ISSN (e): 2250 – 3005 || Vol, 05 || Issue,02 || February – 2015 ||
International Journal of Computational Engineering Research (IJCER)

Character recognition of Devanagari characters using Artificial Neural Network

Richa Patil[1], Varunakshi Bhojane[2]
[1](Information Technology ,PIIT/ Mumbai University, India)
[2](Information Technology PIIT / Mumbai University, India)

ABSTRACT:
In this paper we have presented a novel approach for recognizing the handwritten characters in Devanagari script .This paper describes the method and different techniques used for handwritten character recognition system for Devanagari script .In second stage we extract features of the characters into consideration like types of spine ,intersection points ,directional changes and shirolekha .we have used neural network as a classifier which takes the extracted features as input . The proposed algorithm uses the scripting properties the recognition rate achieved with this technique is 80%.

Keywords *- data collection, extracting feature, image normalization, neural network, Segmentation.*

I. INTRODUCTION

This paper analyses different challenges and approaches that are related to Devanagari handwritten character the script in this paper that is taken into consideration is Devanagari which is also script for Hindi, Marathi, Nepali and Sanskrit languages. About More than 450 people use this script.

II. DEVANAGARI SCRIPT

Since 19th century Devanagari has been commonly used scripting language.it is a standardized script .it is the most adopted writing system in the world which consist of 33 consonant and 11 vowels. The modifiers are another constituent symbols which make recognition of Devanagari script more challenging. since character recognition was based on handwritten characters there are number of factor which come into consideration like

- writing of a person
- width of the pen
- speed of writing

In case of Devanagari language the concept of uppercase and lowercase is absent ,Character in Devanagari are formed by holes,strokes and curves

2 block diagram of recognition system

The schematic block diagram of handwritten Devanagari recognition system consist of various stages is shown below.

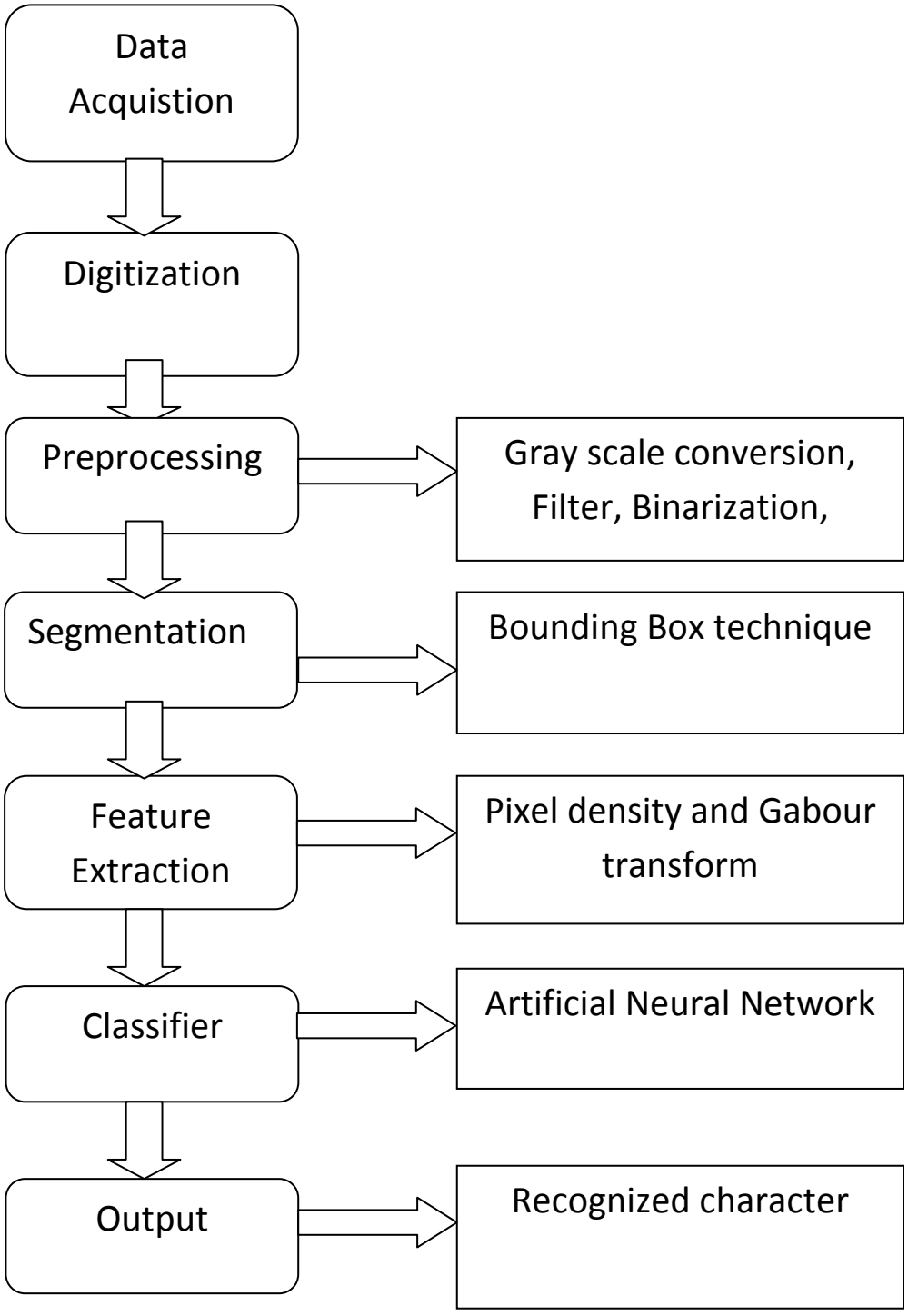

Figure 1

2.1 **Data Acquistion** -Handwritten data samples of character are acquired on paper from different people.

2.2 **Digitization**- The data samples are now scanned using any device which can be a scanner. A flat-bed scanner is used at 300dpi which converts the data on the paper being scanned into a bitmap image [1]

2.3 **Preprocessing** –The digitized image so obtained was in gray tone and stored in TIF format [2]

Preprocessing removes discontinuity and distortion in the input character and convert the character in form of recognizable form .The process which comes under preprocessing are.

- Binarization
- Inversion
- Noise Reduction
- Normalization

2.3.1 Binarization

For testing purpose the character are first written on plain paper with pen or marker pen .These images that are stored via paint are nothing but RGB images. These images are then converted to white and black images with a threshold values of 0.5 to get a true binary image that is 0s for all background pixel and 1s for all foreground pixel.

2.3.2 inversion

The binary images consist of a black foreground as font and white as background. The number pixel in background exceed the one as compared to foreground so we can say number of 1s will atleast be 10 times as compared to number of 0s . We conventionally work with 1s leaving 0s here since we are working only on 1s so we have less number of calculation hence we invert the image so that background is black and foreground is white which gives appearance just like DOS screen.

2.3.3 Noise Reduction

Digital images become prone to different type of noise. Noise is result of error which during image acquisition process gives pixel values that do not reflect true intensities. Noise causes disconnected line segments ,bumps and gaps in lines ,filled loops etc [3].Noise can be introduce by scanner CCD detector, electronic transmission of data .Median filter ,wiener filter and various other filters are used to deal and remove this noise .Gaussian filter is used to smoothing the images and median filter are used to replace the intensity of character image.

2.3.4 normalization

Normalization is the process of converting a random sized images into a standard size image. It is a technique used to reduce shape variation.

2.4 Segmentation – segmentation is a process which analyze the digitalized images provided by a scanning device.This image localize the limit of each character and isolate them from each other .the various steps in segmentation are

(1) The binarized image is checked for Interline space among words.
(2) If interline spaces are detected then the image is segmented into set of paragraphs across the interline gap.
(3) The character in image are now detected and considered as individual objects.
(4) Features are now extracted for these bounding object in feature extraction step.

Figure 2. Character Segmentation

2.5 Feature extraction-

The aim of this step is to recognizing structure and to extract discriminant information from an image of a character and to reduce its dimensions of its representation .In character recognition as in pattern recognition task ,plays a major role in improving the recognition accuracy[4]The feature extraction method we used is Gabor transform which for different sides and we got various output as

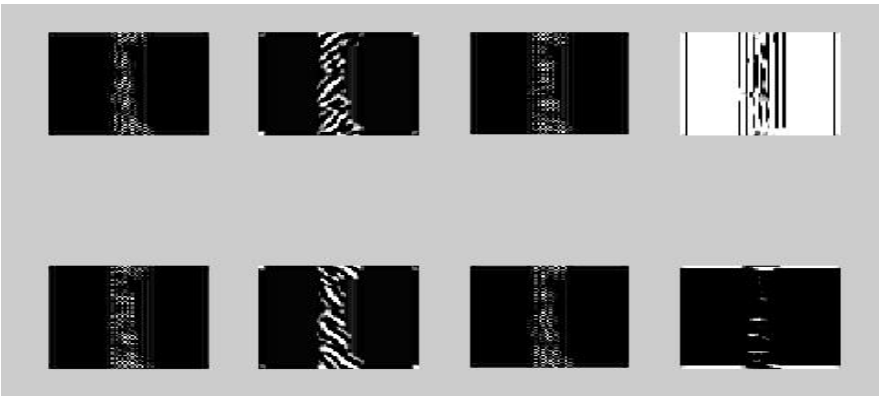

Figure 3 Feature extraction of a character

These technique gives feature like Area, Centroid, Eccentricity , EquiDiameter , Majoraxislength, Extent, Minoraxislength, orientation, perimeter.

2.6 Classification

The extracted features act as an input in the classification process.A bag-of-key feature are used as an input and given to classifier such as ANN.Here we have used Probablistic ANN to recognize character the feature extracted from the character are given as input to ANN.

2.6.1 Probabilistic Neural Network (PNN)

In a PNN, the operations are organized into a multilayered feedforward network with four layers:

- Input layer
- Hidden layer
- Pattern layer/Summation layer
- Output layer

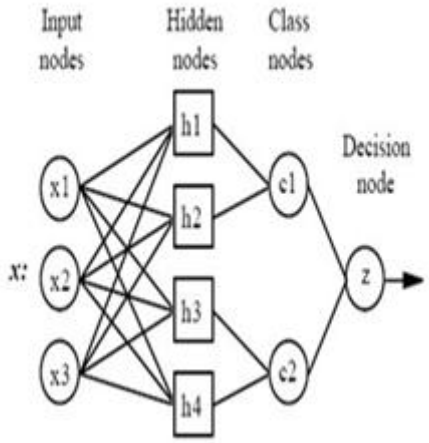

Figure 4 PNN architecture

III. CONCLUSION

A lot of research work exists in the survey for Devanagari Handwritten recognition.. In this paper, we have projected various aspects of each phase of offline Devanagari character recognition process that have been used. Researchers have used many character set for research. The following key challenges are more to be carried out by researchers by increasing number of holes and strokes and mixed words.

REFERENCES

Journal Papers:

[1] Gunjan Singh et al, Recognition of Handwritten Hindi Characters using Backpropagation Neural Network / *(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (4) , 2012,4892-4895*

[2] Sneha U.Bohra et al ./ Handwritten Character Recognition for major Indian Scripts:*A Survey International Journal of Computer Science & Engineering Technology*

[3] Vikas J Dongre,Vijay H Mankar, A Review of Research on Devnagari Character Recognition, *International Journal of Computer Applications (0975 – 8887) Volume 12– No.2, November 2010*

[4] Sushma shelke, Shaila Apte, Multistage Handwritten Marathi Compound Character Recognition Using Neural Networks, *Journal of Pattern Recognition Research 2 (2011)* 253-268.