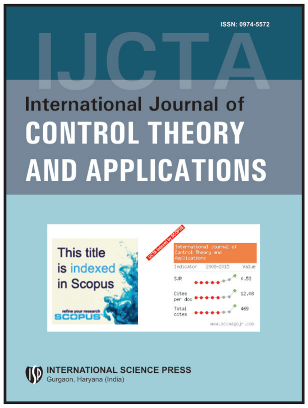

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 24 • 2017

Design of High Efficient and High Recognition Rate for Real Time Handwritten Recognition using HMM and ANN Classification

Vinita Patil[a] and Rajendra Patil[b]

[a]*Research Scholar, Appa IET, Kalburagi. Email: vinita278@yahoo.co.in*
[b]*Associate Professor Department of ECE, Appa IET, Kalburagi. Email: rajurpatil@yahoo.com*

Abstract: Handwritten character recognition has been an active and challenging research problem. Most of the traditional methods have two challenges, due to the large variations of characters and the dependency relationship between characters. First, in real applications, words may be written cursively, so it is hard to identify the words automatically. Even if the words are neat, different people may write the same words in different styles. Since there are large shape variations in human handwriting, recognition accuracy of handwritten words is very difficult. Finally, dependency relationship between characters has not been well utilized. For these reasons handwritten character recognition has become a highly active and challenging research problem. To recognize the handwritten words, in the proposed work the hybrid mode that combines Artificial Neural Network (ANN) and Hidden Markow Model (HMM). Due to the large variations of words, a feature extractor which could obtain transnationally invariant features is necessary. ANNs are well suited for handwritten character recognition problems. The output of each layer of ANN can be regarded as features directly, but the features maybe non-linear separable.

Keywords: Handwritten recognition, HMM, pre-processing, segmentation and ANN.

1. INTRODUCTION

The history of handwriting recognition systems is not complete without mentioning the Optical Character Recognition (OCR) systems which preceded them. Optical Character Recognition (OCR) is a problem recognized as being as old as the computer itself. There have been many papers and technical reports published reviewing the history of OCR technologies. Modern OCR was said to have begun in 1951 due to an invention by M. Sheppard called GISMO, a robot reader-writer. In 1954, a prototype machine developed by J. Rainbow was used to read uppercase typewritten letters at very slow speeds. By 1967, companies such as IBM finally marketed OCR systems. However in the late 60's, these systems were still very expensive, and therefore could only be used by large companies and government agencies. Today, OCR systems are less expensive and can recognize more fonts than ever before. Even so it is important to note that in some situations these commercial packages are not always satisfactory. Senior mentions that problems still exist with unusual character sets, fonts and with documents of poor quality. Research now focuses more on hand-printed numeral, character and joined/cursive handwriting recognition. Unfortunately the success of OCR could not carry on to handwriting recognition, due to the variability in people's handwriting. As for the recognition of isolated handwritten numerals, Suen, details many researchers which have already obtained very promising results using various classification methods. The purpose of the project is to take handwritten Kannada characters as input, process the characters, train the neural network algorithm, to recognize the pattern and modify the character to a beautiful version of the input. Pattern recognition is perhaps the most common use of neural network. The neural network [10] is presented with a target vector and also a vector which contains the pattern information, this could be of image and hand written data. The neural network then attempts to determine if the input data matches a pattern that the neural network has memorized. Here we use the Convolutional Neural Network [11] for training process. Convolutional Neural Network (CNN) has many strengths. First, feature extraction and Classification are integrated into one structure and are fully adaptive. Second, the network extracts 2-D image features at increasing dyadic scales. Third, it is relatively invariant to geometric, local distortions in the image. CNN has been used for in several applications including handwritten recognition, face detection and face recognition.

CRF [12] is to capture the dependency between characters.CRF is used to transition features to relate the neighboring character labels with the features of the corresponding handwritten character segments along with features specific to a character image. Suen[1] mentions that the key to high recognition rates is feature extraction. However, this in itself is a very difficult problem which has led researchers to use more complex methods for preprocessing, feature extraction and classification. Such methods include the use of Neural Networks and Mathematical Morphology such as HMM. In some cases, researchers have constrained their experiments heavily, only using one person's handwriting, while other researchers' experiments were not performed on benchmark databases. Some of the problems and challenges which are faced by researchers today include: developing accurate segmentation, preprocessing, feature extraction and classification techniques. For the first problem, segmentation, the diverse styles and sizes of handwriting both play a large factor in the failure of current techniques. In some cases even a human being would not be able to segment handwriting containing characters which are tightly packed together and illegible. These segmentation systems also have to deal with the variability of handwriting from one person to another, not to mention problems when one writer's handwriting is cursive, while another person's is simply overlapping. Challenges faced for preprocessing deal with the choice of whether to convert raw handwriting into a more efficient form i.e. whether to binarise the handwriting or keep it in grey-scale form.

The segmentation of written words into components characters is often the first step in handwriting recognition systems. In some cases, segmentation is forced on the user by providing boxes for the writing of discrete letters. However in modern continues speech recognition efforts, segmentation of phonemes is not performed either before training or the recognition steps. Instead segmentation occurs simultaneously with recognition. If such a system is adapted for handwriting recognition, the very difficult issue of time consuming would be avoided. The method described in [2] uses combination of template and feature matching technique to recognize the pre segmented character shapes using neural networks. Neural Network approach has been also used for classification and recognition, since writing have a high input dimension and presence of large number of weights in the network. Next they used the sequence of angle features and single hidden layer network for classification, but the performance of these systems was also poor. Main reason for the poor performance the presence of noise in the x-y coordinates, which affected the extracted features greatly. In [3] the recognition system for handwritten manuscripts by writers of the 20th century. The proposed system first applies some preprocessing steps to remove background noise. Next the pages are segmented into individual text lines. After normalization a hidden Markov model based recognizer, supported by a language model, is applied to each text line. In [3] investigate two approaches for training the recognition system. The first approach consists in training the recognizer directly from scratch, while the second adapts it from a recognizer previously trained on a large general off-line handwriting database. The second approach is unconventional in the sense that the language of the texts used for training is different from that used for testing. In paper [3] experiments with several training sets of increasing size found that the overall best strategy is adapting the previously trained recognizer on a writer specific data set of medium size. The final word recognition accuracy obtained with this training strategy is about 80%.

Character recognition has been the subject of intensive research for more than thirty years. However, the problem of handwritten character recognition is far from being satisfactorily resolved despite many research efforts. The greatest difficulty lies in the infinite variations of shapes resulting from the writing habits and styles of the different writers. Stochastic modeling is a flexible and general method for modeling such problems in which dealing with uncertain information is necessary. Hidden Markov Model (HMM) has become the most promising model in speech recognition, mainly due to its ability to handle the sequential dynamic speech signal. Because of this success, it is naturally expected to achieve similar success in handwriting recognition. [4] Propose an efficient scheme for Chinese handwriting recognition in the framework of HMM. Automatic identification of handwritten script facilitates many important applications such as automatic transcription of multilingual documents and search for documents on the Web containing a particular script. The increase in usage of handheld devices which accept handwritten input has created a growing demand for algorithms that can efficiently analyze and retrieve handwritten data. [5] Paper proposes a method to classify words and lines in an online handwritten document into one of the six major scripts: Arabic, Cyrillic, Devnagari, Han, Hebrew, or Roman. The classification is based on 11 different spatial and temporal features extracted from the strokes of the words. The proposed system attains an overall classification accuracy of 87.1 percent at the word level with 5-fold cross validation on a data set containing 13,379 words. The classification accuracy improves to 95 percent as the number of words in the test sample is increased to five, and to 95.5 percent for complete text lines consisting of an average of seven words. It is well known that Hidden Markov Models (HMM) and Dynamic Programming (DP) [6], provide a theoretical framework and practical algorithms for temporal pattern recognition with lexical constraints (even for large vocabularies). The techniques initially developed for ASR are also applicable to Handwriting Recognition (HWR), which shares many features with ASR especially if auto segmentation (from word to letter) is used. Most on-line cursive handwriting recognition systems use a lexical constraint to help improve the recognition performance. In Paper [7], an integrated offline recognition system for unconstrained handwriting is presented. The proposed system consists of seven main modules: skew angle estimation and correction, printed handwritten text discrimination, line segmentation, slant removing, word segmentation, and character segmentation and recognition, stemming from the implementation of already existing algorithms as well as novel algorithms. This system has been tested on the NIST, IAM-DB, and GRUHD databases and has achieved accuracy that varies from 65.6% to 100% depending on the database and the experiment.
Most of Arabic handwriting recognition in previous works focused on recognizing offline script [8]. Much of online recognition focused on isolated Arabic letters only. As far as we could determine, there was littlework that tackled the difficulties of online Arabic cursive handwriting recognition. Al-Emami and Usher developed an online Arabic handwriting recognition system based on decision-tree techniques. The system was tested with 13 Arabic-letter shapes. Alimi developed an online writer dependent system to recognize Arabic cursive words based on neuro-fuzzy approach. The system was tested by one writer on 100 replications of a single word.

As for the delayed strokes, previously work viewed them as features that added complexity to online handwriting recognition. Four methods were proposed to recognize words with delayed strokes. In the first method, delayed strokes were totally discarded from handwriting in the preprocessing phase [8] In the second, delayed strokes were detected in the preprocessing phase and then used in a post processing phase [8]. In the third method, the end of a word was connected to the delayed strokes with a special connecting stroke. This special stroke, which indicated that the pen was raised, resulted in a continuous stroke sequence for the entire handwritten English sentence. Finally, delayed strokes were treated as special characters in the alphabet. So, a word with delayed strokes was given alternative spellings to accommodate different sequences where delayed strokes are drawn in different orders. The research of on-line handwriting recognition started in the 1960s and has been receiving intensive interest from the 1980s. The comprehensive survey before 1990s is made in [9]. As recent survey papers, Plamondon et. al., mainly reviewed the status of western on-line handwriting recognition while Liu et. al., and Jaeger et. al., reviewed that of on-line Chinese and Japanese handwriting recognition. Papers [9] mainly discuss on-line Japanese handwriting recognition.

Problem Statement

Handwriting of different person has different stroke, tilt and other patterns. Each writing word may comprise of different set of characters. Therefore handwriting recognition mainly depends upon extracting such features. Thus HMM is a suitable technique for detecting such features. The system is first normalized with respect to slant, skew, baseline location and height. A sliding window is used to transform a normalized handwritten text line into a sequence of feature vectors. The window is one pixel wide and shifted from left to right over a line of text. At each position of the window, five geometrical features are extracted and HTK is installed, HMM model is built, features extracted are passed into model. The standard HMM has proved to be very useful tool handwriting recognition although they present a poor discriminative power. On the contrary neural networks have been recognized as powerful tools for classification, but they are less efficient to model temporal variations than HMM. Identifying a person based on his handwriting samples is difficult because there are different features for writing such as slant, size, spacing between words, page margin, pressure, line spacing etc, which differ for every person. These are writing dependent and scanning quality dependent. So by extracting only these features handwritings cannot be recognized.

2. METHODOLOGY

The proposed work has been designed using MATLAB2014a and the design consists of five phases such as image acquisition, preprocessing, feature extraction, segmentation and classification as shown in Figure 1. The real time handwritten images are collected from different persons and their own handwritten work and stored in the database for further processing. The one of database image is subjected to preprocessing to enhance an image quality and to remove the noises. The Hidden Markow Model (HMM) is the best algorithm for extraction of features in terms of aging, the extracted features apply for classification to recognize the handwritten words and who's the handwritten one. The phase is discussed in details in the next section.

Preprocessing

The pre-processing is a series of operations performed on the scanned input image. It essentially enhances the image, rendering it suitable for segmentation. The binarization process converts a grayscale image into a binary image using a global threshholding technique.

There are two methods such as average and weighted methods, the average is the average of three colors as shown in the below

Grayscale = $(R + G + B/3)$ and weighted grayscale image = $((0.3 \times R) + (0.59 \times G) + (0.11 \times B))$.

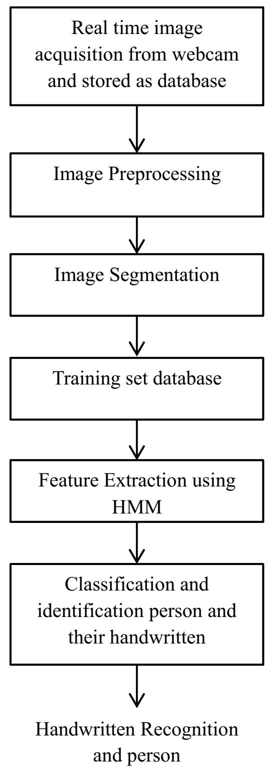

Figure 1: Block diagram of proposed handwritten recognition and person identification

Architectural Design

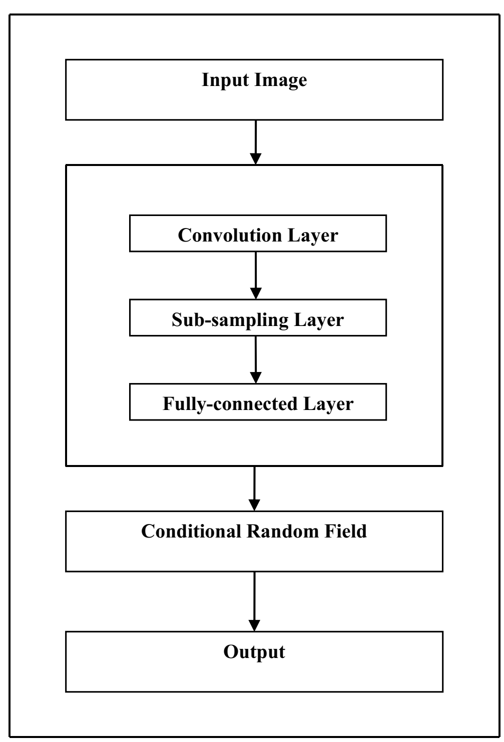

Figure 2: Internal architectural Design of handwritten recognition

The main steps performed by this function are:

1. Calculate a grid size based on the maximum dimension of the image. The minimum grid size is 32 pixels square.
2. If a window size is not specified chose the grid size as the default window size.
3. Identify grid points on the image, starting from top-left corner. Each grid point is separated by grid size pixels.
4. For each grid point calculate the cdf of the region around it, having area equal to window size and cantered at the grid point.
5. After calculating the mappings for each grid point, repeat steps 6 to 8 for each pixel in the input image.
6. For each pixel find the four closest neighbouring grid points that surround that pixel.
7. Using the intensity value of the pixel as an index, find its mapping at the four grid points based on their cdfs.
8. Interpolate among these values to get the mapping at the current pixel location. Map this intensity to the range [min:max) and put it in the output image.

Detection of edges in the binarized image using Sobel technique, dilation the image and filling the holes present in it are the operations performed in the last two stages to produce the pre-processed image suitable for segmentation. The pre-processing is a series of operations performed on scanned input image. Generally smoothing and normalization should be done in this step. The pre-processing also defines a compact representation of the pattern. The character is uniformly resized into pixels. After extracting the character we need to normalize the size of the characters. There are large variations in the sizes of each character hence we need a method to normalize the size. Then this result will be fed to the Feature extraction stage. In this stage it removes loops and difficult strokes and then it segments the word into separate characters. This result will be fed to the recognition phase. After feature extraction it compares the present character with the trained samples and recognizes the right characters and gives the output images. The proposed design part includes the system design, architectural design, and use case diagrams are discussed in the following sections.

The ANN consists of three main types of layers. They are Convolution layers, Sub-sampling layers and Fully-connected layers. Network layers are arranged in a feed-forward structure. Each Convolution layer is followed by a Sub-sampling layer and the last Convolution layer is followed by the Fully-connected layer. ANN extracts features from the raw image in the first layers and classifies the image in the last layer. In a Convolution layer each plane is connected to one or more feature maps of the preceding layer. A connection associated with a convolution mask is a 2-D matrix of adjustable entries called weights. The plane output is a 2-D matrix called a feature map. This name arises because each convolution output indicates the presence of a visual feature at a given pixel location. A convolution layer produces one or more feature maps. Each feature map is then connected to exactly one plane in the next Sub-sampling layer. A Sub-sampling layer has the same number of planes as the preceding convolution layer. A Sub-sampling plane divides its 2-D input into non-overlapping blocks of size 2×2 pixels. In the last convolution layer, each plane is connected to exactly one preceding feature map. This layer uses convolution masks that have exactly the same size as its input feature maps. The result will be fed to conditional random field and it produces the output. An undirected graphical method is learned using the outs of the hierarchical feed-forward layers which serve as feature extractors. This architecture automatically learns powerful nonlinear features from the raw input data. The approach proposed here is superior to the other NN/CRF combined models, because of its easy implementation, deep architecture and robustness to image variance.

Sequence Diagram

A Sequence diagram is an interaction diagram that shows how processes operate with one another and what is their order. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario.

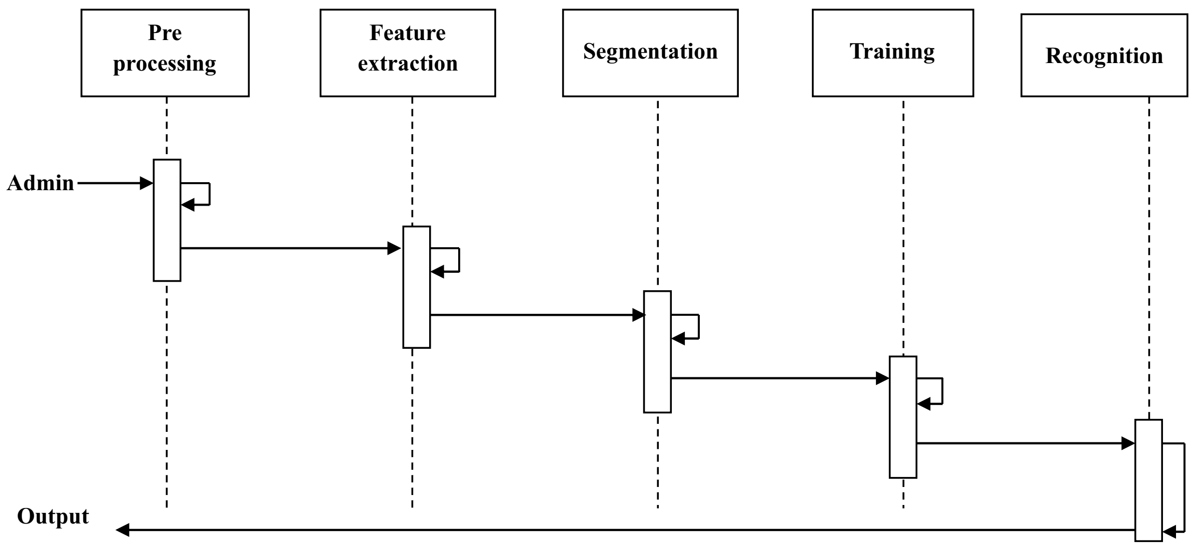

Figure 3: Handwritten Sequence Diagram

Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios. A sequence diagram shows, as parallel vertical lines, different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. The Figure 3 shows the Sequence Diagram for the proposed system.

Control Flow Design of Handwritten

The control flow is constructed by using different types of control flow elements and it includes the task and containers that connect to data source. Here the activity diagram shows the control flow of the proposed system.

Activity Diagram

Activity diagram are graphical representation of workflows of stepwise activities and actions. Activity diagrams are constructed from a limited number of shapes, connected with arrows. A black circle represents the start of the workflow, rounded rectangles represent actions and diamonds represent decisions. Arrows run from the start towards the end and represent the order in which activities happen. An encircled black circle represents the end. The Figure 4 shows the activity diagram of the proposed system.

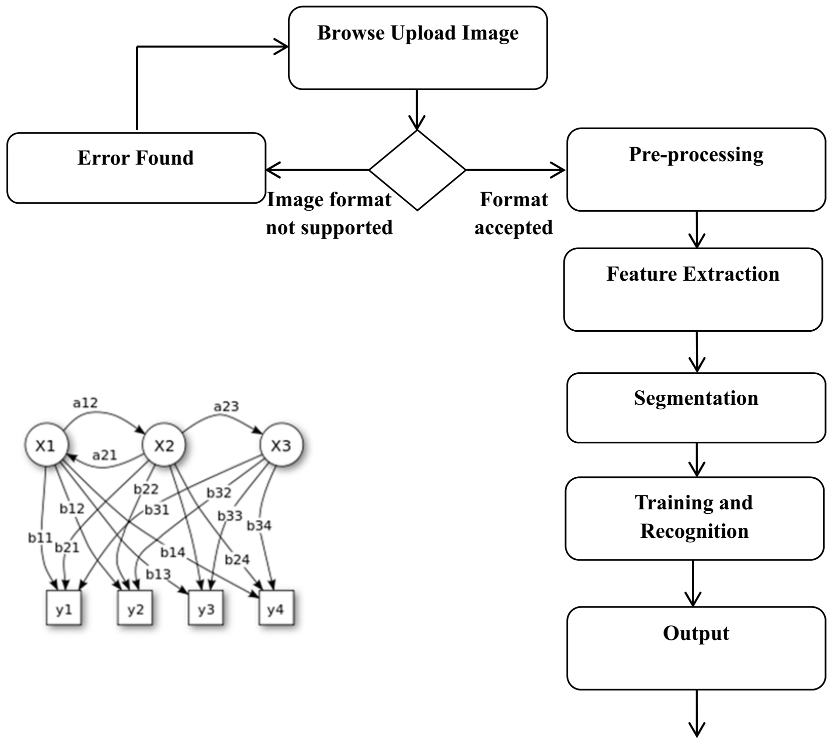

Figure 4: Flow activity Diagram of handwritten recognition

Feature Extraction using HMM

The Segmentation of content into sections through explicit or implicit, split the transform the content for size normalization. Detection of major features from trained data (analytics engine) using top-down approach as follows:

1. In feature extraction stage each character is represented as a feature vector, which becomes its identity. The major goal of feature extraction is to extract a set of features, which maximizes the recognition rate with the least amount of elements.
2. Due to the nature of handwriting with its high degree of variability and imprecision obtaining these features, is a difficult task. Feature extraction methods are based on 3 types of features:
   (i) Statistical
   (ii) Structural
   (iii) Global transformations and moments

Normally feature selection algorithms use an objective function that assigns a measure of quality to each feature set. Often this function calculates the recognition rate of the classifier using the considered feature set on a validation set. For many applications the time complexity of a feature selection algorithm is very high, because the validation set must be large in order to get reliable results and because the number of validated feature sets is large. In the current paper we propose a new objective function that quickly computes an approximation of the recognition rate on a validation set. An HMM based classifier system usually incorporates one individual HMM for each class where each HMM is built up from a set of states. For all these states output distributions and for all pairs of states transition probabilities are defined. The input to an HMM classifier is a sequence of feature vectors $x_1, ..., x_n$. In the recognition, or decoding, phase the input sequence $x_1, ..., x_n$ is mapped to a sequence of states $si_1, ..., si_n$ and for each such mapping a likelihood value is defined by the HMM. The optimal mapping, which maximizes the likelihood, is usually found by means of the Viterbi algorithm. This optimal mapping is equivalent to an optimal path through a graph that is defined by the product of the sequence of input vectors and the states of the HMM. The likelihood of the optimal path is the score of the sequence of feature vectors for the considered HMM. The class corresponding to the HMM with the highest score is the output class as shown in the Figure 5-8.

Classification using Neural Network

From the collected features from the HMM are feed to ANN for training and testing and within 1000 iterations, the ANN will predict the person is recognized or not and their handwritten words.

Training the Neural Network

The construction of the neural network involves three different layers with feed forward architecture because it is the most popular network architecture using still today. The input layer of this network is a set of input units, which accept the elements of input feature vectors. The input units (neurons) are fully connected to the hidden layer with the hidden units. The hidden units (neurons) are also fully connected to the output layer. The output layer produced the response of neural network to the activation pattern applied to the input layer. The information given to a neural net is propagated layer-by-layer from input layer to output layer through one or more hidden layers for percentage of recognition and handwritten person as shown in Figure 9-11.

3. RESULTS AND DISCUSSION

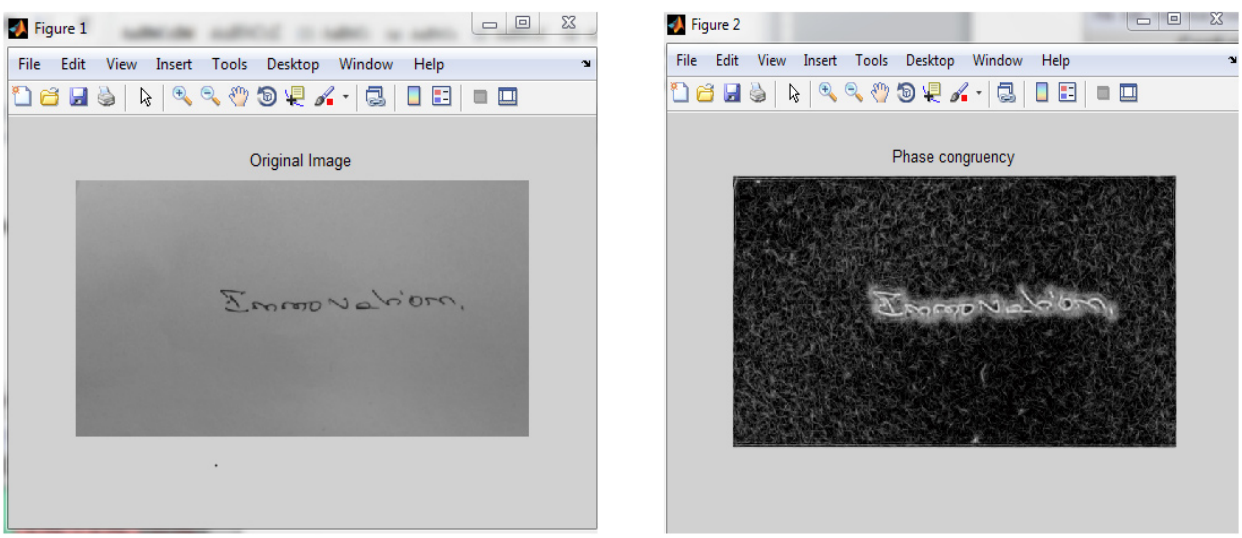

Figure 5: Original handwritten image and its phase congruency image output

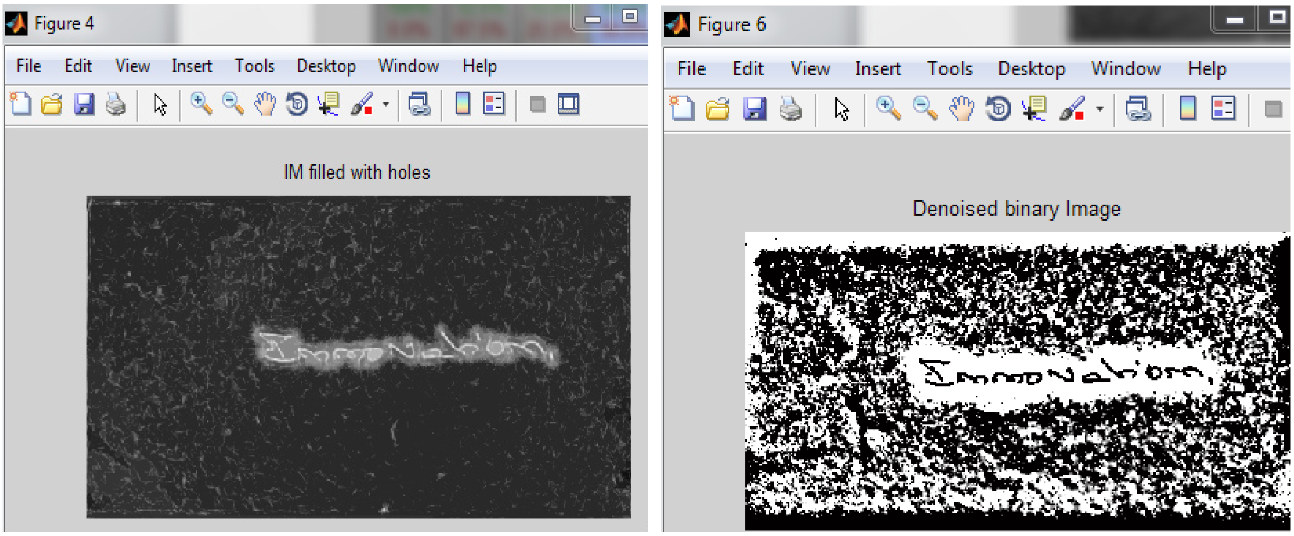

Figure 6: Handwritten image of IM with holes and its denoised binary image

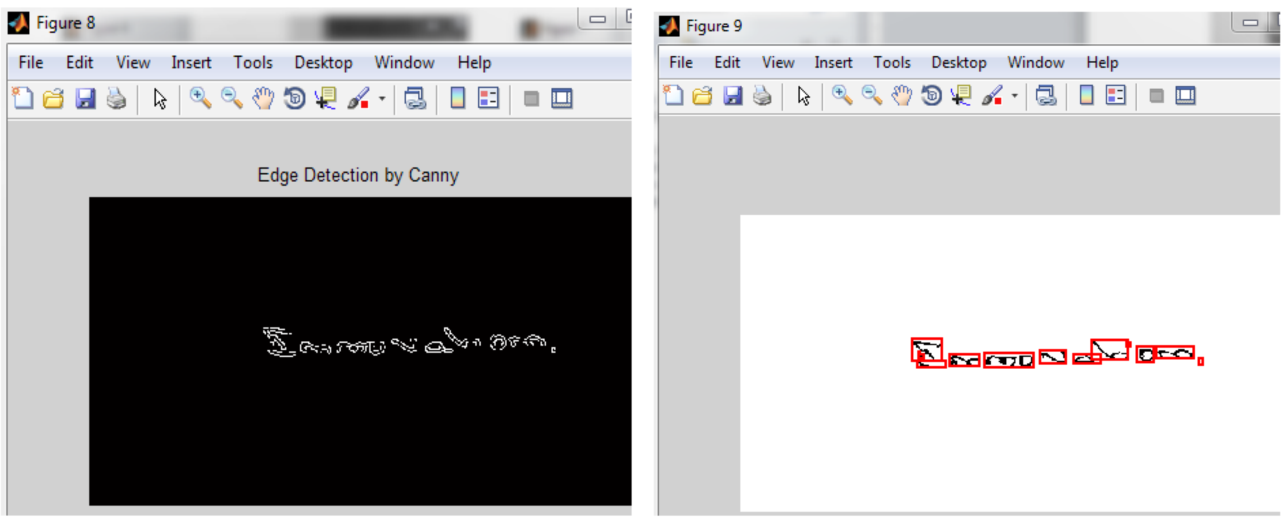

Figure 7. Edge detection of handwritten image and its each character segmentation

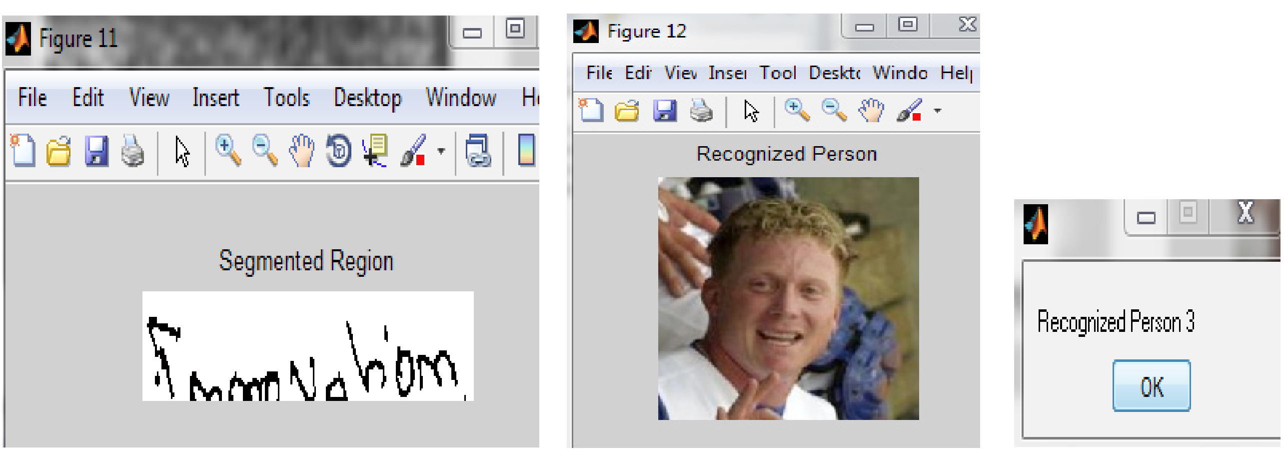

Figure 8: Segmentation region of handwritten recognition and identification of who handwritten

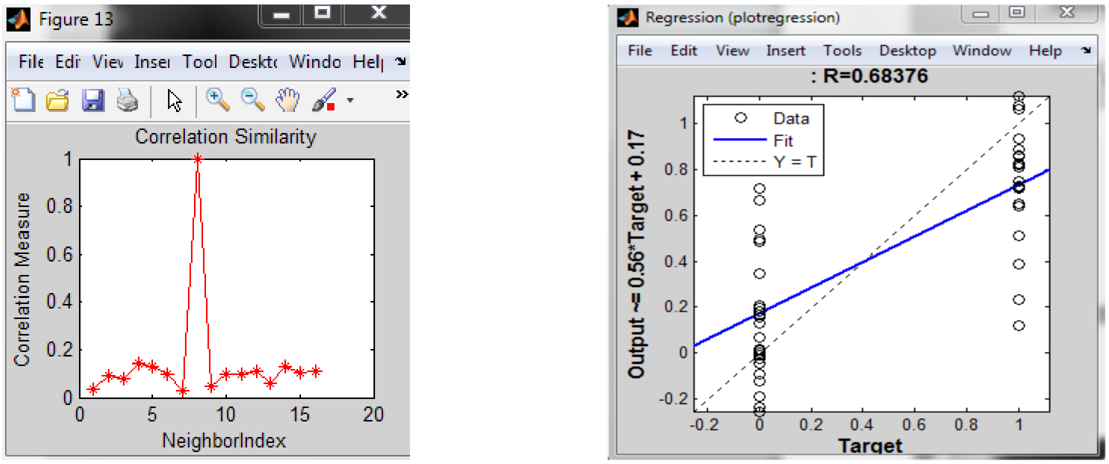

Figure 9: Performance analysis plotted graphs for handwritten recognition based on the similarity

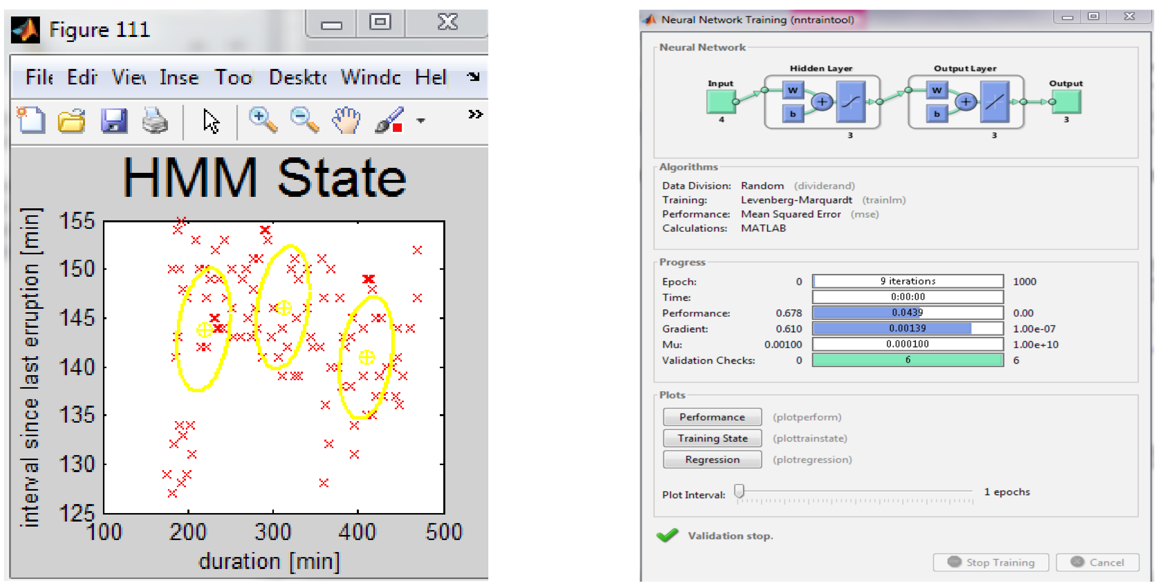

Figure 10: Feature extraction using HMM and its classification using ANN

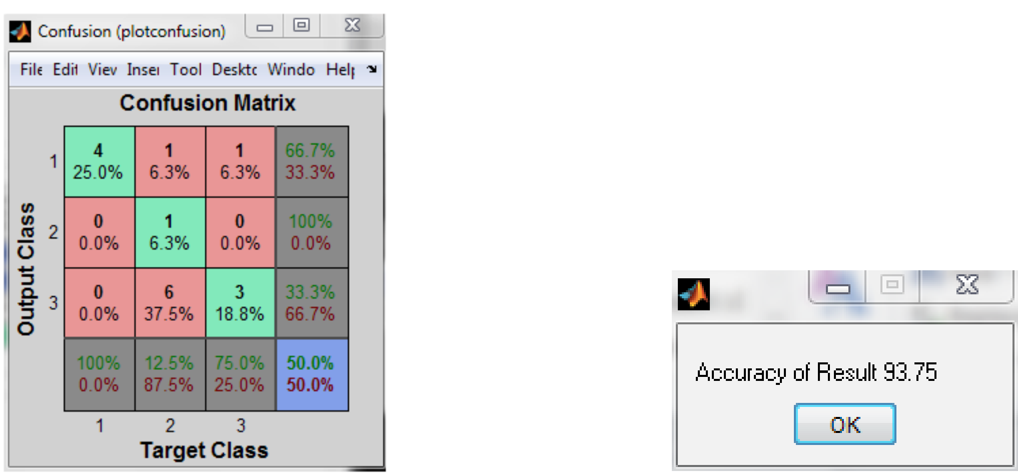

Figure 11: Confusion matrix of ANN and accuracy of handwritten recognition

4. CONCLUSION

Extraction of Features is a crucial component in the real application pertaining to handwriting recognition systems.

Using HMM and ANN architectures the real time written signature are maintained in the database for further process and The results on Arabic and Latin word databases show that the Marti–Bunke and the LHG feature sets are the most efficient, with good complementarity with each other. In future work, this framework will be applied to guide the design of novel features, and it will be extended to compare the nature of various recognition methods in terms of strength and complementarity, HMM with ANN models.

REFERENCES

[1] B. Verma, M. Blumenstein & S. kulkarni-" On-Line and Off-Line Handwriting Recognition:A Comprehensive Survey", 1EEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. Vol. 22, No. 1. JANUARY 2000

[2] Sujata S.K and Kohir "*Handwriting recognition using HMM"* ICN 2010.

[3] Emanuel Inderm uhle1 and Marcus Liwicki and Horst Bunke," *Recognition of Handwritten Historical Documents: HMM-Adaptation vs. Writer Specific Training"* Website: http://www.nb.admin.ch/slb/.

[4] Bing Feng, Xiaoqing Ding, Youshou Wu," *Chinese Handwriting Recognition Using Hidden Markov Models"* 1051-4651/02 $17.00 (c) 2002 IEEE.

[5] Anoop M. Namboodiri, and Anil K. Jain-" *Online Handwritten Script Recognition",* IEEE transactions on pattern analysis and machine intelligence, Vol. 26, No. 1, January 2004.

[6] Alain Lifchitz "*A fast lexically constrained viterbi algorithm for online handwriting recognition*" In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition,September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 313-322.

[7] E. Kavallieratou, N. Fakotakis, G. Kokkinakis-"*An unconstrained handwriting recognition system"* IJDAR (2002) 4: 226–242.

[8] Fadi Biadsy, Jihad El-Sana, Nizar Habash-"*Online Arabic Handwriting Recognition Using Hidden Markov Models"* www.alarabimag.com,www.aljazeera.net.

[9] Masaki Nakagawa Junko Tokuno Bilan Zhu Hideto Oda Akihito Kitadai-" *Recent results of On-line Japanese handwriting recognition and its Applications*".

[10] L. R..Rabiner and B. H. Juang," *An Introduction to Hidden Markov Models",* January 1986 IEEE ASSP Magazine.

[11] Charles C. Tappert and Sung-Hyuk Cha,"*English Language Handwriting Recognition Interfaces*"- Chapter 6 in *Text Entry Systems*, ed. MacKenzie and Tanaka-Ishii, MorganKaufman, 2007.

[12] Christophe Choisy and Abdel Belaid-**"***Handwriting Recognition Using Local Methods for Normalization and Global Methods for Recognition"-ICDAR.*

[13] A. L. Koerich, R. Sabourin, C. Y. Suen-"*Large vocabulary off-line handwriting recognition:A survey"*- Pattern Anal Applic (2003) 6: 97–121 DOI 10.1007/s10044-002- 0169-3.

[14] Z.Hachkar, A. Farchi, B.Mounir, J. EL Abbadi " *Building HMM recognizer for Arabic Digits using HTK".*

[15] Lawrence R Rabiner-"A Tutorial on Hidden Markov Model and selected applications in speech recognition"-Proceedings of IEEE, Vol. 22, No.2, February 1989.

[16] Haikal El Abed, Volker M¨argner, Monji Kherallah, Adel M. Alimi," *ICDAR 2009 Online Arabic Handwriting Recognition Competition*" 2009 10th International Conference on Document Analysis and Recognition.

[17] S.Mori, C.Y. Suen and K. Kamamoto, "historical review of OCR research development." Proc of IEEE.

[18] R. Plamondon and S.N Srihari, ”on-line and off-line handwritten character recognition: A comprehensive survey,”IEEE.

[19] J. Lafferty, A. Macullum, and F. Pereira,. Conditional random fields: Probabilistic models for segmenting and labeling sequential data,. *Eighteenth international conference on Machine Learning.*

[20] Lecun, Y. and Bottou, L. and Bengio, Y. and Haffner, P., “Gradient-based learning applied to document recognition”, Proceedings of the IEEE.