Writer-independent Offline Handwritten Signature Verification using Novel Feature Extraction Techniques

Md. Aminur Rahman
Ahsanullah University of Science and Technology, Dhaka, Bangladesh

Sarker Miraz Mahfuz
Ahsanullah University of Science and Technology, Dhaka, Bangladesh

S. M. Abdullah Al-Mamun
Ahsanullah University of Science and Technology, Dhaka, Bangladesh

ABSTRACT
Signature is critical for authentication and authorization in commercial, financial and legal transactions and fittingly, it is one of the most commonly used biometrics for authentication. Hence, an accurate and efficient signature verification system is required. The objective of signature verification is to discriminate the original signatures from the forged ones. It is a challenging task as even two signatures of the same person possess variations in different areas such as the starting and ending positions, the angle of inclination, relative spacing between letters, height, width etc. Offline signature verification is even more challenging as it is devoid of the dynamic information about the signing process. Although numerous research works have been done in the area of offline signature verification in last decades, it still remains an open research problem. There are three common phases in signature verification system: image preprocessing, feature extraction and verification. In this paper, two novel features have been presented that can be extracted from preprocessed signature images in the feature extraction phase. The proposed features are: i) Stroke angle and average intersected points ii) Pixel density of the signature nucleus. The goal of this research is to strengthen the feature set with the proposed features what will help to get more accurate verification of the signatures.

General Terms
Signature Recognition and Verification, Pattern Recognition, Image Processing.

Keywords
Offline Signature Verification, Biometric Authentication, Forgery Detection, Neural Network, Novel Features.

1. INTRODUCTION
Handwriting is a skill that is highly personal to individuals which consists of graphical marks on the surface in relation to a particular language and that is why the signature of any person is an important biometric characteristic which is usually implemented for personnel identification or document authentication. Signature usage dates from ancient times and is held until today as a means of authorization. In biometric authentication system [1], a person is recognized based on physiological or behavioral traits. The measurements of biological traits, such as fingerprint, face, iris, etc. are used in the first case where the latter case is related to behavioral traits such as voice and the handwritten signature. As it is one of the most widespread biometrics for authentication and authorization in legal transactions, the need for efficient automated solutions for signature recognition and verification has increased in recent years. The recognition part is concerned with the identification of the signature owner while the verification part is liable for the decision whether a signature is genuine (produced by claimed individual) or forged (produced by an imposter). The identification is done by comparing the input signature of a subject with samples from all subjects in the database and the verification is done by comparing the input signature image with samples from the same subject [2].

Signatures can be of different types. In a broad sense, based on form and content signatures can be classified in three types:

(1) Simple: These signatures are the ones where the person just scripts his or her name in a stylish manner. Very often it is very easy to interpret all the characters in these signatures.

(2) Cursive: Cursive signatures are more complex. Though the signatures still contain all the individual characters within the name, they are however drafted in a cursive manner, usually in a single stroke.

(3) Graphical: Signatures are classified as graphical when they portray complex geometric patterns. It is very difficult to deduce the name of the person from a graphical signature as it is more of a sketch of the name of the signer.

The forgery of signatures can be classified in three groups: Random Forgery, Unskilled Forgery and Skilled Forgery [4].

(i) Random Forgery: In this type of forgery, a signature is created by some individual knowing only the name of the person whose signature is to be made. It is the easiest type to detect. It is also known as 'simple forgery'.

(ii) Unskilled Forgery: When a signer, without any prior experience, creates a signature after observing the original signature once or twice.

(iii) Skilled Forgery: This type of forgery is created by someone who may be a professional in replicating signatures and possesses prior experience. The imposter creates a signature only after enough practicing over it.

Examples of different kinds of signatures with skilled and unskilled forgeries have been shown in the Fig 1.

Based on the signatures acquisition techniques, the writer independent (a single model is used to classify images from any user) signature verification systems are categorized into two types: Offline Signature Verification and Online Signature Verification [4]. In online signature verification, the signatures are captured during the signing process and hence, the dynamic information can be extracted. On the other hand,

in offline signature verification, the signatures are captured as digital images once the signing process is over and thus, only static information can be extracted.

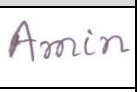
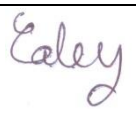
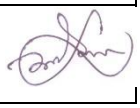
Type	Genuine	Skilled Forgery	Unskilled Forgery
Simple			
Cursive			
Graphical			

Fig 1: Examples of three types of signatures with forgeries

The accuracy of the offline signature verification system largely depends on the image preprocessing and feature extraction techniques. In this paper, the focus has been given on feature extraction phase. Two novel features along with the techniques to extract those features have been proposed. These proposed features have been extracted along with other well-established features to strengthen the feature vector. This paper is organized as follows: the problem statement has been described in section-2 and then section-3 illustrates the methodology. The experimental results have been presented in section-4. Finally, section-5 concludes with the summarization of the entire work presented in this paper.

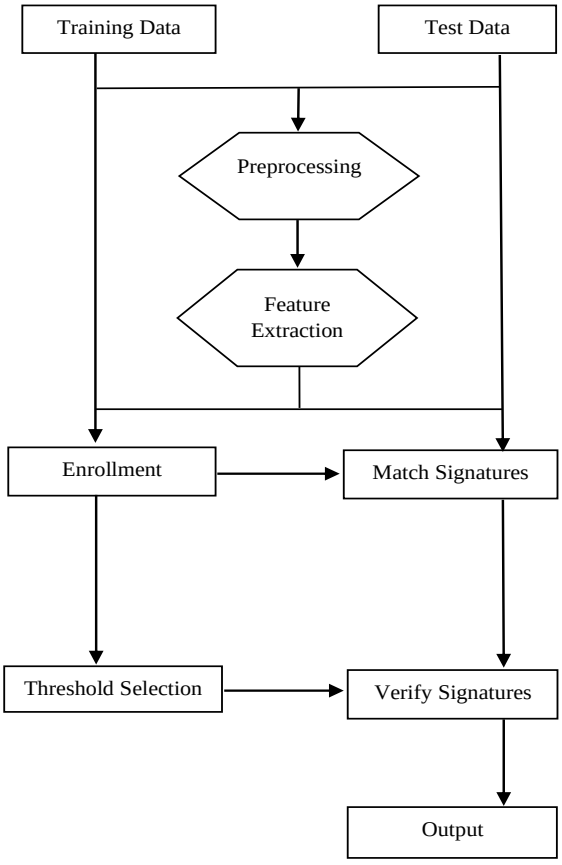


Fig 2: Process of Signature Verification

2. PROBLEM STATEMENT

The identification and verification of someone's signature accurately is the problem in using signature as biometric for authentication. It is a tough job to perform for various reasons. The major drawback of signature is having high intra-class variability, meaning that people cannot sign the same signature over and over again. Every signature has several features like spikes, loops, edges, overall size, slants etc. in it and it is inevitable that even two signatures taken from the same person in same circumstances will have variations between them in those abovementioned features. The problem is accentuated by this non-repetitive nature of variations of signatures.

complex. The problem gets into its worst state when someone tries to imitate the signature of a person with the purpose of fraud or false representation. In this case, a person is targeted and the imposter practices that person's signature over and over again to perfectly resemble the original signature. So skilled forgeries take place.

3. METHODOLOGY

From the studies of previous research works in this area, it has been found that there are some common steps in an offline signature verification process [10]. The steps are:

(a) Preprocessing

(b) Feature Extraction

(c) Classification/Verification

Preprocessing: Preprocessing of the signature images is an important part of the system to verify handwritten signature where inverting into gray level, binarization, cutting edges, thinning, noise removal, orientation of the image etc. are performed to process the image of the signature for further operations. If the preprocessing cannot be done properly then the result of feature extraction and ultimately the verification will not be good enough.

Feature Extraction: To do the verification, some key features of the signatures are required. The features are extracted from the preprocessed signature images in this step. Some of the common, well established features of handwritten signatures are aspect ratio, vertical and horizontal histogram, center of gravity, Hough transform, signature occupancy ratio, baseline slant angle etc.

Verification: Verification is the final step of the whole process where the unique features extracted from the image in the previous step are used to verify the sample signature with the original signature stored in the database. Generally, classifiers for signature verification can be classified into two types: *writer-dependent* and *writer-independent* [3][6]. In case of first approach, a model is trained for each of the user but on the other hand, in case of writer-independent approach, a single model is trained for all the users to classify a query input.

3.1 PREPROCESSING

Preprocessing of the input signature images is required for many reasons such as noise removal, enhancing image features etc. so that the next two steps, feature extraction and verification, can be done with more accuracy. In this research, following preprocessing techniques have been applied on the input signature images.

3.1.1 Grayscale Image

Grayscale image, which is an image with colors ranging from white to black. Darker areas are blacker and lighter areas are whiter and the area in between is gray in different levels. To convert the RGB image that has been used as input image into grayscale image, *rgb2gray* MATLAB function can be applied.

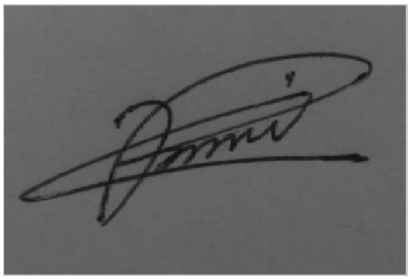


Fig 3: RGB image converted to grayscale image

3.1.2 Resizing

At the outset of preprocessing, resizing the input images was required as the input signature images normally comes at different sizes. Complying with most of the research works in this area, the images were resized as 256*256 [2][4].

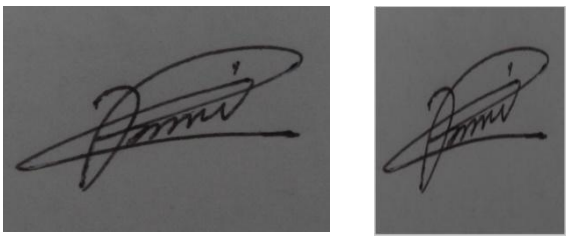


Fig 4: (left) Original image and (right) Resized image (256*256).

3.1.3 Removal of Noise

The images of the signatures mainly contain a type of noise called "Salt Pepper Noise". Two morphological operations called *Erosion* and *Dilation* can be applied to remove those noises from the input images.

Image + Erosion + Dilation = Noise Free Image

3.1.4 Binarization

Binarization means that the image will be represented in a binary format; like by only black and white. There will be no color in between, this is the main difference between a grayscale image and a binarized image. The binarization of a grayscale image can be done by various algorithms, in this research, canny algorithm has been used with the help of *edge* MATLAB function.

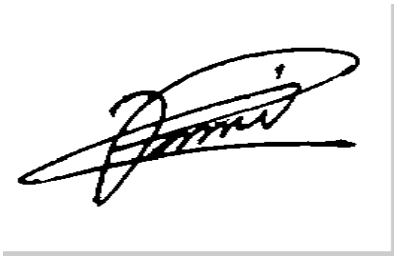


Fig 5: Grayscale image converted to binary image.

3.1.5 Thinning

Thinning is a morphological operation that is used to transform a digital image into a simplified, but topologically equivalent image. This operation can be used on the binary images for skeletonization of the images by removing selected foreground pixels from the binary images. Here, MATLAB function *bwmorph* was used with logical negative version of the binary images from the previous step and 'skel' for algorithm as arguments.

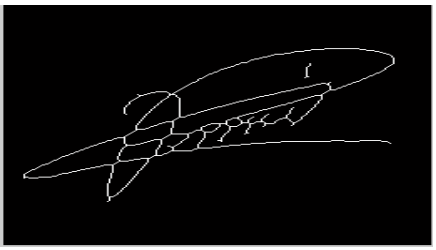


Fig 6: Skeletonization of negative binary image

3.2 FEATURE EXTRACTION

As a result of decades of research in the area of signature verification, a set of feature extraction techniques which are most effective for offline signature verification has already been established [8][9][23]. In this work, those predefined features have been implemented along with the proposed two features. Then a feature vector was created including all those features for the verification phase. The implemented features are as follows:

3.2.1 Aspect ratio

Aspect ratio is the height width ratio of a signature. The ratio is obtained by dividing signature height (h) to signature width (w). The height is the maximum length of the columns obtained from the resized image. Similarly, width is also calculated considering the rows of maximum length. Signature height and width can change but height-to-width ratio of an individual's signature is approximately constant. The aspect ratio, R_i for i^{th} sample signature image can be calculated as follows:

$$R_i = h_i / w_i$$

where h_i is the height of the i^{th} sample signature and w_i is the width of the i^{th} sample signature.

3.2.2 Occupancy ratio

This ratio is defined by the number of pixels which belong to the signature divided by the total pixel count in the signature image. This is also known as *signature density*. The Signature Density, D_i for i^{th} sample signature image can be calculated as follows:

$$D_i = l_i / x_i$$

where l_i is the number of pixels which belong to the signature of the i^{th} sample signature and x_i is the total number of pixels in i^{th} sample signature.

3.2.3 Center of gravity

The center of gravity (CG) of an object is the point at which weight is evenly dispersed and all sides are in balance. The point G (x_g,y_g) on the Fig 7 where lines A and B are crossing is the center of gravity of the signature in the image [9]. The point provides information about the layout of the pixel density. Those A and B lines divide the signature image into vertical and horizontal regions so that the number of pixels is the same in each region.

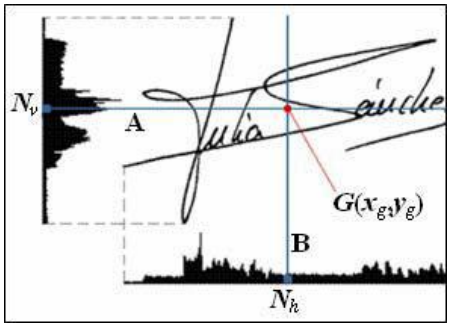


Fig 7: Center of Gravity of a signature image.

3.2.4 Slope of the line joining the centers of gravity of the two halves of signature image

The signature image is divided into left and right halves within its bounding box, and the centers of gravity of the two halves are determined separately [23]. Then the slope of the line joining the two centers is calculated which is an attractive feature to distinguish signatures.

3.2.5 Maximum horizontal and vertical histogram ratio

In this method, the black pixels in each row and column of a signature image are counted [23]. The specific row with maximum black pixels is found and the pixel count is recorded as the maximum horizontal histogram. Similarly, the specific column with maximum black pixels is found and the pixel count is recorded as the maximum vertical histogram. Then the maximum horizontal and vertical histogram ratio, H_i for i^{th} sample signature image can be calculated as follows:

$$H_i = t_i / v_i$$

where t_i is the maximum horizontal histogram of the i^{th} sample signature and v_i is the maximum vertical histogram of the i^{th} sample signature.

3.2.6 Number of edge points

An edge point of a signature is that pixel, which belongs to the signature, has only one neighbor in its 8-neighbor. To extract the edge points in a given signature, generally a 3*3 structuring element is used with all coefficients equal to 1 [5].

3.2.7 Number of cross points

A cross point of a signature is that point, which belongs to the signature, has at least three 8-neighbors. To extract the cross points in a given signature, generally a 3*3 structuring element is used with all coefficients equal to 1 [5].

Along with the abovementioned predefined features, the proposed novel features have also been extracted to make the feature vector more effective for signature verification.

3.2.8 Proposed Features

The proposed features are described in the following sections.

3.2.8.1 Stroke angle and average intersected points

In normal circumstances when someone writes his/her signature down on a piece of paper or something else, the signature always contains an angle which is being called *stroke angle* of the signature. It is found that this angle of the signature remains almost same for any individual. But it will not be true all the time. When someone writes his/her signature in any position s/he is not accustomed to then the stroke angle may vary. But it is assumed that the sample signatures are written in normal circumstances.

Process: Stroke angle of a signature above 45 degree is near impossible. So, the process of finding the stroke angle was formulated for 0 to 45 degree. The steps of the process for an angle between 0 and 45 are given below:

Step 1: First of all, the equation of a line of that selected angle needs to be formed by the straight-line equation ($y = mx + c$) and the slope (m) what can be calculated from the angle in degree.

Step 2: Now that straight line needs to be moved from the bottom of the signature to the top of that so that the line can cover the whole signature. This can be done by varying the constant value (c) of the straight-line equation.

Step 3: The intersected points of the signature by the straight line needs to be counted in every step of the line moving from bottom to top or vice-versa of the signature.

Step 4: After all the iterations for a single angle of 0 to 45 degree, the average of the intersected points is counted by the equation:

Average intersected points = Total intersected points/ no. of iterations needed to cover the whole signature.

Those four steps were executed for every angle between 0 and 45 degree and when it was done, the angle was selected as the stroke angle for which the average intersected points were maximum.

Both the angle and the average value of the intersected points were enlisted in the feature set as features.

3.2.8.2 Pixel density of the nucleus

First of all, the goal was to find out the nucleus of a signature and then the pixel density of that area. To keep the thing simple, it was assumed that 1/8 portion of the signature image would be the size of the nucleus. As 256*256 images of signature were used as samples, the size of the nucleus had to be 32*32. Then the number of pixels in each and every area of 32*32 in a sample image has been counted. In this process the nucleus, that contains maximum number of pixels in it, was found. At last, the density of the nucleus was found by the equation below.

*Pixel density of nucleus = Total pixel in the nucleus/ 32*32*

3.3 VERIFICATION

This is the final step of the signature verification process. After preprocessing the images of signatures appropriately, effective features are extracted from those preprocessed images. Then those extracted features are used to train a model that is used for the verification i.e. for the classification of original or forged signature. Many researchers in the past have tried to improve this verification process by inventing different methods or improving existing methods but not all of them were equally effective or efficient [4][12-21]. Some of the efficient methods for offline signature verification are:

a. Template Matching approach
b. Neural Network (NN) approach,
c. Hidden Markov Model (HMM) approach [23],
d. Statistical approach
e. Structural and Syntactic approach etc.

Neural Network (NN) has been used in this research because from the previous research papers it is found that this approach gives better results than other techniques in offline

signature classification [15-17].

3.3.1 Neural Network Architecture

There are two different classes of network architectures [18]:

i. *Single layer feed forward network:* There is just a single layer of weights in this type of network where the inputs are directly connected to the outputs via a series of weights without any hidden layer in between them [fig 8]. The sum of products of the weights and the inputs is calculated in each neuron node, and if the value is above some threshold, the neuron fires.

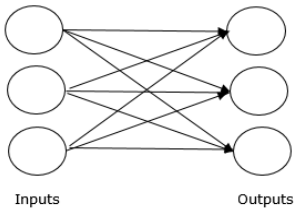


Fig 8: Single layer feed forward network

ii. *Multi-layer feed forward network:* This type of network possesses one or more hidden layers in between inputs and outputs [fig 9]. The neurons of the first hidden layer are supplied with the inputs from the source nodes in the input layer and the outputs of the first hidden layer neurons are feed as inputs to the neurons of the second hidden layer and so on. If every node in each layer of the network is connected to every other node in the adjacent forward layer, then the network is called fully connected. On the other hand, if some of the links are missing, the network is called partially connected.

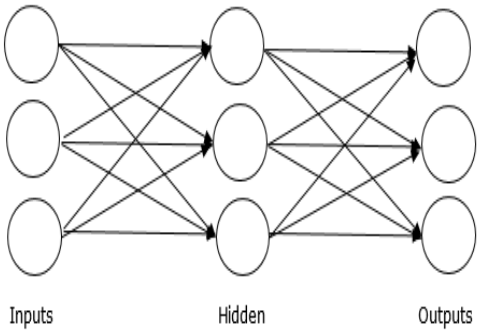


Fig 9: Multi-layer feed forward network

A multi-layer feed forward network with sigmoid hidden layer and softmax output neuron has been used for the verification phase of this investigation. This network can classify vectors well, given enough neurons in its hidden layer. 20 hidden layers have been used here.

4. EXPERIMENTAL RESULTS

The contribution of this research work is in the feature extraction phase where a feature vector has been created with the proposed two novel features. In this section of this paper, the performance of the whole signature verification is presented. To gauge the performance, the 'handwritten signatures' dataset, offered on kaggle.com for free, has been used. The dataset can be found here at, `shorturl.at/wAGL6`. The test was started with 55% of data of total sample data as training set and went up to 85% percent to explore the performance result. It was seen that the performance of the system got better and better along with the increment of training data but after a certain stage the results became more or less stable and that almost stable result was found using 75% of total sample data as training data.

Table 1. Comparison of error rate with training data

Percentage of training data	Error percentage of test data
55	7.05882
60	5.63380
65	5.26315
70	4.65116
75	3.57142

Related figures and graphs are given below for better understanding of the performance of the system using 75% of total sample signature image as training set.

Table 2. The final result of the verification system

	Samples	% of Error
Training	213	3.2863
Validation	43	2.3255
Testing	28	3.5714

The confusion matrices and ROC curves for training phase, validation phase and test phase are given in fig 10 and fig 11 respectively. The last one (bottom-right) is representing the combined result.

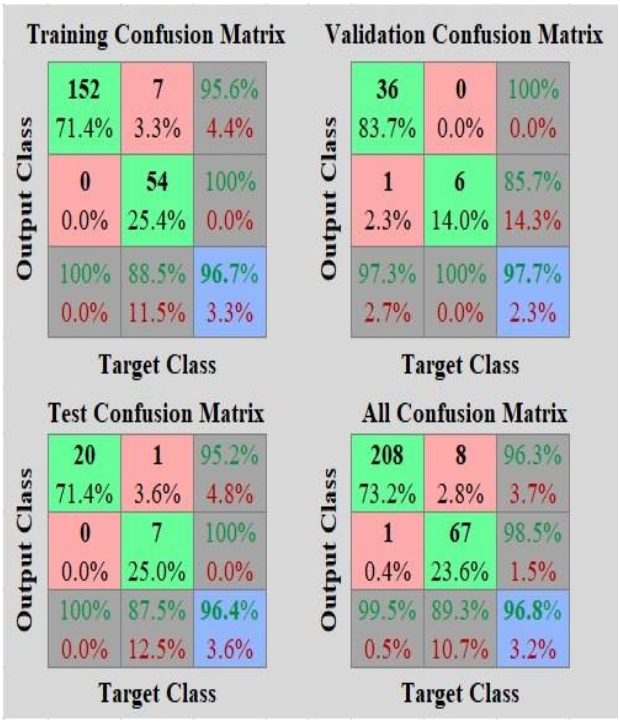


Fig 10: Training, Validation and Test confusion matrices

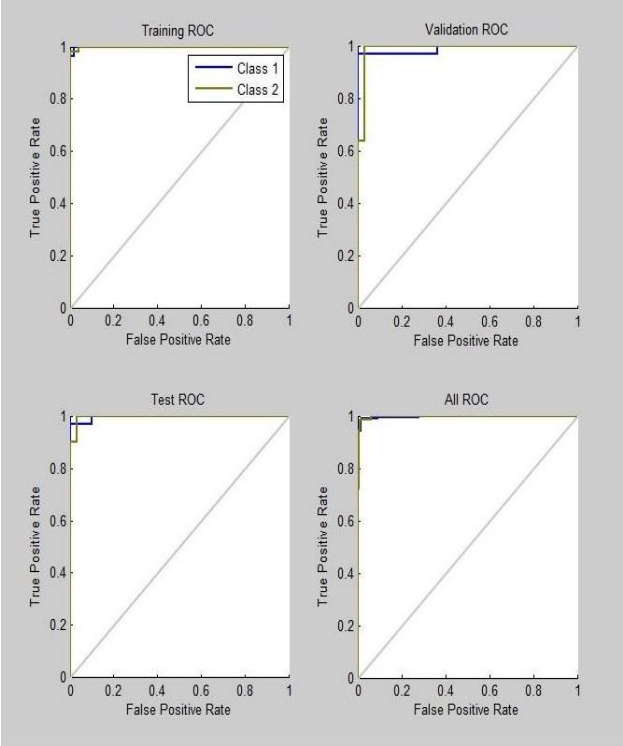


Fig 11: Training, Validation and Test ROC curves

5. CONCLUSION

In this paper, the concentration has been put on feature extracting algorithms because if a good feature set cannot be prepared, the classification of signatures will not be done correctly by the classifiers. Here a feature set has been prepared with the proposed two novel features along with other feature those have been invented over many years by the researchers. At the end, neural network has been used for the classification and impressive result has come out with a very low error rate. Though the system shows very good result, there is still scope for improving the system such as solving the orientation problem of the input sample signature images. It is possible that the orientation of sample input images of the signatures may not be presented in ideal orientation and in that case this system might not work as expected. So, research work can be carried on to fix the orientation of the input images before going into the signature verification steps.

6. REFERENCES

[1] Debnath Bhattacharyya, Rahul Tanjan, Farkhod Aliserov A, and Minkyu Choi, "Biometric Authentication: A Review", *International Journal of u and e - Service, Science and Technology*, 2009.

[2] Siti Norul Huda Sheikh Abdullah and Khairuddin Omar, "State-of-the-Art in Offline Signature Verification System", *International Conference on Pattern Analysis and Intelligent Robotics (ICPAIR)*-2011 IEEE.

[3] D. Bertolini, L. S. Oliveira, E. Justino, and R. Sabourin. "Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers". *Pattern Recognition*, 43(1), January 2010.

[4] Henali P., Shivani D., Pooja D and Abha D. "Review on offline signature recognition and verification techniques". *International Journal of Computer Applications (IJCA)*, June 2018.

[5] Juan Hu and Youbin Chen. "Offline Signature Verification Using Real Adaboost Classifier Combination of Pseudo-dynamic Features". *Document Analysis and Recognition, 12th International Conference on*, pages 1345–1349, August 2013.

[6] Luiz G. Hafemann, Robert Sabourin, and Luiz S. Oliveira. "Writer independent feature learning for Offline Signature Verification using Deep Convolutional Neural Networks". *International Joint Conference on Neural Networks*, pages 2576–2583, July 2016.

[7] M. Pourshahabi, M. Hoseyn Sigari, and H. Pourreza. "Offline handwritten signature identification and verification using contourlet transform". *Soft Computing and Pattern Recognition, Int. conference of*. IEEE, 2009.

[8] J. Ruiz-del Solar, C. Devia, P. Loncomilla, and F. Concha. "Offline Signature Verification Using Local Interest Points and Descriptors". *Progress in Pattern Recognition, Image Analysis and Applications*, number 5197. Springer, 2008.

[9] J. F. Vargas, M. A. Ferrer, C. M. Travieso, and J. B. Alonso. "Offline signature verification based on grey level information using texture features". *Pattern Recognition*, 44(2):375–385, February 2011.

[10] R. Zouari, R. Mokni, and M. Kherallah. "Identification and verification system of offline handwritten signature using fractal approach". *Image Processing, Applications and Systems Conference (IPAS), 2014 First International*, pages 1–4, November 2014.

[11] L. Ravi Kumar and A.Sudhir Babu, "Genuine and Forged Offline Signature Verification Using Back Propagation Neural Networks", *(IJCSIT) International Journal of Computer Science and Information Technologies*, 2011.

[12] B H Shekar and R.K.Bharathi, "Eigen-signature: A Robust and an Efficient Offline Signature Verification Algorithm" *IEEE-International Conference on Recent Trends in Information Technology, ICRTI- 2011* IEEE.

[13] Miguel A. Ferrer, Francisco Vargas, Carlos M. Travieso and Jesus B. Alonso, "Signature Verification using Local Directional Pattern (LDP)" *International Conference on computer security technology (ICCST)*-2010.

[14] Meenakshi K. Kalera, Sargur Srihari, and Aihua Xu. "Offline signature verification and identification using distance statistics". *International Journal of Pattern Recogn.*, 18(07):1339–1360, November 2004.

[15] S.T. Kolhe, S. E. Pawar. "Offline Signature Verification Using Neural Network", *International Journal of Modern Engineering Research (IJMER)*, Vol.2, Issue.3, May-June 2012.

[16] Ashwini Pansare, Shalini Bhatia "Off-line Signature Verification Using Neural Network", *International Journal of Scientific & Engineering Research*, Volume 3, Issue 2, February-2012.

[17] Paigwar Shikha, Shukla Shailja, "Neural Network Based Offline Signature Recognition and Verification System", Research Journal of Engineering Sciences ISSN 2278 – 9472, Vol. 2(2), 11-15, February 2013.

[18] P. Fishwick, "Neural network models in simulation: A comparison with traditional modeling approaches,"

Working Paper, University of Florida, Gainesville, FL, 1989.

[19] Peter Shaohua Deng, Hong-Yuan Mark Liao, Chin Wen Ho, and Hsiao-Rong Tyan. "Wavelet-Based Off-Line Handwritten Signature Verification". *Computer Vision and Image Understanding*, 76(3), December 1999.

[20] A. El-Yacoubi, E. J. R. Justino, R. Sabourin, and F. Bortolozzi. "Offline signature verification using HMMs and cross-validation". *Neural Networks for Signal Processing X, 2000. Proceedings of the 2000 IEEE Signal Processing Society Workshop*, volume 2. IEEE, 2000.

[21] Dr. S. Adebayo Daramola and Prof. T. Samuel Ibiyemi, "Offline Signature Recognition using Hidden Markov Model (HMM)", *International Journal of Computer Applications 10(2):17-22, November 2010*.

[22] S. Ghandali and M.E. Moghaddam. "A Method for Off-line Persian Signature Identification and Verification Using DWT and Image Fusion". *IEEE International Symposium on Signal Processing and Information Technology, ISSPIT*, pages 315–319, December 2008.

[23] Daramola Samuel, Ibiyemi Samuel, "Novel Feature Extraction Technique for Off-line Signature Verification System", *International Journal of Engineering Science and Technology*, Vol. 2(7), 2010, 3137-3143.