✚IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Handwritten Devnagari Special Characters and Words Recognition Using Neural Network

Vaibhav Dedhe[*1], Sandeep Patil[2]
[*1,2] Department of ETC,SSGI,Bhilai, India
Vaibhavdedhe@gmail.com

Abstract

The aim of this paper is to develop software which can recognize off line Devanagari special words which is made up of half consonant and consonant and special characters which is half form of consonant from scanned image of written documents using neural network. This paper will help to support easily digitization of the Devanagri script. Using this methodology for reorganization of the special characters of Devanagri script it is also easy to digitize the books written in Devanagri because ability of this methodology not only recognizes the normal characters but also the special characters. Some conventional methods like feature extraction and edge detection will be used for pre-processing the characters. These characters will be analysed by comparing its features. The process of data training of samples collected by different people will be followed after previous process. The proposed method will provide accuracy up-to 90% for special characters of deavanagri script with less training time.

Keywords: Feature extraction; edge detection; training time.

Introduction

Handwriting recognition is a technique which refers to the detection of written characters. This can also be viewed as we need to detect most right and appropriate character to which given figure matches. Handwritten characters are having many verities and it changes person to person. Offline character recognition refers to the recognition technique where the final figure is given to us [Bertolami, Zimmermann and Bunke, 2006]. [1]

Handwritten character recognition is a very vital field of learning in image processing. It uses images scanned from handwritten document to recognize the matching character and let the computer deal with the input data and information directly.

It has a large range of useful applications in the postal service, financial assistance, taxes, banking, etc. However, since there are features of random written factors and character-mode's instability, recognition of handwritten characters has been a Challenging matte.

The handwriting styles of different persons vary infinitely which makes the expansion of expert systems to recognize handwritten characters very difficult. The most major difficulty in handwriting recognition is the vast variation in personal writing styles. There are also differences in one person's writing style depending on the state, frame of mind of the writer and writing situation.

A recognition system should be insensitive to minor variations and still be able to distinguish unlike but sometimes very similar-looking characters.

India is a country of large population and different languages spoken all over the India and its national language is Hindi, the third most spoken language of the world, is represented in Devanagari script.

The aim of this paper is to develop software which can recognize special characters and word of Devanagari script from scanned image of printed documents.

In general, handwriting recognition is classified into two types as off-line and on-line handwriting recognition methods. In the off-line recognition, the characters is usually captured optically by a scanner and the complete characters are existing as an image. But, in the on-line system, the two dimensional coordinates of consecutive points are represented as a function of time and the order of strokes made by the writer are also available.[2]

Neural Networks tools in recent times being used in a variety of pattern recognition like image, character etc.. Neural network is playing an key role in handwritten character recognition. Many reports of character recognition in different languages have been published but still high recognition precision and minimum training time of handwritten characters

recognition using neural network is a area of interest for researcher.

Methodology

Normally , HCR can be divided into three parts namely pre-processing, feature extraction, and classification.

A. *Devanagri Special Characters*

Its basic set of symbols consists of 34 consonants and 18 vowels, and though Devanagari has a native set of symbols for numerals as shown in fig-1
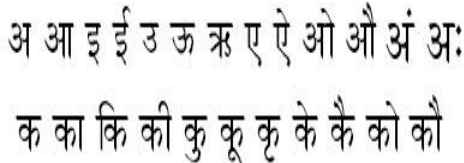

Fig-1: Non compound Devanagari Characters

Below fig shows the special characters of Devanagri Script consist of half form of consonants
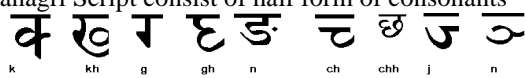
Fig-2: Half Form of Consonant Devanagari Characters

Combination of half consonant and consonant makes special words as shown in fig-
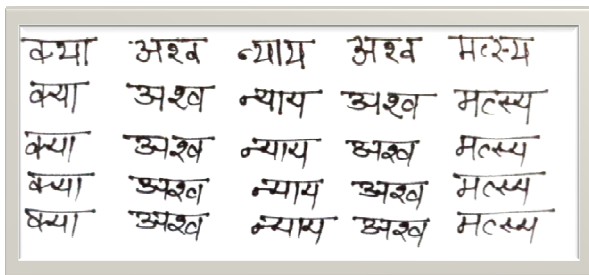

Fig-3: Special words made up of Devnagri half consonant and consonant

B. *Pre-Processing*

Pre-Processing stage is to produce a clean character image. As shown in Fig-4.
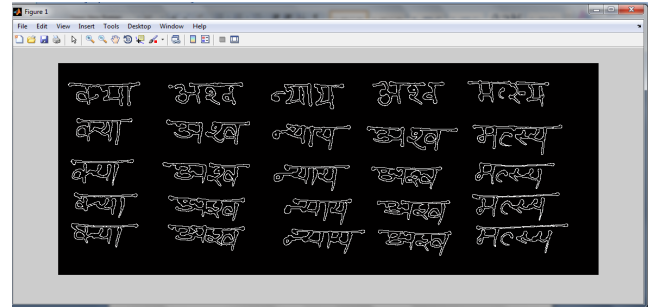

Fig-4: Image of samples after Pre-processing

C. Cropping of Image

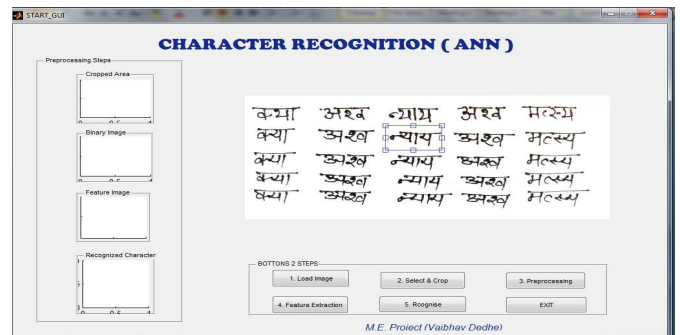In this step we can select any word from the loaded image for recognition process.

Fig-5: Selection of the word from loaded image

D. *Feature Extraction-*

Pre-processing step is followed by feature Extraction or edge detection process Here main features are extracted of the selected word by using an algorithm.
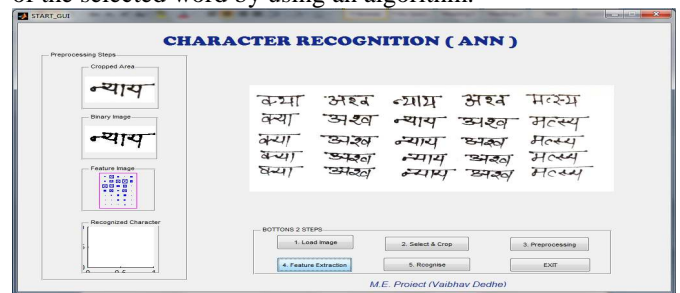

Fig-6: Feature Extraction

E. *Recognized character as a output*

After Feature Extraction we can get the recognized image which was selected –
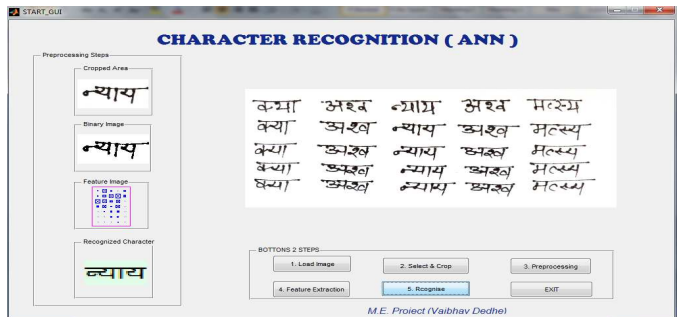
Fig-7: Recognized Word

Block Diagram of System

The Block diagram of whole system is shown in the following Fig- 8

Fig-8 Block Diagram Of The system

The procedure of handwritten character recognition is as follows:

- Get the sample by scanning as input.
- Edge detection and preprocessing operations are performed.
- Boundary Detection Feature Extraction Technique applied.
- Neural network Classification.
- Recognized Character as output.

Future Extraction

In this paper, to extract the information of the boundary of a handwritten character, the eight-neighbor adjacent method has been adopted. This scans the binary image until it finds the boundary. The searching follows according to the clockwise direction. For any foreground pixel B, the set of all foreground pixels connected to it is called connected component containing B. The pixel B and its 8-neighbors are shown in Figure 9. Once a white pixel is detected, it checks another new white pixel and so on. The tracing follows the boundary automatically. When the first pixel is found, the program will be assigned the coordinates of that position to indicate that this is an origin of the boundary. The new found pixel will be assigned as a new reference point and starts the eight-neighbor searching. In this way, the coordinates of the initial point are varied according to the position. As the tracer moves along the boundary of the image, the corresponding coordinates will be stored in an array for the computation of Fourier Descriptors. During the boundary tracing process, the program will always check the condition whether the first coordinates of the

boundary are equal to the last coordinates. Once it is obtained; means the whole boundary has been traced and boundary tracing process completes .
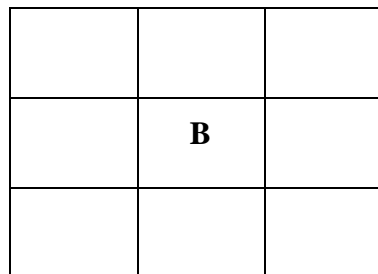

Fig- 9 Pixel A and its 8-neigbour

Neural Network

Recognition –

One well-organized way of solving complex problems is following the lemma "divide and conquer". A complex system may be decomposed into simpler elements, in order to be able to understand it.Recognition of handwritten characters is a very complex problem. A feed forward back propagation neural network is used in this work for classifying and recognizing the handwritten characters. The characters could be written in different size, orientation, thickness, format and dimension. This will give infinite variations. The neural classifier consists of two hidden layers besides an input layer and an output layer. For training, back-propagation algorithm has been implemented.
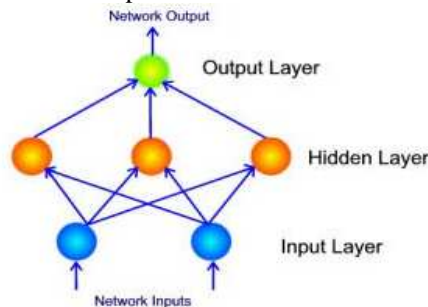

Fig-10: Network Layer

Multilayer Perceptron Layer –

Multilayer networks solve the classification problem for non linear sets by employing hidden layers, whose neurons are not directly connected to the output. The perceptron made up of one or more layers of artificial neurons, the inputs are fed directly to the outputs via a series of weight. In this way it can be considered the simplest kind of feed forward network. Multilayer networks overcome many of the limitation of single layer network. The capabilities of multi-layer

networks stem from the non-linearities used with the units. Each neuron in the network receives inputs from other neurons in the network, or receives inputs from the outside world. The output of the neurons are connected to other neurons or to the outside world. Each input is connected to the neurons by a weight.

The neuron calculates the weighted sum of the inputs which is passed through a non-linear transfer function to produce the actual output for the neuron. The most popular non linear transfer function is the sigmoidal type.

The typical sigmoid function has the form –
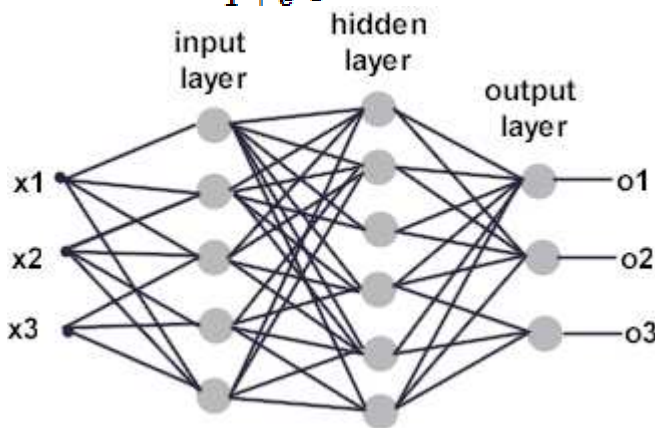
$$f(x) = w \frac{1}{1 + e^{-gx}}$$

Fig-11 Multilayer Perceptron Network

Flowchart of the System

A complete flowchart of special word character recognition is given below in Figure-12
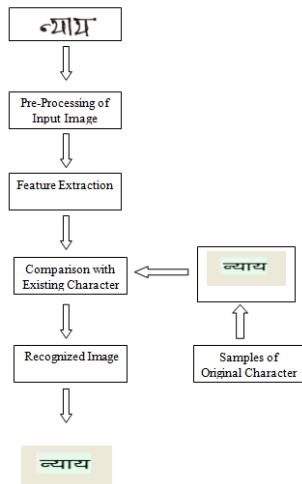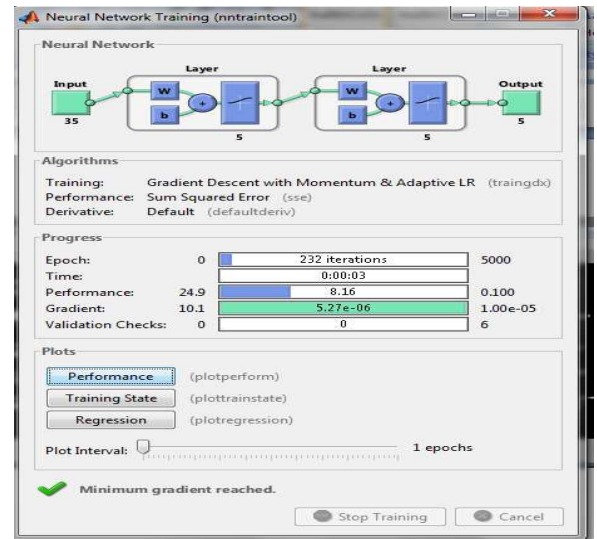

Fig-9 Flowchart of the system

Result

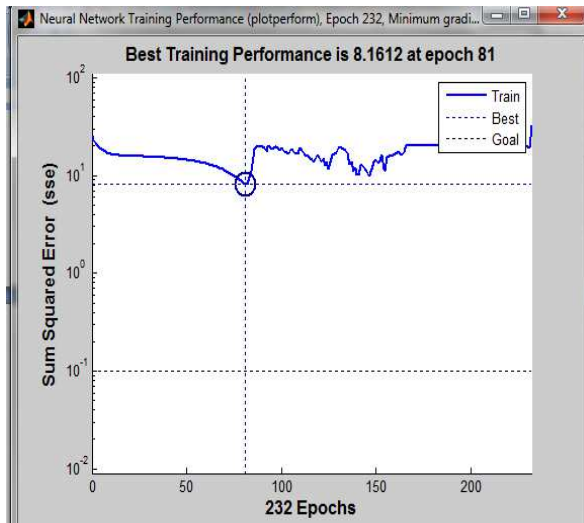| S.No	Image	% Success in Recognition
1	क्या	70
2	मत्स्य	90
3	अश्व	80
4	न्याय	90

Table 1.1 : Different words taken for Recognition and there Accuracy Percentage

We trained the samples for different times and comparing the results on the basis of performance, training state, and regression.
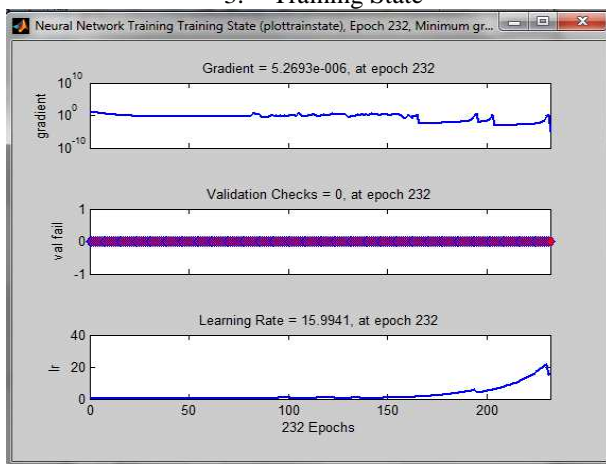
1. Training Set –

2. Training Performance –

3. Training State –

Conclusion & Future Scope

In this paper, a system for recognizing handwritten special words and special characters has been prepared, Off line handwritten Devanagri character recognition is a difficult problem, not only because of the huge variations in person to person handwriting, but also, because of the overlapped and joined characters and also same looking words. A small set of different words and characters of Devanagri script are taken and trained using back propagation algorithm, and testing is done with different sets of samples. It is found that initially accuracy was not satisfactory but after modifying the algorithm and training the samples repetitively accuracy starts increasing. But still few words and characters are there having low accuracy of recognition. Above table shows the result for the system. We get accuracy up-to 90% for some characters.

Future Scope- This project is made for recognition for off-line Dvanagri special characters and words. It can be implemented for on-line characters and words also.

References

[1] " Offline Handwriting Recognition using Genetic Algorithm". Rahul kala, Harsh vajirani, Anupam shukla, and Ritu tiwari. IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010

[2] Anita Pal & Dayashankar Singh, " Handwritten English Character Recognition Using Neural Network ", International Journal of Computer Science & Communication, July-December 2010.

[3] J.Pradeep , E.Srinivasan , S.Himavathi,"Neural Network based Handwritten Character Recognition system without feature extraction",International Conference on Computer, Communication and Electrical Technology – ICCCET,18th & 19th March, 2011.

[4] Sandeep B. Patil, Dr. G.R. Sinha and Vaishali S. Patil, "Isolated Handwritten Devnagri Numerals Recogntion Using HMM", IEEE Catalog Number: CFP1172M-PRT. EAIT in Second International conference on Emerging Applications of Information technology, Kolkota, ISBN: 978-1-4244-9683-9, pp.185-189.

[5] Sandeep B. Patil, G.R. Sinha and Kavita Thakur, 'Isolated Handwritten Devnagri Character Recognition Using Fourier Descriptor and HMM' in International Journal of Pure and Applied Sciences and Technology (IJPAST), ISSN:2229-6107, Volume 8, No 1, January 2012, pp. 69-74.

[6] Alain Biem, "Minimum classification Error Training for on line Handwriting Recognition", IEEE Transaction on Pattern analysis and machine intelligence, 28(7): 1041-1051, 2006.

[7] Huiqin Lin, Wennuan Ou, Tonglin Zhu," The Research of Algorithm for Handwritten Character Recognition in Correcting Assignment System", 2011 Sixth International Conference on Image and Graphics,2011.

[8] Sutha.J, Ramraj.N, "Neural Network Based Offline Tamil Handwritten Character Recognition System", *IEEE International Conference on Computational Intelligence and Multimedia Application, 2007.*

[9] Yuelong Li Jinping Li Li Meng, "Character Recognition Based on Hierarchical RBF Neural

Networks" *Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth* International Conference.

[10] Dayashankar Singh, Maitreyee Dutta and Sarvpal H. Singh, "Neural Network Based Handwritten Hindi Character Recognition", *ACM International Conference (Compute 09), Jan. 9-10, 2009, Bangalore.*

[11] S. K. Parui and B. Shaw. , "Offline handwritten devanagari word recognition: An hmm based approach", In Proc. Of PReMI-2007(Springer), LNCS-4815:528-535, December 2007.

[12] Patil S.B. and Sinha G.R, 'Off-line mixed Devnagri numeral recognition using Artificial Neural Network, Advance in computational Research, Bioinfo journal,ISSN:0975-3273 & E-ISSN: 0975-9085, Volume 4, Issue1, 16-17 March 2012,pp. 38-41.