IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

SIGNATURE VERIFICATION USING IMAGE PROCESSING AND NEURAL NETWORKS

1 B.Akhila, 2 G.Nikhila , 3 A Lakshmi , 4 G.Jahnavi, 5 Mrs .J.Himabindhu.

1Student, 2Student, 3Student, 4Student, 5Assistant Professor

1Vignan's Institute of Engineering for Women,

2Vignan's Institute of Engineering for Women,

3Vignan's Institute of Engineering for Women,

4Vignan's Institute of Engineering for Women,

5Vignan's Institute of Engineering for Women.

Abstract:

The fact that the signature is widely used as a means of personal verification emphasizes the need for an automatic verification system. Verification can be performed either Offline or Online based on the application. Online systems use dynamic information of a signature captured at the time the signature is made. Offline systems work on the scanned image of a signature. We have worked on the Offline Verification of signatures using a set of shape based geometric features. The features that are used are Baseline Slant Angle, Aspect Ratio, Normalized Area, Center of Gravity, number of edge points, number of cross points, and the Slope of the line joining the Centers of Gravity of two halves of a signature image. Before extracting the features, pre-processing of a scanned image is necessary to isolate the signature part and to remove any spurious noise present. The system is initially trained using a database of signatures obtained from those individuals whose signatures have to be authenticated by the system. For each subject a mean signature is obtained integrating the above features derived from a set of his/her genuine sample signatures. This mean signature acts as the template for verification against a claimed test signature. In this paper, we present how the problem has been handled in the past few decades, analyze the recent advancements in the field, and the potential directions for future research.

Keywords: Signature verification, Image processing, Artificial Neural Networks, Pre- processing, Feature Extraction, Back propagation, Histogram of oriented gradients.

INTRODUCTION

For any legal transactions the authorization is done by the signature. So the need of the signature verification increases. The handwritten signatures are unique for individuals and which is impossible to duplicate. The technology is easy to explain and trust. The primary advantage that signature verification systems have over other type's technologies is that signatures are already accepted as the common method of identity verification.

The handwritten signature verifications are of two types Online and the offline.

On-line method uses an electronic technique and a computer to extract information about a signature and takes dynamic information like pressure, velocity, speed of writing etc. for the purpose of verification. In off-line signature verification involves less electronic control and uses signature images captured by scanner or camera.

An off-line signature verification system uses features extracted from scanned signature image. The features used for offline signature verification are much simpler. In this only the pixel image needs to be evaluated. But, the off-line systems are difficult to design as many desirable characteristics such as the order of strokes, the velocity and other dynamic information are not available in the off-line case. The verification process has to wholly rely on the features that can be extracted from the trace of the static signature images. In the area of Handwritten Signature Verification (HSV), specially offline HSV, different technologies have been used and still the area is being explored.

PROBLEM STATEMENT

The problem of automatic handwritten signature verifications commonly modelled as a verification task: given a learning set L, that contains genuine signatures from a set of users, a model is trained. This model is then used for verification: a user claims an identity and provides a query signature X new. The model is used to classify the signature as genuine(belonging to the claimed individual) or forgery (created by someone else). To evaluate the performance of the system, we consider a test set T , consisting of genuine signatures and forgeries. The signatures are acquired in an enrolment phase, while the second phase is referred to operations (or classification) phase.

LITERATURE SURVEY

Vigorous research has been pursued in handwriting analysis and pattern matching for a number of years. In the area of Handwritten Signature Verification (HSV), especially offline HSV, different technologies have been used and still the area is being explored. In this section we review some of the recent papers on offline HSV.

The approaches used by different researchers differ in the type of features extracted, the training method, and the classification and verification model used.

Hidden Markov Models Approach

Hidden Markov Model (HMM) is one of the most widely used models for sequence analysis in signature verification. Handwritten signature is a sequence of vectors of values related to each point of signature in its trajectory. Therefore, a well chosen set of feature vectors for HMM could lead to the design of an efficient signature verification system. These Models are stochastic models which have the capacity to absorb the variability between patterns and their similarities. In HMM stochastic matching (model and the signature) is involved. This matching is done by steps of probability distribution of features involved in the signatures or the probability of how the original signature is calculated. If the results show a higher probability than the test signatures probability, then the signatures is by the original person, otherwise the signatures are rejected. A HMM is used to model each writer signature. The method achieves an AER of 18.4% for a set of 440 genuine signatures from 32 writers with 132 skilled forgeries.

Neural Networks Approach

The main reasons for the widespread usage of neural networks (NNs) in pattern recognition are their power and ease of use. A simple approach is to firstly extract a feature set representing the signature (details like length, height, duration, etc.), with several samples from different signers. The second step is for the NN to learn the relationship between a signature and its class (either "genuine" or "forgery"). Once this relationship has been learned, the network can be presented with test signatures that can be classified as belonging to a particular signer. NNs therefore are highly suited to modelling global aspects of handwritten signatures. The proposed system uses structure features from the signatures contour, modified direction feature and additional features like surface area, length skew and centroid feature in which a signature is divided into two halves and for each half a position of the centre of gravity is calculated in reference to the horizontal axis. For classification and verification two approaches are compared the Resilient Back propagation (RBP) neural network and Radial Basic Function(RBF) using a database of 2106 signatures containing 936 genuine and 1170 forgeries. These two classifiers register 91.21% and 88% true verification respectively.

Template matching approach

Fang et al. proposed two methods for the detection of skilled forgeries using template matching. One method is based on the optimal matching of the one-dimensional projection profiles of the signature patterns and the other is based on the elastic matching of the strokes in the two dimensional signature patterns. Given a test signature to be verified, the positional variations are compared with the statistics of the training set and a decision based on a distance measure is made. Both binary and grey-level signature images are tested. The average verification error rate of 18.1% was achieved when the local peaks of the vertical projection profiles of grey-level signature images were used for matching and with the full estimated covariance matrix incorporated.

Statistical approach

Using statistical knowledge, the relation, deviation, etc between two or more data items can easily be found out. To find out the relation between some set of data items we generally follow the concept of Correlation Coefficients. In general statistical usage refers to the departure of two variables from independence. To verify an entered signature with the help of an average signature, which is obtained from the set of, previously collected signatures, this approach follows the concept of correlation to find out the amount of divergence in between them. In this approach various features are extracted which include global features like image gradient, statistical features derived from distribution of pixels of a signature and geometric and topographical descriptors like local correspondence to trace of the signature. The classification involves obtaining variations between the signatures of the same writer and obtaining a distribution in distance space. For any questioned signature the method obtains a distribution which is compared with the available known and a probability of similarity is obtained using a statistical Kolmorogorv-Smirnov test.

Using only 4 genuine samples for learning, the method achieves 84% accuracy which can be improved to 89% when the genuine signature sample size is increased. This method does not use the set of forgery signatures in the training/learning.

Support Vector Machine

Support Vector Machines (SVMs) are machine learning algorithms that uses a high dimensional feature space and estimate differences between classes of given data to generalize unseen data. The system uses global, directional and grid features of the signature and SVM for classification and verification. The database of 1320 signatures is used from 70 writers. 40 writers are used for training with each signing 8 signatures thus a total of 320 signatures for training. For initial testing, the approach uses 8 original signatures and 8 forgeries and achieves FRR 2% and FAR 11%.

EXISTING SYSTEM AND DRAWBACKS

Many researchers has been pursued in handwriting analysis and pattern matching for a number of years. In the area of Handwritten Signature Verification (HSV), especially offline HSV, different technologies have been used and still the area is being explored. It takes more time to classify the genuine or forged.

The major disadvantage of signature verification is that it uses large dataset for greater accuracy.

It takes more time to verify the original signature and forgery signature , because sometimes forgeries are same like original signatures.

PROPOSED SYSTEM

we present a model in which Neural Network classifier is used for verification.

Signatures from database are pre-processed prior to feature extraction. Features are extracted from pre-processed signature image. These extracted features are then used to train a neural network.

In verification stage, on test signatures pre-processing and feature extraction is performed. These extracted features are then applied as input to a trained neural network which will classify it as a genuine or forged signature.

Pre- processing:

The signature is first captured and transformed into a format that can be processed by a computer. Now it's ready for pre-processing. In pre-processing stage, the RGB image of the signature is converted into grayscale and then to binary image. The pre-processing stage includes:

Denoising: One of the fundamental challenges in the field of image processing and computer vision is image denoising, where the underlying goal is to estimate the original image by suppressing noise from a noise contaminated version of the image. Image noise may be caused by different intrinsic (i.e., sensor) and extrinsic (i.e., environment) conditions which are often not possible to avoid in practical situations. An alternative approach to the problem of image denoising based on data-adaptive stochastic optimization via Markov-Chain Monte Carlo sampling.

Fig: Before Denoising

Fig: After Denoising

Color invention: The true color image RGB is converted to the grayscale intensity image by saturation information while retaining the luminance.

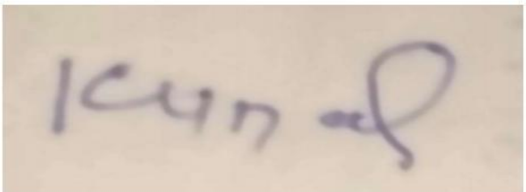

Fig: A sample signature to be processed

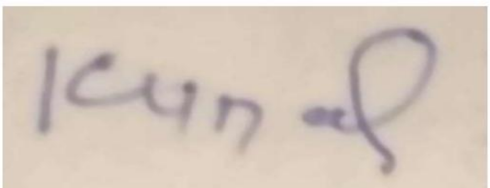

Fig: A Gray intensity image

Grayscale Images: A gray-scale image is a data matrix whose values represent intensities within some range where each element of the matrix corresponds to one image pixel. It is simply one in which the only colors are shades of gray. The reason for differentiating such images from any other sort of color image is that less information needs to be provided for each pixel. In fact a `gray' color is one in which the red, green and blue components all have equal intensity in RGB space, and so it is only necessary to specify a single intensity value for each pixel, as opposed to the three intensities needed to specify each pixel in a full color image.

Grayscale images are very common, in part because much of today's display and image capture hardware can only support 8-bit images. In addition, grayscale images are entirely sufficient for many tasks and so there is no need to use more complicated and harder-to-process color images.

Image Filtering and Binarization: Any image when resample is filtered by a low pass FIR filter. This is done to avoid aliasing. This aliasing occurs because of sampling the data at a rate lower than twice the largest frequency component of the data. So a low pass filter will remove the image high frequency components. And for this purpose the filter used. Now the grayscale image is segmented to get a binary image of objects. In a binary image, each pixel assumes one of only two discrete values: 1 or 0. A binary image is stored as a logical array.

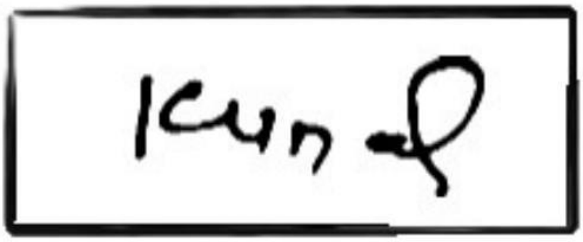

Fig : Binary Image interpreting the bit value of 0 as black and 1 as white

Future Extraction:

We have experimented with two separate features: histogram of oriented gradients (HOG) relative to the dominant orientation and local binary patterns .

Histogram of oriented gradients(HOG): Histogram Orientation Gradient (HOG) is used for feature shape representation, which was introduced by Dalal and Triggs at the CVPR conference in 2005 . HOG is basically used for person detector, which stands for Histograms of Oriented Gradients. In this research, HOG has been adopted to be as a feature extraction technique to recognize and authenticate the signature image.

Gradient Computating: The first step of calculation in many feature detectors in image pre-processing is to ensure normalized color and gamma values. As Dalal and Triggs point out, however, this step can be omitted in HOG descriptor computation, as the ensuing descriptor normalization essentially achieves the same result. Image pre-processing thus provides little impact on performance. Instead, the first step of calculation is the computation of the gradient values. The most common method is to apply the 1-D centered, point discrete derivative mask in one or both of the horizontal and vertical directions. Specifically, this method requires filtering the color or intensity data of the image with the following filter kernels:

Dalal and Triggs tested other, more complex masks, such as the 3x3 Sobel mask or diagonal masks, but these masks generally performed more poorly in detecting humans in images. They also experimented with Gaussian smoothing before applying the derivative mask, but similarly found that omission of any smoothing performed better in practice.

Orientation Bining: The second step of calculation is creating the cell histograms. Each pixel within the cell casts a weighted vote for an orientation-based histogram channel based on the values found in the gradient computation. The cells themselves can either be rectangular or radial in shape, and the histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, depending on whether the gradient is “unsigned” or “signed”. Dalal and Triggs found that unsigned gradients used in conjunction with 9 histogram channels performed best in their human detection experiments. As for the vote weight, pixel contribution can either be the gradient magnitude itself, or some function of the magnitude. In tests, the gradient magnitude itself generally produces the best results. Other options for the vote weight could include the square root or square of the gradient magnitude, or some clipped version of the magnitude.

Descriptive Blocks: To account for changes in illumination and contrast, the gradient strengths must be locally normalized, which requires grouping the cells together into larger, spatially connected blocks. The HOG descriptor is then the concatenated vector of the components of the normalized cell histograms from all of the block regions. These blocks typically overlap, meaning that each cell contributes more than once to the final descriptor. Two main block geometries exist: rectangular R-HOG blocks and circular C-HOG blocks. R-HOG blocks are generally square grids, represented by three parameters: the number of cells per block, the number of pixels per cell, and the number of channels per cell histogram. C-HOG blocks can be described with four parameters: the number of angular and radial bins, the radius of the center bin, and the expansion factor for the radius of additional radial bins

Fig: Offline signature sample as biometric

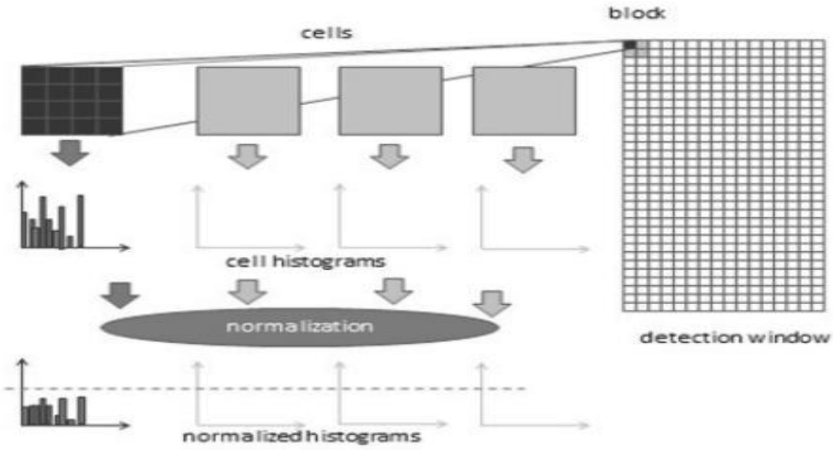

Fig : Demonstrate the HOG alogorithm

Object recognition: HOG descriptors may be used for object recognition by providing them as features to a machine learning algorithm. Dalal and Triggs used HOG descriptors as features in a support vector machine (SVM)[; however, HOG descriptors are not tied to a specific machine learning algorithm.

Note that while complex features give more information, simple features such as gradient orientation are more robust to normal variations found in a signature .

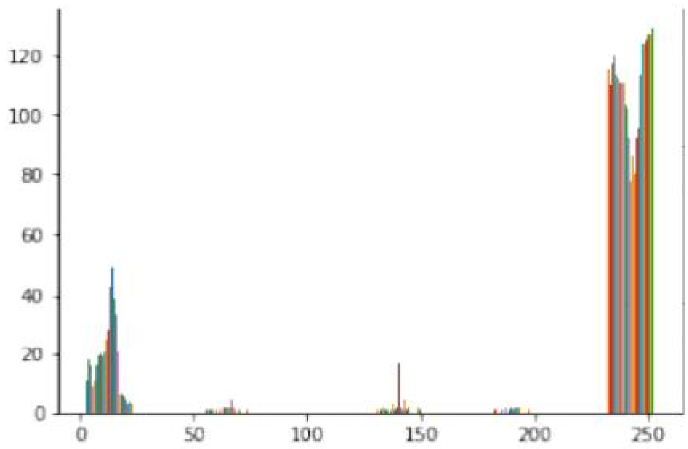

Fig: Histogram of Image pixel

5.3 ANN Training :

Artificial Neural Network or ANN resembles the human brain in learning through training and data storage.

The ANN is created and trained through a given input/ target data training pattern. During the learning process, the neural network output is compared with the target value and a network weight correction via a learning algorithm is performed in such a way to minimize an error function between the two values..

The mean-squared error (MSE) is a commonly used error function which tries to minimize the average error between the network's output and the target value.

Twelve exact signatures and twelve forged signatures train the network and they were enough to give very good results in verification.

Table contains all the information related to the design of the neural network. Both original and forgery signatures are used for training the network. Testing signatures are also available.

Parameter	**Value**
Number of layers	2
Number of neurons Output layer	1
Number of inputs	12
Learning rate	Default
Transfer function First layer	Sigmoid
Transfer function second layer	Sigmoid
Initial weights	Randomized
Initial biases	Randomized
Max number of epochs	1000
Error goal	0.0001
Number of patterns for original signature	12
Number of patterns for fake signature	12
Number of tested signatures	24
Number of tested original signatures	12
Number of tested fake signatures	12

Table : Neural Network Specification

Visualization of Result

Through the following interface, the user can select signature image of interest from available database.

Then train the network with the content of this database and 'matched ' or 'not matched' status which indicates if the signature is exact or forged.

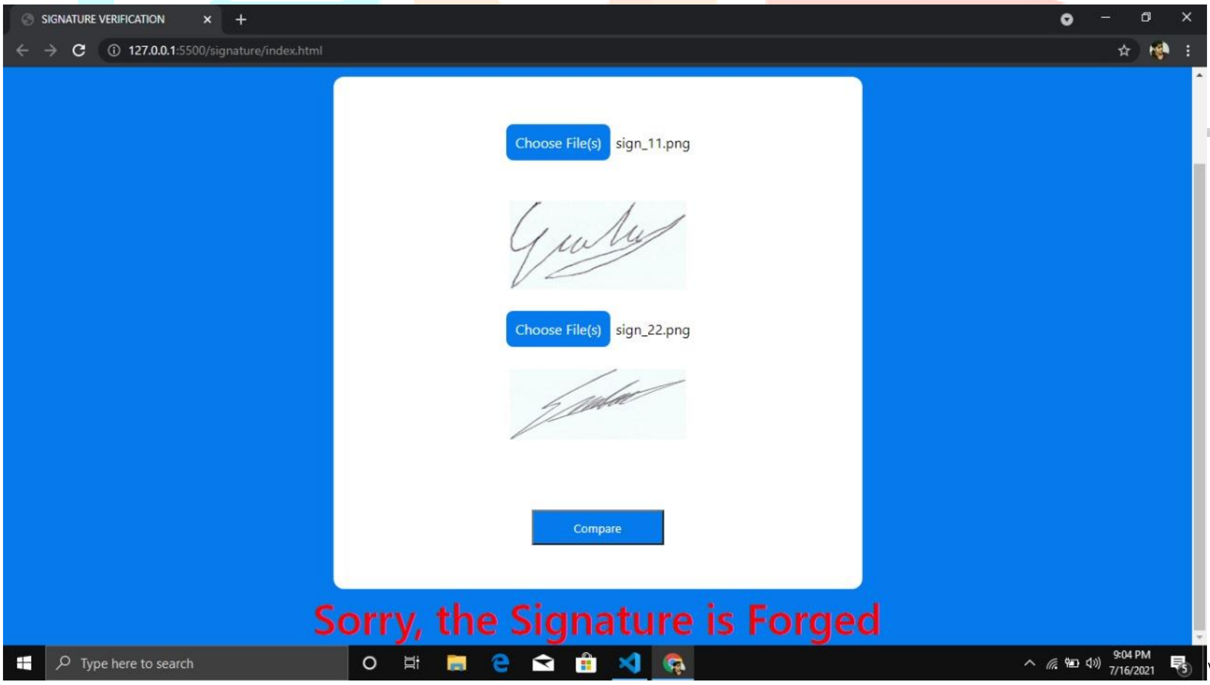

Fig: Output SCreen

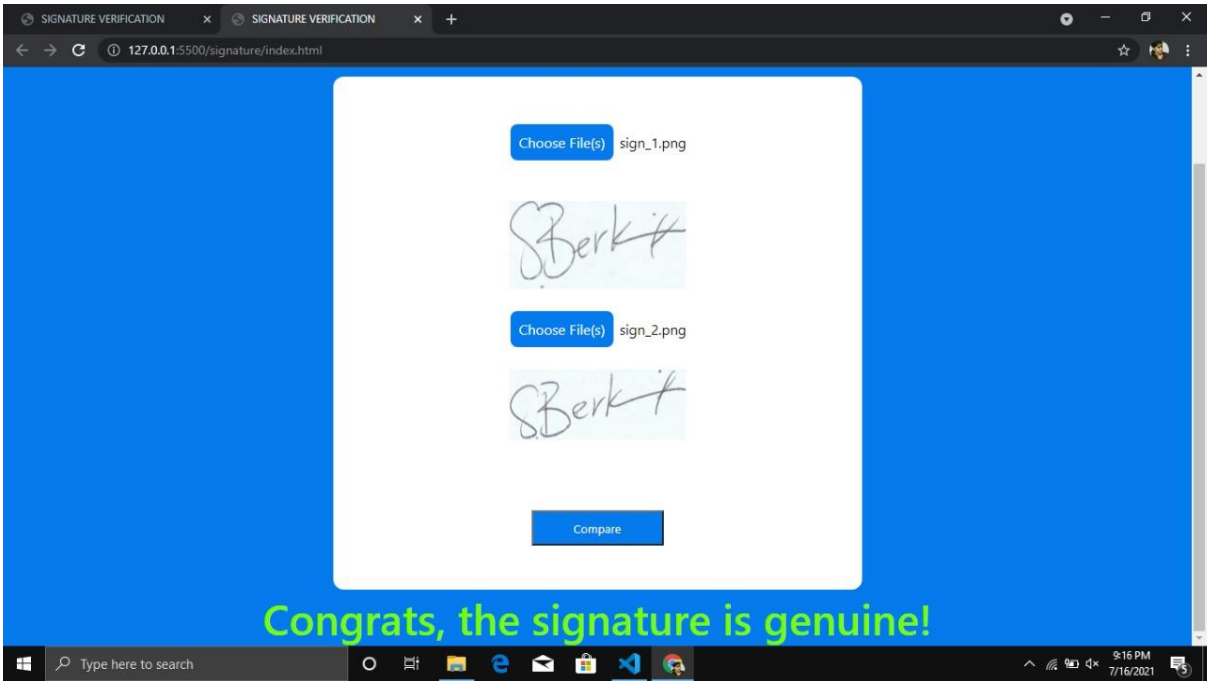

Fig : Output Screen

CONCLUSION

Neural networks have demonstrated their success in many applications due to their ability to solve some problems with relative ease of use and the model-free property they enjoy. One of the main features, which can be attributed to ANN, is its ability to learn nonlinear problem offline with selective training, which can lead to sufficiently accurate response.

Application of Artificial Neural Network (ANN) to the above mentioned problem has attained increasing importance mainly due to the efficiency of present day computers. In addition, the times of simulation and testing in the ANN application are minimal. And the verification system based on ANN is able to learn different kinds of signature datasets.

Moreover, the use of large data is not required to show the capability of learning for this sort of problem, we have chosen only twelve genuine signatures and twelve forged ones for training, and we get very good results. However for real practice use, larger training data can increase the robustness of the system.

After training, the best classification accuracies were achieved. The classification ratio exceeds 93%. The algorithm we supported uses simple geometric features to characterize signatures that effectively serve to classify signature as exact or forged. The system is robust and can detect random, simple and semi-skilled forgeries. We have no clear idea about its performance in case of very skilled forgeries because we are not skillful imitating signatures to the extent being considered as skilled forgeries.

FUTURE WORK

In terms of feature work , this project will be implemented and in a futher implementation stage will opt for better training and verification method to improve the accuracy of the offline signature verification system

REFERENCES

[1] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01, CVPR '05, pages 886–893, Washington, DC, USA, 2005. IEEE Computer Society.

[2] T. Ojala, M. Pietikainen, and D. Harwood. Performance evaluation of texture measures with classification based on kullback discrimination of distributions. pages A:582–585, 1994.

[3] J. F. Vargas, M. A. Ferrer, C. M. Travieso, and J. B. Alonso. Off-line signature verification based on grey level information using texture features. Pattern Recogn., 44:375–385, February 2011.

[4] M. A. Ferrer, J. B. Alonso, and C. M. Travieso. Offline geometric parameters for automatic signature verification using fixed point arithmetic. IEEE Trans. Pattern Anal. Mach. Intell., 27(6):993–997, 2005.

[5] C. M. T. J. B. Francisco Vargas, Miguel A. Ferrer. Off-line handwritten signature gpds-960 corpus. In IAPR 9'th International Conference on Document Analysis and Recognition, pages 764–768, September 2007.

[6] J. K. Guo. Forgery detection by local correspondence. PhD thesis, College Park, MD, USA, 2000. Director-Rosenfeld, Azriel.

[7] Wikipedia. https://en.wikipedia.org/wiki/Artificial_neural_network

[8] Alexander Wong, Stochastic image denoising based on Markov-chain Monte Carlo sampling

[9] "Histograms of Oriented Gradients for Human Detection" ,http://lear.inrialpes.fr/people/triggs/pubs/Dalalcvpr05.pdf

[10] M. Trikha, M. Singhal, and M. Dutta, “Signature Verification using Normalized Static Features and Neural Network Classification,” International Journal of Electrical and Computer Engineering, vol. 6, no. 6, pp. 2665, 2016.

[11] M. Singhal, M. Trikha, and M. Dutta, “Time Independent Signature Verification using Normalized Weighted Coefficients,” International Journal of Electrical and Computer Engineering, vol. 6, no. 6, pp. 2658, 2016.