3%	4 13.5%	0 0.0%	83.0% 23.2%
	3	0 0.0%	1 3.2%	3 10%	72.2% 23%
		96.4% 5.6%	84.6% 25.0%	75.7% 27%	91.3% 11.0%
		1	2	3	Target Class

Confusion matrix

B. Accuracy

In math, science, and engineering, accuracy refers to how close a measurement is to the true value. Accuracy is one of the most frequently used in some areas like Precision and Recall is also very frequent. In below we give the matlab graph for Accuracy, Precision, Recall and F- measure Calculation.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

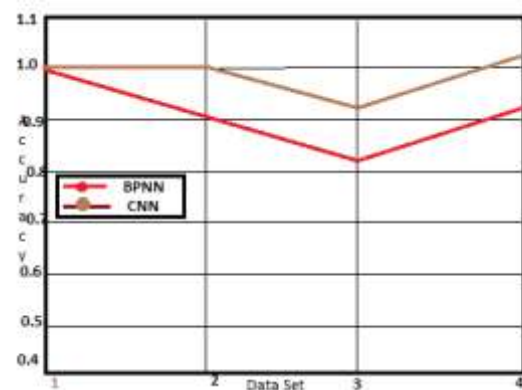


Fig. 5: Graph of Accuracy

Fig.5 shows the accuracy graph of our system. Here we use CNN for feature extraction and feed forward network for character recognition. Below the CNN value show the accuracy that lie between 0.8 to 1. The red graph show the accuracy when only apply the feed forward network. In our System we apply CNN with Feed forward network that provide better accuracy.

C. Precision:

Precision is how consistent results are when measurements are repeated. Precise values differ from each other because of random error, which is a form of observational error. In the field of text retrieval, precision is the fraction of retrieved character which is relevant to the query. Precision is also called positive predictive value. Precision can be find as :

$$\text{Precision} = TP / (True\ Positive + False\ Positive)$$

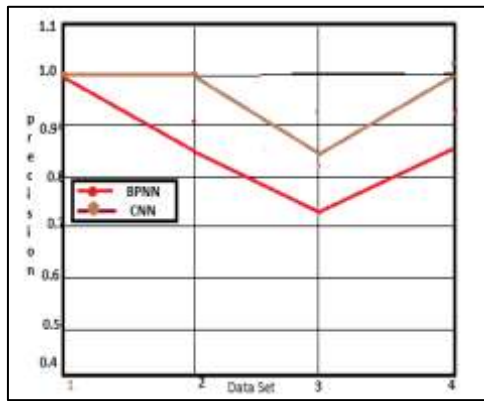


Fig. 6: Graph of Precision

Fig. 6 shows the precision value of graph. Here we use CNN with feed forward network for deep learning based character recognition which value is shown in brown graph that lies between 0.8 to 1 which is better than the graph value which is only using feed forward network for character recognition.

D. Recall:

Recall can be defined as the ratio of the total number of correctly classified positive examples divide to the total number of positive examples. It is also called positive sensitivity value or true positive rate, which can be calculated by the following equation :

$$\text{Recall} = \frac{TP}{TP + FN}$$

TP is the number of True positive samples FN is the number of False negative samples

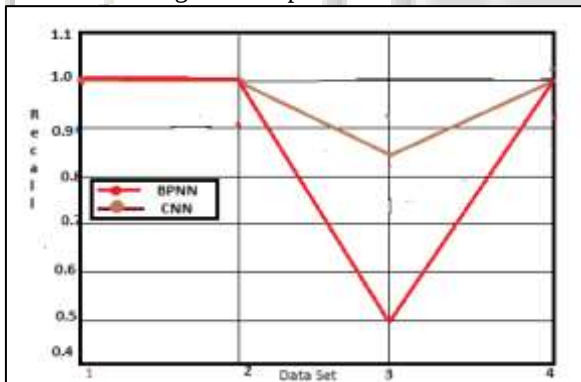


Fig. 7: Graph of Recall

Fig.7 shows the recall value graph. Where the brown graph having value between 0.85 to 1 shows the graph of deep learning based character recognition using CNN with feed forward network

VII. F-MEASURE:

Since we have two measures (Precision and Recall) it helps to have a measurement that represents both of them. We calculate an F-measure which uses Harmonic Mean in place of Arithmetic Mean as it punishes the extreme values more. Traditional F-measure or balanced F-score are shown as:

$$F - \text{measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

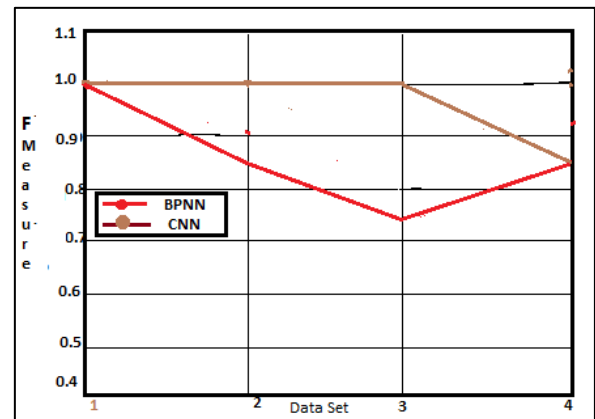


Fig. 8: Graph of F-Measure

Fig.8 shows the comparison graph between the recognition system using CNN with feed forward network and feed forward network only. CNN with feed forward network value lies between 0.85 to 1 which is better than the value having feed forward network.

VIII. CONCLUSION

This paper introduce a novel method of detection and recognition character from natural scene images using deep learning based technique. Choosing proper feature extraction technique and classifier which provide accuracy in text detection and character recognition. For feature extraction we are use convolutional neural network and in the last layer of this network we a new technology i.e diagonal based feature learning technique. With the proposed diagonal-based feature extraction classification

Accuracy is optimum. Diagonal base feature extraction in supervised learning the technique which is use is convolution neural network in last but one layer provide accurate result comparing to the other segmentation based technique. After the extracting feature for training it is send to the feed forward network and then testing take place. Here feed forward network is working as a classifier of neural network. Apart from efforts towards a general solution to all sorts of images, these algorithms can be trained and adapted to more specific scenarios, Techniques of unsupervised learning is the future enhancement of this paper.

REFERENCES

- [1] Shubhangi D. C. and Prof. P.S. Hiremath, "Handwritten English Character And Digit Recognition Using Multiclass SVM Classifier And Using Structural Micro Features", International Journal of Recent Trends in Engineering, Vol 2, No. 2, November 2009 193
- [2] Richard G. Casey and Eric 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
- [3] Anil. K. Jain, and Torfin Taxt, "Feature extraction method for character recognition- A Survey," Pattern Recognition, vol. 29, no. 4, pp. 641-662, 1996.

- [4] Saidane,z.garcia,c,"automatic scene text recognition using a convolution neural network".In: International Conference on Document Analysis and Recognition,vol. 2.pp. 874-879 (2007).
- [5] Vibhor Goel, Anand Mishra, Karteek Alahari, C. V. Jawahar. Whole is Greater than Sum of Parts: Recognizing Scene Text Words. International Conference on Document Analysis and Recognition, Aug 2013,Washington DC, United states.
- [6] Adam Coates, Blake Carpenter, Carl Case, Sanjeev Satheesh, Bipin Suresh, Tao Wang, David J. Wu, Andrew Y. Ng. Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning.International conference on document analysis and recognition 2011.
- [7] U. Pal, T. Wakabayashi, N. Sharma and F. Kimura, " Handwritten Numeral Recognition of Six Popular Indian Scripts", Ninth international conference on document analysis and recognition ICDAR 07,Vol.2'pp.749-753,2007.
- [8] Pradeep, E.Srinivasan and S.Himavathi, " Diagonal Based Feature Extraction For Handwritten Alphabets Recognition System Using Neural Network", (IJCSIT), Vol 3, No 1, Feb 2011 .
- [9] Dinesh Acharya U., N. V. Subbareddy & Krishnamoorthi, " Combined Classifiers In Recognition Of Handwritten Kannada Numerals: A Hybrid Approach", International Journal of Information Technology and Knowledge Management July-December 2009, Volume 2, No. 2, pp. 305-311
- [10] S.V. Rajashekararadhya and Dr P. Vanaja Ranjan, "Efficient Zone Based Feature Extraction Algorithm for Handwritten Numeral Recognition of Four Popular South Indian Scripts", Journal of Theoretical and Applied Information Technology, Vol4, No12 © 2005 - 2008 JATIT.
- [11] A deep learning based character recognition system from multimedia document by Usha Yadav