Survey and Analysis of Devnagari Character Recognition Techniques using Neural Networks

3 authors, including:

Neha Sahu
Mody University of Science and Technology
5 PUBLICATIONS **26** CITATIONS

SEE PROFILE

Indu Kashyap
Manav Rachna International University
20 PUBLICATIONS **65** CITATIONS

SEE PROFILE

Survey and Analysis of Devnagari Character Recognition Techniques using Neural Networks

Neha Sahu
Student, CS deptt
FET, Manav Rachna
International University,
Faridabad

R. K. Rathy, PhD.
Prof, CS deptt.
FET, Manav Rachna
International University,
Faridabad

Indu Kashyap
Asst Prof, CS deptt.
FET, Manav Rachna
International University,
Faridabad

ABSTRACT

English Character Recognition techniques have been studied extensively in the last few years and its progress and success rate is quite high. But for regional languages these are still emerging and their success rate is moderate. There are millions of people who speak Hindi and use Devnagari script for writing. As digital documentation in Devnagari script is gaining popularity. Research in Optical Character Recognition (OCR) is very essential especially with an eye on its applications in banks, post offices, defense organizations, library automation, etc. Devnagari Optical Character Recognition needs more attention as it is national language and there is less development in this field due to complexity in the script.

This paper describes the current techniques being used for DOCR. The overview of the system is explained with the available techniques and their current status.

Keywords

Devnagari Character Recognition, Off-line Handwriting Recognition, Segmentation, Feature Extraction, OCR.

1. INTRODUCTION

Optical Character Recognition is a technique that works on the system based on how data is collected and processed leading to reorganization of characters in the text. All these systems can be classified in two categories: (i) Systems classified according to the data acquisition techniques, i.e. Online or Offline. (ii) Systems classified according to the text type i.e. print- written or hand-written. The main objective of DOCR system is to recognize Devnagari characters which are in the form of digital image. This is done by searching a match between the features extracted from the given character's image and the library of image models. The library helps in distinction of features between the character images. This eliminates the confusion for correct character recognition.

2. FEATURES OF DEVNAGARI SCRIPT

Devnagari script has 11 vowels and 33 consonants. They are called basic characters. Vowels can be written as independent letters, or by using a variety of diacritical marks called *hallant* which are written above, below, before or after the consonant they belong to. When vowels are written in this way they are known as *modifiers* and the characters so formed are called *conjuncts*. When a vowel follows a consonant, it is written in its respective *maatraa* form, which is appended to the consonant. Sometimes two or more consonants can combine and take new shapes. These new shape clusters are known as

compound characters. A sample of Devnagari character set is provided in figure 1 to 6 [1].

Fig 1: Vowels And Corresponding Modifiers.

Fig 2: Consonants

Fig 3: Half Form Of Consonants With Vertical Bar.

Fig 4: Some Special Combination Of Half-Consonant And Consonant.

Fig 6: Special Symbols

Architecture of DOCR: The block diagram below will show all the steps in DOCR.
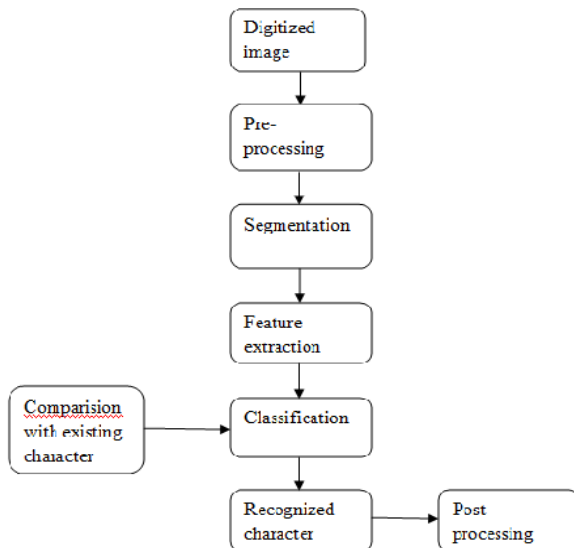
Fig 7: Architecture if DOCR System

3. PREPROCESSING OF IMAGE

Whenever data is collected for recognition, it is collected as an optically scanned image of the paper document. Text digitization is done using a flat bed scanner having resolution between 100 and 600 dpi and stored in a file of picture elements,

Fig 8: Steps in preprocessing of Images

called pixels. These pixels may have values: OFF (0) or ON (1) for binary images, 0– 255 for gray-scale images, and 3 channels of 0–255 colour i.e. RGB values for colour images. This collected raw data must be further analyzed to get useful information. Pre-processing essentially enhances the image rendering it suitable for segmentation. Such processing includes the following:

3.1 RGB to Gray Conversion:

The scanned image is stored as a JPEG image but images of other formats like BMP, TIFF etc are also used for recognition. All these images are in RGB format. In order to start working with an image it must be converted into a gray scale image. A gray scale image represents an image as a matrix where every element has a value corresponding to how bright/dark the pixel at the corresponding position should be colored. The value of a pixel lies between 0 to 1 or between 0 to 255 depending upon its class.

3.2 Thresholding/Binarization:

Binarization process converts a gray scale image into a binary image using global thresholding technique. This image format also stores an image as a matrix but can only color a pixel black or white. It assigns a 0 for black and a 1 for white. The image is then inverted to obtain image such that object pixels are represented by 1 and background pixels by 0 because it is preferred to work with 1s and leave the 0s aside. Also, smaller number 1s will mean lesser calculations.

3.3 Noise Reduction:

The noise introduced by the optical scanning device or the writing instrument, causes disconnected line segments, bumps and gaps in lines, filled loops etc. The distortion including local variations, rounding of corners, dilation and erosion, is also a problem. Median filter is a process that replaces the value of a pixel by the median of gray levels in the neighborhood of that pixel.

3.4 Skew Detection and Correction:

Skew Detection refers to the tilt in the bitmapped image of the scanned paper for OCR. It is usually caused if the paper is not fed straight into the scanner. Most of the algorithms are sensitive to the orientation (or skew) of the input document image, making it necessary to develop algorithms, which can detect and correct the skew automatically. Skewed lines are made horizontal by calculating skew angle and making proper correction in the raw image.

3.5 Thinning:

Thinning is a morphological operation process that is used to remove selected foreground pixels from the binary images and thin the images to single-pixel width level so that their contours are brought out more vividly. In this way, the attributes to be studied later will not be affected by the uneven thickness of edges or lines in the symbol. Various standard functions are now available in MATLAB for above operations.

3.6 Edge Detection, Dilation and Filling:

The boundary detection of image is done to enable easier subsequent detection of pertinent features and objects of interest. Detection of edges in the binarized image is done using sobel technique. After locating the edges the image is dilated and the holes present in the image are filled. These are the operations performed in the last two stages to produce the pre-processed image suitable for segmentation .

4. SEGMENTATION

Segmentation is one of the most important phases of DOCR system. Segmentation subdivides an image into its constituent regions or objects which are certainly characters. This is needed because classifier recognizes these characters only. The process of segmentation mainly follows the following pattern:
1) Identify the text lines in the page.
2) Identify the words in individual line.
3) Finally identify individual character in each word.
Histogram based technique is one technique used for segmentation of lines and words [1]. Once words are segmented from lines, head line is located and removed. After that word is divided into three horizontal parts known as upper zone, middle zone and lower zone. Individual characters are separated from each zone by applying vertical scanning.

Fig 9: Segmentation of Images

5. CHARACTER FEATURE EXTRACTION

Next step is extracting features from the segmented characters and to distinguish it from other characters by comparing it with already stored patterns of all characters in the library. 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 so far. Various feature extraction methods are classified in three major groups:

5.1 Global Transformation and Series Expansion:

It includes few methods like Fourier Transforms, Wavelets and Moments [6].

5.2 Geometrical and Topological Features:

It includes various methods like Extracting and Counting Topological Structures [4] [7], Measuring and Approximating the Geometrical Properties [7-8] and coding [8-9] .

Various other techniques have also been used by various researchers few among them are briefly described below:

5.3 Shadow Features of Character:

For computing shadow features [10], the rectangular boundary enclosing the character image is divided into eight octants, for each octant shadow of character segment is computed on two perpendicular sides so a total of 24 shadow features are obtained. Shadow is basically the length of the projection on the sides. [3].

5.4 Chain Code Histogram of Character Contour:

For a scaled binary image, first the contour points of the character image are located. If any of the 4-connected neighbor points is a background point then the background point, is considered as contour point [3].

5.5 View based features:

This feature extraction method examines four "views" of each character extracting from them a characteristic vector, which describes the given character. The view is a set of points that plot one of four projections of the object (top, bottom, left and right) – it consists of pixels belonging to the contour of the character and having extreme values of one of its coordinates [3].

5.6 Diagonal feature extraction scheme:

This scheme is used in [2] where every character image of size 90x 60 pixels is divided into 54 equal zones, each of size 10x10 pixels. The features are extracted from each zone pixels by moving along the diagonals of its respective 10x10 pixels.

6. CHARACTER CLASSIFICATION:

DOCR classification techniques can be classified as follows.
1. Template Matching.
2. Statistical Techniques.
3. Neural Networks.
4. Support Vector Machine (SVM) algorithms.
5. Combination classifier.

The above approaches are neither necessarily independent nor disjoint from each other. Occasionally, a CR technique in one approach can also be considered to be a member of other approaches
.

6.1 Template Matching:

This is the simplest way of character recognition, based on matching the stored prototypes against the character or word to be recognized. The matching operation determines the degree of similarity between two vectors (group of pixels, shapes, curvature etc.) A gray-level or binary input character is compared to a standard set of stored prototypes. The recognition rate of this method is very sensitive to noise and image deformation. For improved classification Deformable Templates and Elastic Matching are used [6].

6.2 Statistical Techniques:

Statistical decision theory is concerned with statistical decision functions and a set of optimality criteria, which maximizes the probability of the observed pattern given the model of a certain class. The major statistical methods, applied in the OCR field are Nearest Neighbor (NN) [11-12], Hidden Markov Modeling (HMM) [8], Fuzzy Set Reasoning [14], Quadratic classifier [15].

6.3 Neural Networks Techniques:

Character classification problem is related to heuristic logic as human beings can recognize characters and documents by their learning and experience. Hence neural networks which are more or less heuristic in nature are extremely suitable for this kind of problem. Various types of neural networks are used for OCR classification.

A neural network is a computing architecture that consists of massively parallel interconnection of adaptive 'neural' processors. Because of its parallel nature, it can perform computations at a higher rate compared to the classical techniques. Because of its adaptive nature, it can adapt to changes in the data and learn the characteristics of input signal. Output from one node is fed to another one in the network and the final decision depends on the complex interaction of all nodes.

Several approaches exist for training of neural networks viz. error correction, Boltzman, Hebbian and competitive learning. They cover binary and continuous valued input, as well as supervised and unsupervised learning.

Neural network architectures can be classified as, feed-forward and feedback (recurrent) networks. The most common neural networks used in the OCR systems are the multilayer perceptron (MLP) of the feed forward networks and the Kohonen's Self Organizing Map (SOM) of the feedback networks. MLP is proposed by U. Bhattacharya *et al.* [11-12]. K. Y. Rajput *et al.* [16] used back propagation type NN classifier. Genetic algorithm based feature selection and classification along with fusion of NN and Fuzzy logic is reported in English [8] but no any work is reported for Indian languages.

6.4 Support Vector Machine Classifier:

It is primarily a two-class classifier. Width of the margin between the classes is the optimization criterion, i.e. empty area around the decision boundary defined by the distance to the nearest training patterns. These patterns, called support vectors, finally define the classification function. Their number is minimized by maximizing the margin. The support vectors replace the prototypes with the main difference between SVM and traditional template matching techniques is that they characterize the classes by a decision boundary. Moreover, this decision boundary is not just defined by the minimum distance function, but by a more general possibly nonlinear, combination of these distances.

6.5 Combination Classifier:

Various classification methods have their own superiorities and weaknesses. Hence many times multiple classifiers are combined together to solve a given classification problem. Different classifiers trained on the same data may not only differ in their global performances, but they also may show strong local differences. Each classifier may have its own region in the feature space where it performs the best. Some classifiers such as neural networks show different results with different initializations due to the randomness inherent in the training procedure. Instead of selecting the best network and discarding the others, one can combine various networks, thereby taking advantage of all the attempts to learn from the data.

7. COMPARATIVE ANALYSES OF DIFFERENT APPROACHES

Various methods proposed by many researchers have achieved different levels of success and accuracy rates depending upon the technique used. Based upon numeral results by various researchers

Table 1: Comparison of Numeral Results by Researchers.

S.N.	Method proposed by	Data size	Accuracy obtained
1	R. Bajaj *et al*. [9]	400	89.6%
2	R. J. Ramteke *et al*. [6]	169	92.28%
3	U. Bhattacharya *et al*. [18]	16273	95.64%
4	N. Sharma *et al*. [15]	22,556	98.86%

Based on character results, Kumar and Singh [19] obtained 80% accuracy with 200 data size and N. Sharma *et al*. [15] 80.36% with 11270.

8. CONCLUSION AND FUTURE WORK

This paper has concentrated on an appreciation of principles and methods. Current research employs models not only of characters, but also words and phrases, and even entire documents. In order to have high reliability in character recognition, segmentation and classification have to be treated in an integrated manner to obtain more accuracy in complex cases. Present work has not attempted to compare the effectiveness of various algorithms. It would be difficult to assess techniques separate from the systems for which they were developed

9. REFERENCES

[1] Vikas J Dongre ,Vijay H Mankar, "devnagari document segmentation using histogram approach", *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT),* Vol.1, No.3, August 2011, pp. 46-53.

[2] J.Pradeep, E.Srinivasan, S.Himavathi, "Diagonal Based Feature Extraction For Handwritten Character Recognition System Using Neural Network",*Electronics Computer Technology (ICECT),*2011, Volume-4, 2011, pp. 364-368.

[3] S. Arora, D. Bhattacharjee, M. Nasipuri , D.K. Basu, M.Kundu, L.Malik, "Study of Different Features on Handwritten Devnagari Character", *Second International Conference on Emerging Trends in Engineering and Technology, ICETET-09,* 2009, pp. 929-933.

[4] D. Trier, A. K. Jain, T. Taxt, "Feature Extraction Method for Character Recognition - A Survey*", Pattern recognition,* vol.29, no.4, pp.641-662, 1996.

[5] J´erˆome Revaud, Guillaume Lavou´e and Atilla Baskurt, "Optimal similarity and rotation angle retrieval using Zernike moments", July 16, 2007.

[6] R. J. Ramteke, S. C. Mehrotra, "Recognition of Handwritten Devnagari Numerals*", International Journal of Computer Processing of Oriental Languages,* 2008.

[7] Santanu Chaudhury, Geetika Sethi, Anand Vyas, Gaurav Harit, "Devising Interactive Access Techniques for Indian Language Document Images", *(ICDAR 2003).*

[8] Tapan K Bhowmik, Swapan K Parui Utpal Roy, "Discriminative HMM Training with GA for Handwritten Word Recognition", *IEEE,* 2008.

[9] Reena Bajaj, Lipika Dey , Santanu Chaudhury, "Devnagari numeral recognition by combining decision of multiple connectionist classifiers", *Sadhana* Vol. 27, Part 1, pp. 59–72, February 2002.

[10] S. Basu, N.Das, R. Sarkar, M. Kundu, M. Nasipuri, D.K. Basu, "Handwritten Bangla alphabet recognition using MLP based classifier", NCCPB, Bangladesh, 2005.

[11] Sandhya Arora *et al.*, "Performance Comparison of SVM and ANN for Handwritten Devnagari Character Recognition", *IJCSI International Journal of Computer Science Issues*, Vol. 7, Issue 3, May 2010.

[12] U. Bhattacharya, S. Vajda, A. Mallick, B. B. Chaudhuri, A. Belaid, "On the Choice of Training Set, Architecture and Combination Rule of Multiple MLP Classifiers for Multiresolution Recognition of Handwritten Characters", *9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004).*