AI techniques in Handwriting Recognition

Sanam Narula
Department of Computer Science
University of Waterloo
Waterloo, Ontario
Email: s5narula@uwaterloo.ca

***Abstract*—In this paper, we have discussed various AI techniques used in handwriting recognition. The main aim of the paper is to explain the process of handwriting recognition and to discuss various techniques used for it. We will explain the components of a handwriting recognition system. This work will mainly focus on offline handwriting recognition, however, we will give some insight into the working of on-line recognition systems. In the end, we will compare various techniques used in the process of handwriting recognition.**

***Keywords*—Handwriting Recognition, Offline handwriting recognition, segmentation**

I. Introduction

Handwriting recognition is the process of conversion of handwritten text into machine-readable form. A major hurdle in handwriting recognition is the variation of the handwriting styles, which can be completely different for different writers [9].

Handwriting is a skill that is personal to an individual. It is also used as an accepted biometric for identity verification. Fingerprints and iris verification require the installation of costly equipment. This makes it difficult to use these on a day to day processing places such as banks [4]. Although, the handwriting recognition may not be as robust as fingerprint or iris recognition but still it serves the purpose in a myriad of fields. Signature is found to be the foremost authentic parameter within the field of authentication.

Handwriting interpretation is the task of determining the meaning of a body of handwriting, e.g., a handwritten address. Handwriting identification is the task of determining the author of a sample of handwriting from a set of writers. These two processes determine the special nature of the writing, while handwriting recognition and interpretation are processes whose objectives are to filter out the variations so as to determine the message [10]. It removes the noise from input in order to make it more refined.

Artificial Intelligence techniques are used in Classification and Recognition component of handwriting recognition. Many approaches have been proposed and implemented by researchers such as ANN, HMM, etc.. These techniques have different accuracy rates in different scenarios. We will discuss these techniques in more detail in section 4. The field of offline handwriting recognition has advanced greatly in the past decade and is still a progressing field.

In section 2, we will explain the motivation for research in the field of handwriting recognition. In section 3 and section 4, we will explain different types of handwriting recognition and different phases which a document go through during the process of handwriting recognition. We will explain artificial intelligence techniques used for handwriting recognition. In section 5, we will provide a comparison of various techniques explained in section 4. In section 6, we will provide a conclusion of our work.

II. Motivation

Handwriting recognition is finding its application in various domains. The use of artificial intelligence in the field of handwriting recognition is making the recognition process more accurate and exciting. The increase in the interest of several researcher and companies has led to widespread use of handwriting recognition software. The applications in the domain of handwriting recognition are:

- Recognition of the age of a person by using samples of his/her handwriting: In various forensic investigations, it is helpful to classify the age of the person [2]. In [2], authors used a SVM classifier and achieved an accuracy of 70 percent in the age group determination.
- Recognition of gender (man or woman) by handwriting: In [8], the authors used the effect of Euler number as a feature and used Back Propagation Neural Network(BPNN) to determine the gender of the writer. They reported an accuracy of 80 percent while using their method for determining the gender.
- Writer Verification: It is the task of determining if two handwriting samples were written by the same writers. This is really important especially in the fields like Questioned Document Examination [15].
- Signature Verification: One of the most widely used applications of handwriting recognition is signature verification. Used in banks and other financial institutions. The use of handwriting recognition software in these institutions is apt because it is not possible to have a handwriting expert available at every place. The software which mainly uses artificial intelligence techniques has a fair accuracy level which helps authorities to fight against forgery.
- Bank-Check Processing: Handwriting recognition techniques are used in banks for not only signature verification but also for reading courtesy amounts from checks [10].
- Automatic conversion of handwritten documents to digital form: Information acquisition from customers

is done mainly using documents. These documents or forms are normally handwritten. These documents are converted into digital form in order to ease the information retrieval and storage.

III. Types of Handwriting Recognition

Handwriting data is converted into digital form either by scanning the document or by writing with a special pen on an electronic surface. Based on different approaches, handwriting recognition can be divided mainly into 2 categories:

A. Online Handwriting Recognition

Online handwriting verification tracks down the path and other time-varying sequence variables using specially designed tablets or other devices. In Online recognition system, the 2D coordinate of successive points of writing are stored as a function of time. It examines the behavioral components of the handwriting such as stroke order, speed, and pressure. It measures the physical activity of writing. The recognition rates reported are much higher for the online case in comparison with the offline case.

Figure 1 shows a typical setup of Online handwriting recognition system in which a tablet is used. The user has to write on the tablet using a pen. In applications where the field is of writer recognition, the tablet senses the pressure, speed, and other handwriting features in order to match a person and verify his signature.

Fig. 1: On-line handwriting recognition setup.

B. Offline Handwriting recognition

Offline handwriting recognition is the automatic conversion of text into an image. The data obtained by this form is regarded as a static representation of handwriting. A Camera or a scanner is used for image acquisition. In case of offline handwriting recognition, document analysis is the necessary preliminary step that locates appropriate text when complex, two-dimensional spatial layouts are employed [11].

IV. Phases of Handwriting Recognition

Handwriting recognition is a series of different phases. Each phase refines the input data and tries to convert it into a form which is best for classification and recognition phase. Various phases in handwriting recognition process are shown in Figure 3.

A. Image Acquisition

The input image is taken through a camera or some scanner. The image should have a specific format such as JPEG; BMT etc. The input captured may be in gray, color or binary form.

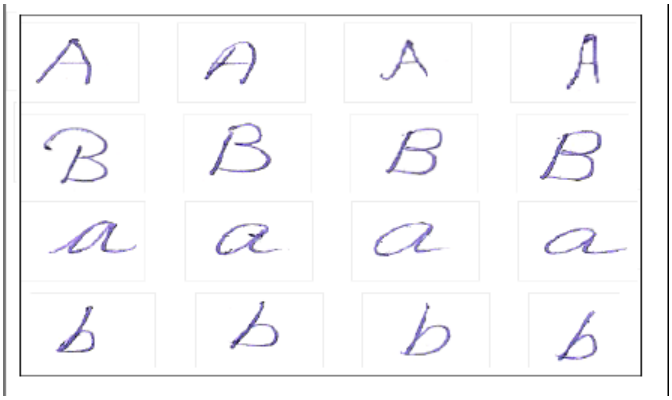

Fig. 2: Image Acquisition

Figure 2 shows a sample data set which is acquired in the system. This helps in recognizing the text written by a specific writer. In cases, where the handwriting recognizer software are installed for text recognition such as reading courtesy amounts from the check, the image of written samples is provided to the system.

B. Preprocessing

The task of preprocessing relates to the removal of noise and variation in handwritten word patterns. It compensates for some of the natural variation that arises between writers. It is necessary to perform several preprocessing operations prior to recognizing text in scanned documents [10]. It essentially enhances the image quality, making it suitable for further processing. Some of the common operations performed are:

- Noise Removal: Noise removal is a topic in document analysis that has been dealt extensively for typed or machine-printed documents. For handwritten documents, the connectivity of strokes has to be preserved [12]. Improper scanning of document or malfunction of the camera or low quality of documents can add noise.

  Various Smoothing filters can be applied to remove the noise from the image. Smoothing is used for blurring and reducing noise, and removal of small details from the image extracting large objects [9]. Morphological operations are also used as a tool in image processing for extracting components of the image that are useful. Morphological operations can also be used to remove

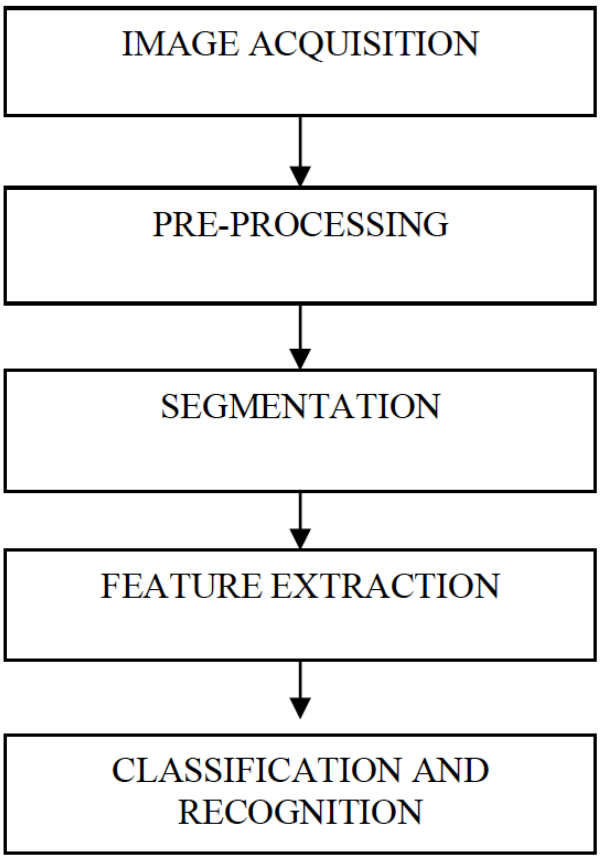

Fig. 3: Phases in handwriting Recognition [13]

the noise on the documents due to the low quality of paper and ink.

- Binarization and Thresholding: It is a crucial step in handwriting recognition. Binarization process converts a grayscale image into a binary image. Good binarization facilitates segmentation and recognition of characters [13].

  Thresholding is used in binarization of an image. It replaces an image pixel with a white pixel or a black pixel depending on the threshold. It helps in extracting the foreground data from background data. It is often used to increase the processing speed. There are two types of thresholding.

  Global thresholding picks one threshold value for the entire document image. It is based on an estimation of the background level from the intensity histogram of the image. Adaptive thresholding is used for images in which different regions of the image may require different threshold values [1].

- Skew Correction: The improper scanning angle of the document leads to skew problems. It is not necessary that the documents are aligned horizontally while scanning. This condition leads to the skewness of document which should be removed before segmentation process.

  There are various methods that are used for skew correction. Some methods rely on detecting connected components and finding the average angles connecting their centroids. The methods which are majorly used are projection profile analysis, Hough transforms, nearest neighbor clustering, cross-correlation, piecewise covering by parallelogram etc [9].

  Figure 4 shows the effect on a text after skew correction has been applied on it.

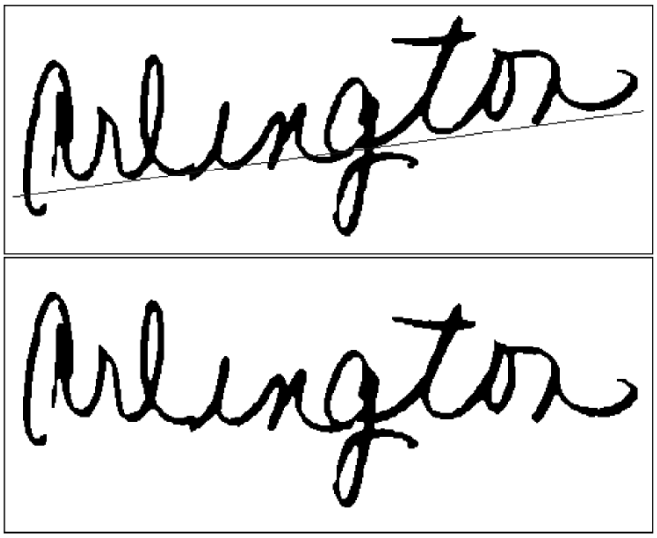

Fig. 4: Skew Correction [14]

- Slant Estimation and Normalization: Cursively written text can exhibit significant slant. This is corrected using slant correction methods. Slant correction does not affect the connectivity of the word and the resulting words are natural. Slant estimation is commonly accomplished by calculating the average angle of near vertical elements.

  Normalization is applied in order to remove the deformities from the document which get introduced due to the variations of different writers. The goal of normalization is to remove or reduce these instances of variability. Some of the common normalization are applied on variability introduced due to baseline drift, writing slant and variable size of the writing in the document [12].

There are many other preprocessing filters which can be applied to the input in order to make it more suitable or more refined for next phases. Once the preprocessing is complete, the next phase is of segmentation.

C. Segmentation

Segmentation is the process of breaking the input data into smaller logical units. Segmentation of unconstrained handwritten word into different zones and characters is more difficult

as compared to that of printed documents. This is mainly because of variability in the inter-character distance, skew, slant, and size. There could be a case where components of two consecutive characters may be touching or overlapping each other. This situation complicates the segmentation task [13].

Segmentation is mainly categorized into 3 categories:

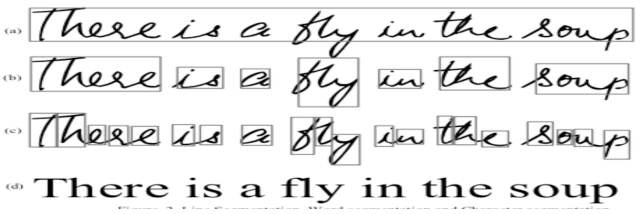

Fig. 5: Segmentation [10]

1) Line Segmentation: It decomposes the document into logical units. It is the most critical part of the document analysis. Figure 5(a) shows the result of a line segmentation. Segmentation of handwritten document is much more difficult than a machine-printed document. In the case of a machine-printed document, the segmentation can be easily accomplished by examining the horizontal histogram profile at a small range of skew angles. The task is more difficult in the handwritten domain.

   In the case of a handwritten document, lines of text might be undulate up and down and ascenders and descenders frequently intersect characters of neighboring lines. One method is based on the notion that people write on an imaginary line which forms the core upon which each word of the line resides. The local minima points approximate this imaginary baseline from each component. A clustering technique is used to group the minima of all the components to identify the different handwritten lines [11].

2) Word Segmentation: Line separation is usually followed by a procedure that separates the line into words. Most of the approaches present for word segmentation focus on identifying physical gaps using only the components. These methods assume that gaps between words are larger than the gaps that are present in between characters. Figure 5(b) shows the result of word segmentation.

3) Character Segmentation: It is basically the process of decomposing an image of a sequence of characters into sub-images of individual symbols. The word is fragmented into sub-letters. In the case of handwritten documents, there could be exceptions because of writing styles with leading and trailing ligatures. These exceptions make the process of character segmentation very challenging. One of the methods used for segmentation incorporates cues that humans use and does not rely solely on the one-dimensional distance between the components.

   We can see in the case of figure 5, the writing style of the writer is characterized by the variation of spacing between adjacent characters as a function of the corresponding characters themselves. The notion of expecting greater space between characters with leading and trailing ligatures is enclosed into the segmentation scheme [10].

D. Feature Extraction

It is the process to retrieve the most important information from the raw data. Features should be extracted which maximizes the recognition rate with the least amount of elements. These features are essential for classification stage as these features help in recognizing the specific character.

This is an important stage as its successful operation improves the recognition rate and reduces the misclassification [9]. Each character is represented as a feature vector, which becomes its identity. The widely used feature extraction methods are Template matching, Deformable templates, Contour Profiles, Zoning, Geometric moment invariants etc..

Features are divided into different categories:

1) Statistical features: Statistical features are derived from the statistical distribution of points. These features provide high speed. Writing style variation is also handled by these features [16]. Some of the major statistical features are:

- Zoning: Zoning divides the image into NxM zones. For each zone, features are extracted to form a vector. The goal of zoning is to obtain the local characteristics instead of global characteristics. Figure 6 shows the result of applying zoning feature on an image.

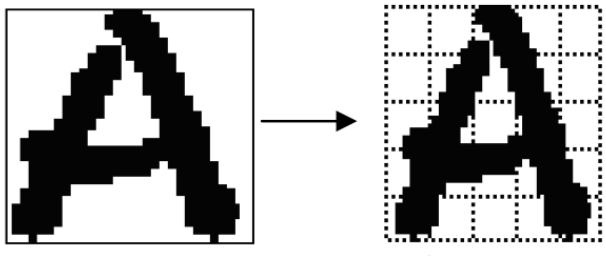

Fig. 6: Zoning [13]

- Crossing and Distances: Crossing count the number of transitions from background to foreground pixels along vertical and horizontal lines through the character image.

  Distances calculate the distances of first image pixel detected from the upper and lower boundaries of the image along vertical lines and from the left and right boundaries along horizontal lines.

- Characteristics Loci: In this feature, for every white point in the background of character vertical and horizontal vectors are generated. The number of times the lines intersect with these vectors is used as a feature [1].

2) Structural feature: Structural features give information about the structure of the image. Structural features describe the geometrical and topological properties of a character, like crossing points, branches, loops, stroke length, stroke width, up, down, left and right projection profiles etc.

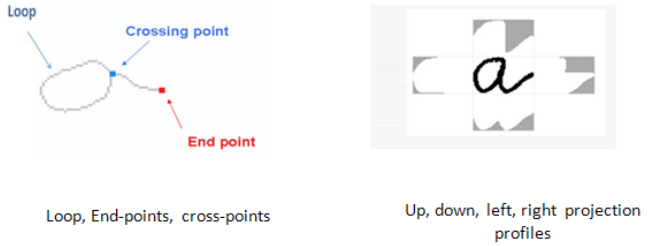

Fig. 7: Structural Features [9]

3) Global transformation feature: Global transformation based features give a nice representation of an image. It is the translation of an image from the spatial domain to frequency domain. This feature stores information contained in an image in the form of few coefficients, thus it performs energy compactness.

Various types of global transformation based features are Discrete Fourier Transform, Discrete Cosine Transform, Discrete Wavelet Transform etc.

E. Classification and Recognition

This is the decision-making part of a recognition system. It uses the features extracted from the previous stage as an input. The techniques of classification are broadly categorized into 2 types:

1) Supervised Learning: In supervised learning, the training data with correct class is used to train the model. This model is used to test the data for correct classification. The training data includes both the input and the desired results. The model undergoes learning process and based on this learning it classifies the test data. Some examples of supervised learning are SVM, HMM, Template based matching etc.

2) Unsupervised Learning: In unsupervised learning, no training data is provided. The model classifies test data based on statistical properties and by their spatial grouping and considering their nearest neighbor. Some examples of unsupervised learning are Clustering, k means etc [9].

The techniques which are majorly used for classification and recognition are:

1) Template Matching: Template matching involves determining similarities between a given template and windows of the same size in an image and identifying the window that produces the highest similarity measure. Once a number of corresponding templates are found their centers are used as corresponding points to determine the registration parameters. Template matching is a simple and robust approach but its efficiency is affected if the image is distorted.

Template matching can be studied in two types:

- Deformable Templates and Elastic Matching: This is used in the domain of object recognition. An image deformation is used to match an unknown image against a database of known images. Two characters are matched by deforming the shape of one, to fit the edge power of the other [3].

  The idea of elastic matching is to match the unknown symbol against all possible elastic stretching and compression of each prototype in the most optimal way. A dissimilarity measure is derived from the amount of bend needed, the decency of fit of the edges and the interior overlap between the distorted shapes [9].

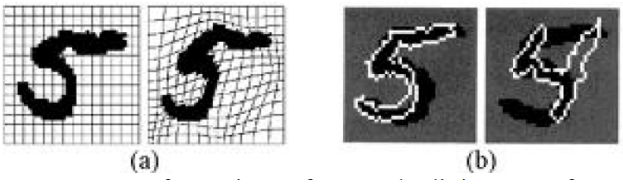

Fig. 8: (a): Deformations of a sample digit, (b) Deformed Template superimposed on target image, with dissimilarity measures [3]

- Direct Matching: A gray-level or binary input character is directly compared to a standard set of stored prototypes. The matching techniques can be as simple as a one-to-one comparison or as complex as decision tree analysis in which only selected pixels are tested.

  Various kind of similarity measures can be used such as Euclidean, Mahalanobis, Jaccard or Yule similarity measures etc. The direct matching method is intuitive but the recognition quality can be very easily affected by noise.

2) Hidden Markov Model: Markov models are used for both online and offline handwriting recognition. Hidden Markov models effectively capture correlations in the sequence of characters in each word.

A hidden Markov model is essentially a Markov chain with one significant difference. In a Markov chain, a state is associated with a unique, deterministic output value, whereas in the case of an HMM, each state is associated with a probability distribution over all the possible output values [20].

Usually, the HMM model can be represented as:

$$\lambda = (A, B, \pi)$$

where A = the state transition matrix.
B = the observation probability matrix.
π = the initial probability distribution for the states.

The approach is especially attractive for handwriting recognition where reliable character segmentation is extremely hard to achieve. In HMM modeling, characters are represented as a succession of states with left-right transitions and a self-transition. The HMM are used with language models in order to achieve greater efficiency.

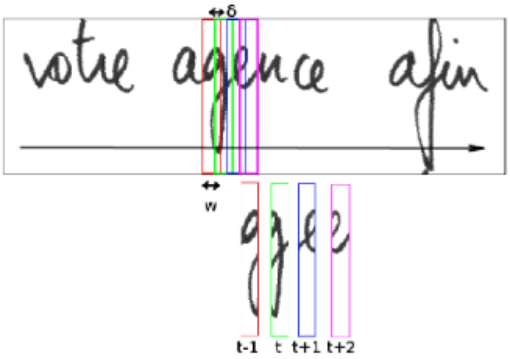

Fig. 9: Extraction of features using sliding window

In order to make HMM technique applicable to images of printed or handwritten text, the images must be transformed into a suitable sequence representation. This is achieved by a sliding window along the text line in the direction of writing. Features are extracted at each position from the small patch of the text image covered by the sliding window [5]. In Figure 9 we can see how the text is iterated and features are extracted using a sliding window.

3) K-Nearest Neighbor (KNN): K- Nearest neighbor is a class of statistical classifiers. Statistical classifiers are rooted in the Bayes decision rule.

In KNN, an incoming pattern is classified using the cluster, whose center is the minimum distance from the pattern over all the clusters. The training vector is usually a multidimensional array. Each row in an array contains feature values and corresponding class label of the training images.

Test vector assigns the class label based on the Euclidean distance measures and number of neighbors (k) considering it does not involve a prior information about the data. This is also called the concept of majority voting of neighbors, an object is classified with being assigned to the class most common amongst its k nearest neighbors, where k is a positive integer. If k = 1, then the object is simply assigned to the class of its nearest neighbor.

As we can see in Figure 10. The test sample (red object) should be classified either to the first class(blue squares)or to the second class (green circles). If k = 3 it is assigned to the first class because there are 2 square and only 1 circle inside the inner circle.

If k = 7, it is assigned also to the first class.

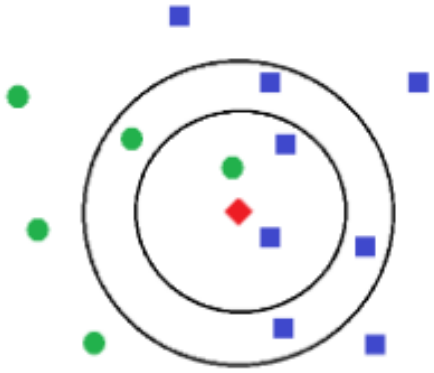

Fig. 10: An example of k-NN classification [6].

4) Artificial Neural Networks(ANN): The term Artificial Neural Networks (ANN) refers to a class of algorithms for which the design of the computational architecture and the process involved is inspired by knowledge of information processing in the human brain. The information is processed by a large amount of simple processing elements which exchange information in the form of activation levels.

An activation level is usually a real-valued parameter, reflecting the state of a neuron or unit, similar to the physiological firing rate of a biological neuron measured at a given time instance [12]. Different classes of Artificial Neural Networks are widely used in the field of handwriting recognition. ANN are fast and reliable tools for classification.

ANN are classified into two major sets:

1) Feed-forward: Nodes are organized into layers. The input layer of nodes, one or more hidden layers, and an output layer. Back propagation is a learning rule for the training of multilayer feed-forward neural network. To train a Back propagation neural network, it must be exposed to a training data set and the answers or correct interpretations of the set.

2) Feedback (recurrent) networks: A Feedback network or a recurrent neural network (RNN) is a class of artificial neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit dynamic temporal behavior.

   Unlike feed-forward neural networks, RNNs can use their internal memory to process arbitrary sequences of inputs. This makes them applicable to tasks such as unsegmented connected handwriting recognition [19]

The majority common ANN used in the handwriting recognition systems are the Multilayer Perceptron (MLP) of the feed forward networks. Figure 14 shows a Feed-forward neural network which consists of 3 layers of input, output and hidden.

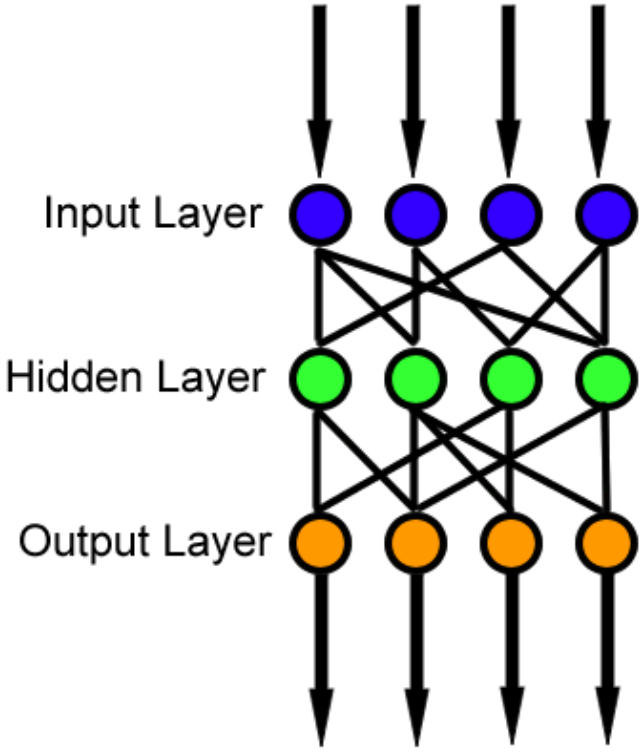

Fig. 11: Feed-Forward Neural network [18].

Multilayer Perceptron (MLP) is a class of networks which consists of multiple layers of computational units interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. The training of such networks is supervised, meaning that the classes of shapes are well defined.

The input features are mapped onto the input units of the MLP, and the output units of the network are defined to represent a particular class of characters. A training set is presented several times to the network until a satisfying error between target and obtained output activations has been achieved [12].

The classification performance of the MLP is tested against a new and unseen test set.

V. Comparitive Analysis

Although, we explained few major techniques used for handwriting recognition, there are much more. These techniques are employed depending on the application domain and data.

Figure 15 shows a comparative analysis between various techniques that we have discussed in the paper. Every technique has some advantage and disadvantage. It really depends on the domain as to what should we choose.

Approaches	Advantages	Disadvantages
Template Matching	Simple and Robust	It can detect only unskilled forgery.
Statistical Approach(K-NN)	It is widely used and can detect skilled forgery as well.	It is expensive to implement.
Hidden Markov Model	It is easy to implement which results in quick verification.	It is also expensive because of the learning requirements,
Neural Network Approach	Fast and reliable. It gives very high results.	It cannot be retained if new data is added.

Fig. 12: Comparative Analysis [18].

VI. Conclusion

Handwriting recognition is a part of a large domain which has applications in graphology and forensic science. Handwriting recognition is a challenging task as same user's handwriting tends to differ depending upon different scenarios. Handwriting is a pattern associated with a user. In this work, we tried to explain the process of handwriting recognition. We listed some famous AI techniques used in the domain of handwriting recognition. We provided a comparative analysis of these techniques.

Handwriting recognition is a field in which research has been going on for many years but still the field is young and the researchers are trying to come up with new methods and techniques which can increase the accuracy and applicability of handwriting recognition.

References

[1] Nariz Arica An Offline Character Recognition System for Free Style Handwritting 1998.

[2] N. Bouadjenek, H. Nemmour and Y. Chibani, "Age, gender and handedness prediction from handwriting using gradient features," Document Analysis and Recognition (ICDAR), 2015 13th International Conference on, Tunis, 2015, pp. 1116-1120. doi: 10.1109/ICDAR.2015.7333934

[3] Mohamed Cheriet, Nawwaf Kharma, Cheng-Lin Liu, Ching Y. Suen, Character Recognition Systems: A Guide for students and Practitioners, (John Wiley and Sons, Inc., Hoboken, New Jersey, 2007).

[4] Maged M.M. Fahmy, Online handwritten signature verification system based on DWT features extraction and neural network classification, Ain Shams Engineering Journal, Volume 1, Issue 1, September 2010, Pages 59-70, ISSN 2090-4479, http://dx.doi.org/10.1016/j.asej.2010.09.007.

[5] G. Fink and T. Plotz, "On the Use of Context-Dependent Modeling Units for HMM-Based Offline Handwriting Recognition," Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Parana, 2007, pp. 729-733. doi: 10.1109/ICDAR.2007.4377011

[6] Mohammad Imrul Jubair and Prianka Banik. Article: A Simplified Method for Handwritten Character Recognition from Document Image. International Journal of Computer Applications 51(14):50-54, August 2012

[7] P. Kumawat, A. Khatri and B. Nagaria, "Comparative Analysis of Offline Handwriting Recognition Using Invariant Moments with HMM and Combined SVM-HMM Classifier," Communication Systems and Network Technologies (CSNT), 2013 International Conference on, Gwalior, 2013, pp. 140-143.

[8] P. Maji, S. Chatterjee, S. Chakraborty, N. Kausar, S. Samanta and N. Dey, "Effect of Euler number as a feature in gender recognition system from offline handwritten signature using neural networks," Computing for

Sustainable Global Development (INDIACom), 2015 2nd International Conference on, New Delhi, 2015, pp. 1869-1873.

[9] Monica Patel and Shital P. Thakkar, Handwritten Character Recognition in English: A Survey, International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2,pp no. 345 350, February 2015

[10] P.Shankar Rao and J.Aditya, Handwriting Recognition - Offline Approach

[11] Plamondon, Rjean, and Sargur N. Srihari. "Online and offline handwriting recognition: a comprehensive survey." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.1 (2000): 63-84

[12] Plamondon, R., Lopresti, D.P., Schomaker, L.R.B. Srihari, R. (1999).On-line handwriting recognition. In: J.G. Webster (Ed.). Wiley Encyclopedia of Electrical and Electronics Engineering, pp. 123-146, New York: Wiley.

[13] Vijay Laxmi Sahu, Babita Kubde, "Offline Handwritten Character Recognition Techniques using neural network: A Review", International journal of science and Research (IJSR), pp. 1-8.

[14] P. Slavik and V. Govindaraju, "Equivalence of different methods for slant and skew corrections in word recognition applications," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 3, pp. 323-326, Mar 2001. doi: 10.1109/34.910885

[15] S. N. Srihari, M. J. Beal, K. Bandi, V. Shah and P. Krishnamurthy, "A statistical model for writer verification," Eighth International Conference on Document Analysis and Recognition (ICDAR'05), 2005, pp. 1105-1109 Vol. 2. doi: 10.1109/ICDAR.2005.33

[16] C.Y. Suen, M. Berthod and S. Mori, Automatic Recognition of Handprinted-Characters-the State of the Art in Proceedings of the IEEE, Vol: 68, No: 4, 1980.

[17] Wikipedia contributors. "Handwriting Recognition." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 28 July. 2016.

[18] Wikipedia contributors. "Feedforward neural network." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 31 July. 2016.

[19] Wikipedia contributors. "Recurren neural network." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 31 July. 2016.

[20] Dong Xiang, Huahua Yan, Xianqiao Chen and Yanfen Cheng, "Offline Arabic handwriting recognition system based on HMM," Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on, Chengdu, 2010, pp. 526-529. doi: 10.1109/ICCSIT.2010.5564429