AN ENHANCED SIGNATURE VERIFICATION SYSTEM USING KNN

S.Kuppusamy[1], M.Bala Krishnan[2], S.Mukthair Basha[3], M.Indhumathy[4]

[1,2,3]Dept. of Information Technology, Rajiv Gandhi College of Engineering and Technology, Puducherry, India.
[4]Assistant Professor, Dept. of Information Technology, Rajiv Gandhi College of Engineering and Technology, Puducherry, India

---***---

Abstract - *Handwritten signatures have proved to be important in authenticating a person's identity, who is signing the document. Here propose a system for signature verification using KNN. Nowadays signature is a basic and important verification system for every individual. Everyone has a unique signature and every individual can differ from others. An automated verification process would enable banks and other financial institutions to significantly reduce check and money order forgeries, which account for a large monetary loss each year. Simulation are carried out using Matlab.*

***Key Words: Handwritten, signature verification, KNN** …*

1.INTRODUCTION

Handwriting is a skill that is highly personal to individuals and consists of graphical marks on the surface in relation to a particular language. Many researchers have been done on this topic. Signatures of the same person can vary with time and state of mind. A method proposed a signature verification system which extracts certain dynamic features derived from velocity and acceleration of the pen together with other global parameters like total time taken, number of pen-ups. The features are modeled by fitting probability density functions i.e., by estimating the mean and variance, which could probably take care of the variations of the features of the signatures of the same person with respect to time and state of mind. Handwritten signature is a form of identification for a person A method is introduced where a signature image is first segmented (vertical and horizontal) and then data is extracted from individual blocks. Here these data is then compared with the test signature. Signatures are composed of special characters and flourishes and therefore most of the time they can be unreadable. Also intrapersonal variations and the differences make it necessary to analyze them as complete images and not as letters and words put together.

The handwritten signature is a particularly important type of biometric trait, mainly due to its ubiquitous use to verify a person's identity in legal, financial and administrative areas. One of the reasons for its widespread use is that the process to collect handwritten signatures is non-invasive, and people are familiar with the use of signatures in their daily life. Signature verification systems aim to automatically discriminate if the biometric sample is indeed of a claimed individual. In other words, they are used to classify query signatures as genuine or forgeries. Forgeries are commonly classified in three types: random, simple and skilled (or simulated) forgeries. In the case of random forgeries, the forger has no information about the user or his signature and uses his own signature instead. In this case, the forgery contains a different semantic meaning than the genuine signatures from the user, presenting a very different overall shape. In the case of simple forgeries, the forger has knowledge of the user's name, but not about the user's signature. In this case, the forgery may present more similarities to the genuine signature, in particular for users that sign with their full name, or part of it. In skilled forgeries, the forger has access for both the user's name and signature, and often practices imitating the user's signature. This result in forgeries that have higher resemblance to the genuine signature, and therefore are harder to detect. Depending on the acquisition method, signature verification systems are divided in two categories: online (dynamic) and offline (static). In the online case, an acquisition device, such as a digitizing table, is used to acquire the user's signature. The data is collected as a sequence over time, containing the position of the pen, and in some cases including additional information such as the pen inclination, pressure, etc. In offline signature verification, the signature is acquired after the writing process is completed. In this case, the signature is represented as a digital image.

FIG.1 SIGNATURE

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

2. RELATED WORK

This work proposes a novel system for off-line handwritten signature verification. A new descriptor founded on a quad-tree structure of the Histogram Of Templates (HOT) is introduced. For the verification step, we propose a robust implementation of the Artificial Immune Recognition System (AIRS). This classifier is inspired from the natural immune system, which generates antibodies to protect the human body against antigens. The AIRS training develops new memory cells that are subsequently used to recognize data through a k Nearest Neighbor (kNN) classification. Presently, to get a robust verification, the kNN classification is substituted by a Support Vector (SV) decision, yielding the AIRSV classifier [1]. Recognition of signature is a method of identification, whereas verification takes the decision about its genuineness. Though recognition and verification both play important role in forensic sciences, however, recognition is of special importance to the banking sectors. In this paper, we present a methodology to analyse 3D signatures captured using Leap motion sensor with the help of a new feature-set extracted using convex hull vertices enclosing the signature. We have used *k*-NN and HMM classifiers to classify signatures. Experiments carried out using our dataset as well as publicly available datasets reveal that the proposed feature-set can reduce the computational burden significantly as compared to existing features. It has been observed that a 10-fold computational gain can be achieved with non-noticeable loss in performance using the proposed feature-set as compared with the existing high-level features due to significant reduction in the feature vector size [2]. In this work a new online signature verification system based on Mellin transform in combination with an MFCC is presented. In the first step we extract signals x(t) and y(t) from each signature and then the novel pre-processing algorithm by Mellin transform is performed. The key property of Mellin transform is the scale invariance which makes the features insensitive to different signature scale. The feature is extracted by Mel Frequency Cepstral Coefficient (MFCC). Subsequently, feature extraction is used to extract coefficient for each signature to construct a feature vector. These vectors are then fed into two classifiers: Neural network with multi-layer perception architecture and linear classifier used in conjunction with PCA and then results are compared. In order to evaluate the effectiveness of the system several experiments are carried out. Online signature database from signature verification competition (SVC) 2004 is used during all of the tests [3]. A system which does computing and is combines with basic, and highly coincidental processing elements which use the data to get a highly relevant and faster response from the inputs taken. Artificial neural network models are a subpart of the machine learning models which are motivated by the functioning of the brain. Neural networks generally work like the neurons of the brain and the connected neurons will work in a network process to collect and process the data for providing the necessary output. There will be an input layer to the system which consists of all the patterns in which the system should process and also the necessary inputs and it communicates with the hidden layer as shown in the below figure and the hidden layers use the patterns and inputs by the input layer and are used to find out a relevant function for the task to be performed and then they communicate with the output layers to display the final output.

2.1 FEEDFORWARD MECHANISM:

This mechanism does not form circles like many artificial neural networks. This mechanism goes in a single way from the input to the hidden layers to the output and do not form any loops or circles in the process.

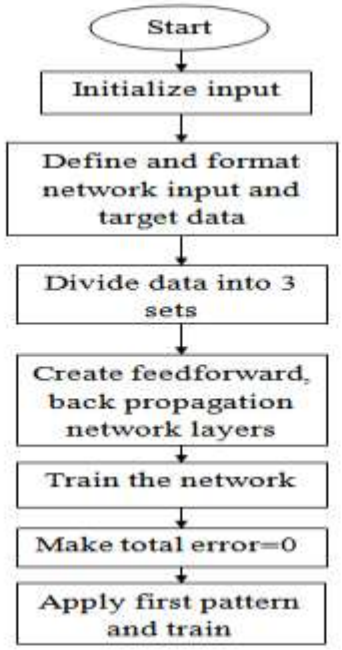

FIG.2 FLOWCHART

The important factor in preprocessing stage is to build standard signature which is prepared for extraction of features. Image processing application, pre-processing is required to remove discrepancies, from the input image.

2.2 NORMALIZATION OF SIGNATURE

It is possible that signature can be fractured due to imperfections in image scanning and capturing. It is also possible that the dimensions of signature can vary from person to person and even the same person can sometimes have different sizes based on the mood and environmental factors. So a process is required to overcome the size variation problem and achieve a standard signature size for all signatures.

2.3 THINNING

It is possible that the signature is written on different pen and the thickness thus varies from one pen to another. The purpose of thinning is to eliminate thickness differences in signature by making all of them one pixel thick. Thinning is used to enhance the object's global properties and to transform the input image into a compact form.

2.4 FEATURE EXTRACTION

This method is used for extracting the necessary and essential features from the input image. A feature vector is created from the features extracted. Each signature has a unique feature vector. These features are extracted as follows

The feature extraction module uses moment invariants to extract texture features of the image using central moment and derived invariant moment.

2.5 NEURAL NETWORK TRAINING

The features extracted are fed to natural network as inputs. Before that the networks are trained with data sets. Each neural network has a corresponding user to it. So a user has two neural networks one with feedforward mechanism and the other with feedback mechanism. The user's features are given as input to both the neural networks and the output is recorded.

In this method it construct a neural network by optimizing some existing neural networks and it will have a use the data structure tree along with nodes similar to human eye which has neurons and it used for recognition of patterns.

3. PROPOSED SYSTEM

The signature verification system using k-nearest neighbor is proposed. In pattern recognition, the k-nearest neighbor algorithm is a method for classifying objects based on closest training examples in the feature space. The intuition underlying nearest neighbour classification is quite straightforward, signatures are classified based on the class of their nearest neighbours.

3.1 KNN

K Nearest Neighbor(KNN) is a very simple, easy to understand, versatile and one of the topmost machine learning algorithms. In KNN, K is the number of nearest neighbors. The number of neighbors is the core deciding factor. K is generally an odd number if the number of classes is 2. When K=1, then the algorithm is known as the nearest neighbor algorithm

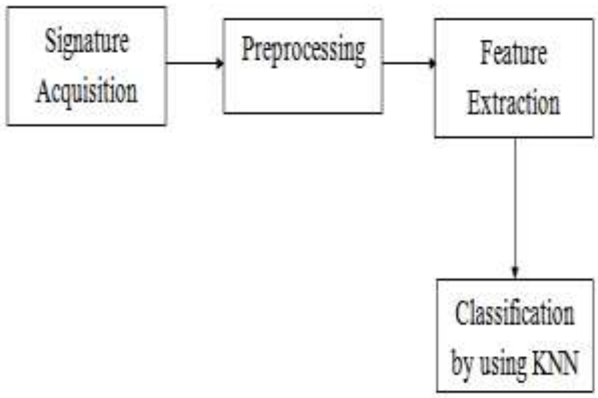

FIG.3 BLOCK DIAGRAM

3.2 IMAGE ACQUISITION

Here For every person we have collected n signature samples for database. It is better if we can collect more signature samples for database. Then for verification collect test signatures against the sample signatures. These test signatures we have to verify if it is genuine or forgery. Each of the signatures (Samples and corresponding test) has to take within a same sized area on paper by pen and collect the image of that particular area.

3.3 PREPROCESSING

The preprocessing of the signature images is related to the removal of noises, and thinning. The goal of thinning is to eliminate the thickness differences of pen by making the image one pixel thick. To remove noises and enhance, the images are preprocessed by filtering techniques. The principle objective of the image enhancement is to process an image for a specific task so that the processed image is better viewed than the original image. Preprocessing is process which helps us to reduce the background noise. Intensity of the image should be normalized. By Enhancing input image or image captured by digital camera, is to remove the background noise, image can get enhanced visual appearance of input images. By this enhancement process artifact image can be highlighted. Image preprocessing is used to create an enhanced and please full version of the captured image. The image preprocessing steps used in the system are:

1) Conversion of RGB image to gray image

2) Resizing of the image

3) Filtering of the image.

In RGB color model, each color appears in its primary spectral components of red, green, and blue. The color of a pixel is made up of three components; red, green, and blue (RGB). The disadvantages of RGB models are, it requires large space to store and it will take more time to process. So there is a need for converting the RGB model to Gray model. Resizing is an important step in image preprocessing. The acquired image is resized according to the requirement of the system. Resizing is nothing but, changing the dimensions of an image. The captured image is resized using some resizing methods according to the requirement of the system.

3.4 FEATURE EXTRACTION

Extracted features in this stage are used for clustering the signature images for verification stage. Features will have to be extract from both sample images and Test image.

3.5 SIGNATURE HEIGHT WIDTH RATIO:

The ratio is obtained by dividing signature height to signature width. The height is the maximum length of the columns obtained from the cropped image. Similarly the width is also calculated considering the row of maximum length. Signature height and width can change. But height-to-width ratios of an individual's signatures are approximately constant.

3.6 SIGNATURE OCCUPANCY RATIO:

It is the ratio of number of pixels which belong to the signature to the total pixels in the signature image. This feature provides information about the signature density.

3.7 DISTANCE RATIO CALCULATION AT BOUNDARY:

After cropping, the pixels in closest proximity to the boundaries (left, right, upper & bottom) are determined and their distance from the left & bottom boundaries are evaluated, i.e. for the upper leftmost pixel its distance from bottom boundary(L1) & for the bottom left most pixel the distance from right boundary is calculated(L2). These values are used later in verification process.

3.8 CLASSIFICATION

The classifier we have used is KNN which stands for k-nearest neighbours. It is basically a classification algorithm that means it assigns a class to a test image based on its feature values. A person may have many sample signature images. We create separate clusters for set of sample signatures for each person. Here we use K-Nearest Neighbors' (KNN) clustering Technique for verifying a test signature belongs which cluster. The k-nearest neighbours' algorithm uses Euclidian distance method to find the distance between two training points. Thus using Euclidian distance we find k nearest neighbouring training points of our test point based on its features and the class with maximum number of occurrences is taken as the decision class for that test image and is assigned to that image.

It is often useful to take more than one neighbour into account so the technique is more commonly referred to as k-nearest neighbour (k-NN) classification where k nearest neighbours are used in determining the class. Since the training signatures are needed at run-time, i.e. they need to be in memory at run-time; it is sometimes also called memory-based classification. Because induction is delayed to run time, it is considered a lazy learning technique. Because classification is based directly on the training signatures it is also called example-based classification or case-based classification. So, k–NN classification has two stages; the first is the determination of the nearest neighbours and the second is the determination of the class using those neighbours. This approach to classification is of particular importance today because issues of poor run-time performance is not such a problem these days with the computational power that is available. During the enrollment phase, a set of reference signatures are used to determine user dependent parameters characterizing the variance within the reference signatures. The reference set of signatures, together with these parameters, are stored with a unique user identifier in the system's database. In the training phase we choose a number of genuine and forged signatures for training the K-NN classifier. In the verification phase when a test signature is input to the system, it is compared to each of the reference signatures of the claimed person. The person is authenticated if the resulting dissimilarity measure is below or equals a threshold value of the classifier, otherwise denied.

4. IMPLEMENTATION

In the project work, the experiments are carried using Matlab coding. We have prepared a GUI layout with a list of menus. Clicking on each menu will perform an independent function.

FIG.4 OUTPUT GUI

FIG.5 MOST LIKELY GENUINE

FIG.6 DEFINITELY FORGED

5. CONCLUSION

In this project, we used Image processing which is one of the most trending and most used domain nowadays for functions like image detection, fingerprint verification etc. This project helps in controlling human errors in signature verification and also makes the signature verification accurate, easy and faster. It also makes the work easier for understanding and executing it by anyone without any knowledge of image processing. If any bank or any company uses this system the customers will feel much more secure and trustworthy. Thus, here propose that this system brings a change in the working of several banks, companies etc. It is better if future works extracts more features that may provide a combination to achieve higher accuracy. Future works should include the use of different features and classifiers such as deep learning neural network. By increasing the numbers of hidden layers, the performance of the neural network can be expected to be better but time for training and testing may increase.

REFERENCES

1. S. Yin, A. Jin, Y. Han, and B. Yan, “Image-based handwritten signature verification using hybrid methods of discrete Radon transform , principal component analysis and probabilistic neural network,” Appl. Soft Comput. J., vol. 40, pp. 274–282, 2016.
2. K. Wrobel, R. Doroz, P. Porwik, J. Naruniec, and M. Kowalski, “Engineering Applications of Artificial Intelligence Using a Probabilistic Neural Network for lip-based biometric verification,” Eng. Appl. Artif. Intell., vol. 64, no. January, pp. 112–127, 2017.
3. D. Suryani, E. Irwansyah, R. Chindra, D. Suryani, E. Irwansyah, and R. Chindra, “ScienceDirect O fflfflineine Signature Signature Recognition Recognition and and Verification Verification System System using usingfficient Fuzzy Kohonen Clustering Network (EFKCN) Algorithm E fficient Fuzzy Kohonen Clustering Network (EFKCN) Algorithm,” Procedia Comput. Sci., vol. 116, pp. 621–628, 2017.
4. Y. Serdouk, H. Nemmour, and Y. Chibani, “New off-line Handwritten Signature Verification method based on Artificial Immune Recognition System,” Expert Syst. Appl., vol. 51, pp. 186–194, 2016.
5. Y. Serdouk, H. Nemmour, and Y. Chibani, “Handwritten signature verification using the quad-tree histogram of templates and a Support Vector-based artificial immune classification ☆,” Image Vis. Comput., vol. 66, pp. 26–35, 2017.
6. P. Porwik, R. Doroz, and T. Orczyk, “Signatures veri fi cation based on PNN classi fi eroptimised by PSO algorithm,” vol. 60, pp. 998–1014, 2016.
7. S. Kumar, D. Prosad, and P. Pratim, “Fast recognition and verification of 3D air signatures using convex hulls,” Expert Syst. Appl., vol. 100, pp. 106–119, 2018.

8. N. Khera and S. A. Khan, “Microelectronics Reliability Prognostics of aluminum electrolytic capacitors using arti fi cial neural network approach,” Microelectron. Reliab., vol. 81, no. October 2017, pp. 328–336, 2018.
9. A. Fallah, M. Jamaati, and A. Soleamani, “A new online signature verification system based on combining Mellintransform , MFCC and neural network,” Digit. Signal Process., vol. 21, no. 2, pp. 404–416, 2011.
10. D. Dabrowski, “Condition monitoring of planetary gearbox by hardware implementation of artificial neural networks,” Measurement, vol. 91, pp. 295–308, 2016.
[11] Vance Faber , “Clustering And the Continuous K-means Algorithm”, Los Alamos Science, Number 22 , 1994

[12] P.S.Bradley, Usama M.Fayyad, “Refining Initial Points for K-Means Clustering”, International Conference on Machine Learning (ICML98), J. Shavlik (ed.), pp. 91-99, 1998

[13] Dorin Comaniciu, Peter Meer, " Mean Shift: A Robust Approach Toward Feature Space Analysis ", IEEE transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, May, 2002

[14] Pierre-Alain Mo[illegible]ic, Jean -Emmanuel Haugeard, Guillaume Pittel, "Image Clustering Based on a Shared Nearest NeighborsApproach for Tagged Collections",Proceedings of the 2008 international conference on Content-based image and video retrieval, pp:269-278 ISBN:978-1-60558-070-8 , 2008

[15] Ujjwal Maulik, Sanghamitra Bandyopadhyay, "Performance Evaluation of Some Clustering Algorithms and Validity Indices", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 12, DECEMBER 2002

[16] H. Liu1, J. Li and M. A. Chapman, "Automated Road Extraction from Satellite Imagery Using Hybrid Genetic Algorithms and Cluster Analysis", Journal of Environmental Informatics 1 (2) 40-47, 2003

[17] M.Liwicki and H.Bunke, "Combining Online and off-line Systems for Handwriting Recognition", 9th International conference on Document Analysis and Recognition (ICDAR 2007) , Curitiba, Brazil, Vol-I, pp:372-376

[18] Vu Nguyen, Michael Blumenstein, Vallipuram Muthukkumarasamy,Graham Leedham, "Offline Signature Verification Using Enhanced Modified Direction Features in Conjunction with Neural Classifiers and Support Vector Machines", 9th International conference on Document Analysis and Recognition (ICDAR 2007) , Curitiba, Brazil, Vol-II, pp:734-738.