# A Hybrid Approach towards Cost Effective Model for Handwritten Character Recognition

Nupur Chauhan
Student, M. tech
Dept. of Computer Science and Engineering
Graphic Era University
Dehradun, India

Manish Sharma
Assistant Professor
Dept. of Computer Science and Engineering
Graphic Era University
Dehradun, India

Pooja Singh
Student, M. tech
Dept. of Computer Science and Engineering
Graphic Era University
Dehradun, India

## ABSTRACT
Handwritten character is gaining a lot of attention in the area of pattern recognition as its applications in various fields are increasing day by day. HCR system is providing us with a key factor to a paperless environment. Feature Extraction is a key part for a cost effective model for handwritten character recognition. Effective features improve the recognition rate and misclassification. A hybrid model provides better performance in comparison of the individual. Convolution neural networks are viewed to be more efficient to optimize the recognition ability of HCR system.

## General Terms
Pattern Recognition, image processing, HCR

## Keywords
Features, classification, cost, convolution neural network

## 1. INTRODUCTION
Character recognition is an active area in the pattern recognition. Character recognition, usually abbreviated to optical character recognition or shortened OCR. It is the electronic translation of images of handwritten, typewritten or printed text which are usually captured by a scanner or camera, into machine-editable text. A text document can be classified into either handwritten document or machine printed document. People generally wish to convert text documents to some sort of digital representation so that they can scan in a document and have the text of that document available in a word processor. It has many practical applications like postal address and zip code recognition, forms processing, automatic processing of bank cheques, identification of human being or its personality and behavior through his handwritten characters, etc. We have multiple approaches for recognizing handwritten characters but here our focus is on cost of recognition model in terms of time and accuracy. Every HCR system goes under four phases, i.e. pre-processing, segmentation, feature extraction and recognition or classification.

We can categorize all the above steps into three phases. In the first phase document image is taken as input and pre-processing is done on this input image. In pre-processing we have different steps like firstly we have to convert RGB image into gray scale image then binarise it (Otsu's method) then appropriate operations are applied like noise removal, smoothing, image enhancement, dimensionality reduction etc. After pre-processing segmentation is done i.e. lines are segmented from document image then words and then particular characters are extracted. Now particular characters are further processed in the second phase. Features are extracted of these characters. These features are used to create models for recognition. These models come into second phase. In third phase recognition process is done using some classifiers. Figure 1 represents all the three phases.

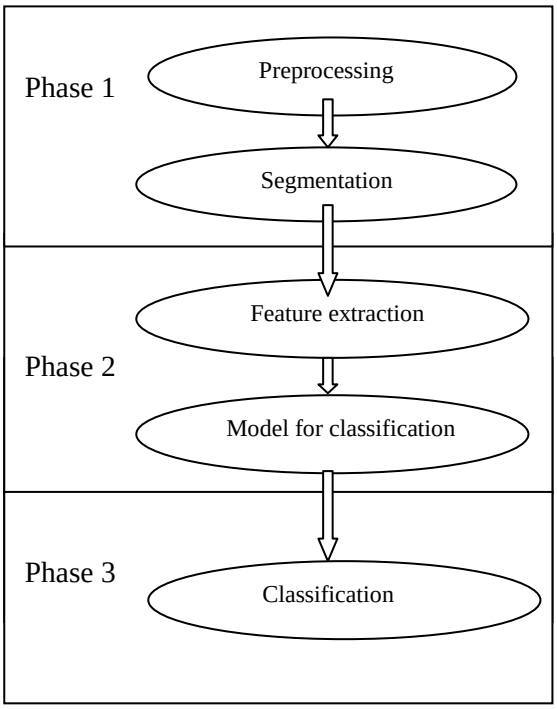


**Figure 1: A generic model for character recognition process**

The rest of the paper is organized as follows. Section 2 describes the various ways to obtain the input images of handwritten document and section 3 describes various ways to pre-process the input text i.e. phase 1. Section 4 describes phase 2 which includes feature extraction methods to build up recognition model. Section 5 describes about phase 3 which includes recognition of characters as final output.

## 2. DATA COLLECTION
Input image can be obtained in three ways 1) we can use a standardized database, like IAM handwriting database, MNIST database, 2) we can capture input image directly from a webcam or 3) we can create a handwritten database ourselves by taking samples of handwritten characters from different persons.

## 2.1 IAM Handwriting Database

This database comprises various forms of handwritten English text which can be used for training and testing of handwritten text recognizer and for performing writer identification and verification experiments. In this database handwritten text is scanned images having a resolution of 300 dpi and saved as PNG images with 256 gray levels. The IAM handwriting database 3.0 contains handwriting samples of 657 writers, 1539 pages of scanned text, 5685 isolated and labeled sentences, 13353 isolated and labeled text lines and 115320 isolated and labeled words.

## 2.2 MNIST Database

The MNIST (Mixed National Institute of Standards and Technology) database is a large database of handwritten digits. It contains 60,000 training and 10,000 testing images.

# 3. PRE-PROCESSING

Preprocessing converts an input character into a standard format for making it suitable for further processes and for extracting information from the image.

## 3.1 Noise removal

It involves noise removal using certain filtering operations like Spatial filter. Proper filter like median filter, mean filter, min-max filter, Gaussian filter, etc. may be applied to remove noise from the document.

## 3.2 Binarization

Binarization is binary (0/1) representation of image from gray scale. Otsu method is a global binarization technique and also has a short running time. The Mean running time for the Otsu's binarization method was 2.0 seconds and this is one of the lowest running times.

## 3.3 Dimensionality Reduction

Dimensionality of image can be reduced using PCA (a well-known method for dimension reduction).

## 3.4 Normalization

The images just before the segmentation stage are converted to certain standard sizes. Linear normalization maintains dimension of input vector but distort the shape of the image whereas non-linear transformation doesn't distort the shape of image to maintain input vector dimension.

## 3.5 Slant Correction

When the document is scanned, it may not be perfectly horizontally aligned, so it is required to align it by performing slant angle correction. Then the slant correction is performed on the image after that it is resized to a standard size. For this near-vertical strokes are located in the character and then average slant of the character is estimated from these strokes. Then shear transformation is used for slant correction in the character image.

## 3.6 Thinning

Morphological skeletonization is performed on resized images. Skeletonization means redundant pixels which do not form the backbone of image are deleted.

## 3.7 Spur removal

Due to some small irregularities in the boundary of the character image skeleton of the image can have some undesirable short spurs. For removal of these spurs pruning (thinning) is used.

## 3.8 Image enhancement

Here the filtered images are enhanced using certain high boost filter makes and histogram equalization technique.

## 3.9 Segmentation

Segmentation is the process of decomposition of an image into sub images. The first step of the segmentation process is segmenting the text region into lines, also called as line segmentation. The RLSA (run length smoothing algorithm) can be used for line detection. Then words are segmented from lines. Word segmentation is easier than line segmentation and character segmentation. The space between two words is generally more than three pixels. Words are segmented by the projection based method [8]. The possible segmentation regions in the text image are determined by using projection profiles and topographic features extracted from the grayscale images.

# 4. FEATURE EXTRACTION

Feature extraction plays very important role in the pattern recognition process. The performance of the recognition system depends on how efficient are the feature sets. Efficient features are those which are useful in better classification. We categorized feature extraction in phase two where a model for classification is developed. It is the main focus of a system that to develop an efficient model for classification so that characters can be classified easily and takes less time for recognition. So we can say that the cost of the recognition system depends on phase two. There are various techniques for extracting features. Olarik Surinta et. al [1] compares feature based and pixel based methods for recognizing Bangla digits. They compare four techniques for feature extraction. Contour Angular Technique (CAT), Hotspot Technique (HOT), Gray pixel based method (GBP) and Black and white downscale method (BWS). The best feature extraction technique CAT outperforms the best pixel based method when the training dataset is not very large. When the training dataset size increases, the best pixel-based method slightly outperforms this feature extraction method. Table 1 shows the comparison of these

Feature extraction is the main part of any pattern recognition application. Feature extraction techniques like Linear Discriminant Analysis (LDA), Principle Component Analysis (PCA), Independent Component Analysis (ICA), Chain Code (CC), zoning, Scale Invariant Feature Extraction (SIFT), Gradient based features, Histogram might be applied to extract the features of individual characters. These features are used to train the recognition system. Global methods like DCT and DWT are faster, whereas local methods are sensitive to noise, improper segmentation, image size and writing style.

**Table 1. Comparison between pixel based and feature based methods**

| Technique | Description | Type | No. of features | Time Taken for classification |
|---|---|---|---|---|
| **CAT** | The method divides the handwritten character into 16 non-overlapping blocks and considers the contour of the handwritten image as 8-directional codes | Feature based | 128 | 37ms |
| **HOT** | Distance between evenly spaced hotspots and the closest black pixels in each direction is used to describe the whole handwritten image | Feature based | 100 | 35ms |
| **BWS** | The black and white handwritten image is partitioned into 9×9 non-overlapping blocks. For each block the number of black pixels is computed | Pixel Based | 81 | 29ms |
| **GPB** | Uses the raw pixel intensities of the handwritten images | Pixel Based | 784 | 73ms |

J. Pradeep et. al [2] proposed diagonal based feature extraction in which every character image of size 90x 60 pixels is divided into 54 equal zones, where size of each zone is 10x10 pixels. The features are extracted from each zone pixels by moving along the diagonals of its respective 10X10 pixels. Each zone has19 diagonal lines and every single diagonal line is summed to get a single sub-feature, so total 19 sub-features are obtained from the each zone. These 19 sub-features values are averaged to form a single feature value and placed in the corresponding zone. This procedure is repeated for the all the zones. The feature values corresponding to the zones where diagonals are empty of foreground pixels are zero. Finally, 54 features are extracted for each character. In addition, 9 features are taking by averaging the values placed in zones row wise and 6 features are taking by averaging the values in zones column wise. So, every character is represented by a total of 69 (54 +15) features. From the test results it is identified that the diagonal method of feature extraction yields the highest recognition accuracy of 98% for 54 features and 99% for 69 features.

El-Sayed M. El-Alfy [3] used 16 non-Gaussian topological features. These features are as follows:

**Convexity based 4 features:**

- Whether there is a curve on the upper right,
- Whether there is a curve on upper left,
- Whether there is a curve on lower right,
- Whether there is a curve on lower left

**Width based 3 features:**

- Whether width of the pixel is greater than 7 pixels on top 6,
- Bottom 6 and
- All rows

**Loop related 5 features:**

- No. of loops,
- The top loop location,
- Second loop location,
- Width of the horizontal bar below the top loop,
- No. of pixels in the top row, which have two regions

**Energy distribution related 3 features:**

- Measure of ratio of no. of pixels in right half to the no. of pixels in the left half of image
- Ratio of no. of pixels in the top 10 rows to the no. of pixels in bottom 10 rows
- Ratio of no. of pixels in upper half to no. of pixels in lower half

**Slope related 1 feature:**

- Sum square error for a straight line fit to the rightmost pixel of the character image

Feature extraction using Euler no. may be efficient as compared to former methods. In this approach first the Euler Number is computed. Based on the result of computation whether the result is positive or negative or zero, the character set is divided into three groups. Euler Number is defined as the number of connected components in the image minus the number of holes. This will divide the characters into 3 groups [9].

- Euler number equal to one and this contains: s, S, f, F, G, h, H, j, J, k, K, l, L, z, Z, x, X, c, C, v, V, n, N, m, M, w, W, E, r, t, T, y, Y, u, U, i, I.
- Euler number equal to zero and this contains: q, Q, R, o, O, p, P, a, A, d, D, g, b.
- Euler number equal to minus one and this contains: B [4].

Feature extraction using convolution network is presented in [5]. Convolution networks are similar to LeNet 5 in architecture (Figure 2). Convolution neural networks handle images directly. The advantage of using convolution network is that learning features with convolutional neural networks is better than using handcrafted features for handwritten word recognition.

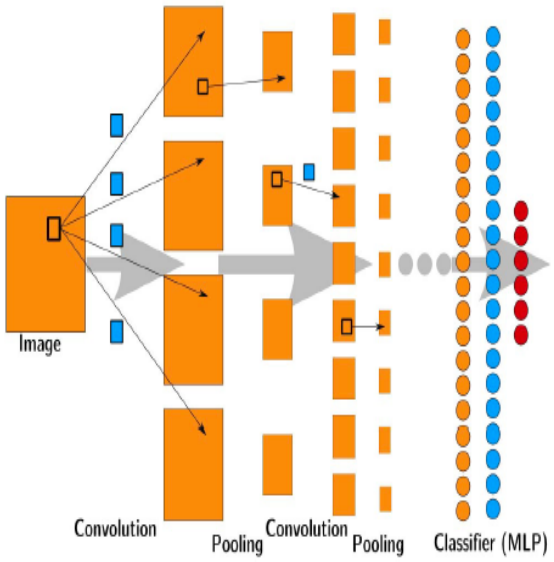


**Figure.2. Convolution neural network [5]**

HMMs are combined with convolution neural networks and outperform the systems based on pixels or handcrafted features. These two can be combined in two ways. First one is a hybrid combination in which Convolution neural network is trained to predict the HMM state t corresponding to a given observation o (i.e. p(t|o)). Second one is a tandem scheme in which neural network extract features, used to train a new standard GMM-HMM.

## 5. CLASSIFICATION

Now when we have extracted features which represent individual characters, classifiers are trained using these features. There are many classifiers like SVM, K-NN, HMM, ANN etc.

There is no such measure to compare them that one is superior to another. Combining multiple character recognition techniques gives better results. R.Radha et. al [6] proposes a new hybrid classification technique for recognizing printed digits. Combining multiple classifiers gives better results in comparison of using single classifiers. [1] Used majority voting technique to combine the output of different classifiers.

The Convolution neural network combines outputs of various neural networks and gives better results [7]. Convolutional neural networks have multiple layers of collection of small neuron which look at the small parts of the input image. The results of these collections are then combined so that they overlap to get a better representation of the original image. This procedure is repeated for every such layer. Due to this, convolution neural networks are robust to translation of the input image. Convolutional networks include pooling layers which can be global or local. These layers combine the outputs of neuron clusters from previous layers. They also consist of various combinations of convolutional layers and fully connected layers, and point wise nonlinearity is applied at the end of or after each layer. One major advantage of convolutional networks is that it uses shared weight in convolutional layers, thus the same filter (weights bank) is used for each pixel in the layer. This reduces required memory size and also improves performance. An error rate of 0.23 percent is achieved with the convolution neural networks on the MNIST database, which is the lowest, achieved on the database as of February 2012.

## 6. CONCLUSION

The research studies offered in this paper shows various aspects and models of handwritten character recognition systems. The whole system is divided here into three phases. Phase two is responsible for efficient system where features are extracted to prepare models for recognition. A comparative analysis is presented between feature based and pixel based methods and feature based methods are found to be more effective in comparison of pixel based methods when dataset is of small size, but as the size of a dataset increases pixel based methods perform well. Efficient features lead to efficient recognition. Here various techniques are discussed among them feature extraction and classification using convolution networks is quite efficient.

## 7. ACKNOWLEDGMENTS

The authors acknowledge Graphic Era University and Mr. Manish Sharma for the contribution and support towards completing this research work.

## 8. REFERENCES

[1] Olarik Surinta, Lambert Schomaker and Marco Wiering, A Comparison of Feature and Pixel-based Methods for Recognizing Handwritten Bangla Digits, 12th International Conference on Document Analysis and Recognition 2013.

[2] J. Pradeep, E. Srinivasan and S. Himavathi, DIAGONAL BASED FEATURE EXTRACTION FOR HANDWRITTEN ALPHABETS RECOGNITION SYSTEM USING NEURAL NETWORK, International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 1, Feb 2011.

[3] El-Sayed M. El-Alfy, A Hierarchical GMDH-Based Polynomial Neural Network for Handwritten Numeral Recognition Using Topological Features, ©2010 IEEE.

[4] Om Prakash Sharma, M. K. Ghose, Krishna Bikram Shah, An Improved Zone Based Hybrid Feature Extraction Model using Euler Number, International Journal of Soft Computing and Engineering (IJSCE'12), ISSN 2231-2307, Volume -II, Issue- II, Article no-96, pp. 154-158.

[5] Theodore Bluch, Hermann Ney, Christopher Kermovant, Feature extraction with convolutional neural networks for handwritten word recognition, 2013 12th International Conference on Document Analysis and Recognition © 2013 IEEE.

[6] R. Radha, R.R Aparna, Digit Recognition Using Hybrid Classifier, 2014 World Congress on Computing and Communication Technologies © 2013 IEEE

[7] Parag Dhawan, Snehlata Dongre, D. J. Tide, Hybrid GMDH Model for Handwritten Character Recognition, ©2013 IEEE.

[8] Saiprakash Palakollu, Renu Dhir, Rajneesh Rani, Handwritten Hindi Text Segmentation Techniques for Lines and Characters, Proceedings of the World Congress on Engineering and Computer Science 2012 Vol I, WCECS 2012, October 24-26, 2012, San Francisco, USA.