Offline Signature Verification Using Neural Network

Dr. Kiran Y.C[1], Ms. Nirmita Nagaraj[2]

[1]Professor, Dept. of Computer Science and engineering, B.N.M Institute of Technology, Karnataka, India
[2]Student, Dept. of Computer Science and engineering, B.N.M Institute of Technology, Karnataka, India

---***---

Abstract - *— Signature is a distinct or a characteristic of a person's name as a source of identification. Signature is required to approve, accept or to oblige anything by the individual on the document. Hence it considered as one of the Biometric. These days, signatures are pruned to forgeries hence, an effective offline signature verification using neural network as a classifier is proposed in this paper. The test signature is identified first to see who the signer is and then the test signature undergoes the verification process to check if it is a forged or a genuine signature. Once the image is uploaded into the system, it undergoes pre-processing and feature extraction. The trained result is sent to the neural classifier along with the extracted features and the verification process is done to obtain the final result. The output is shown in the form of matched percentage. The performance of the model shows that it works efficiently with fewer number of images.*

Key Words*:* **Biometric, offline signature verification, neural network, test signature, trained result.**

1. INTRODUCTION

One of the distinguishing feature for a person's identification through many ages has been signature. Signature is one of the Biometric for person's identity. A person's name or a simple short form of the name can be a handwritten depiction that acts as a proof of identification. One of the accepted form of conformation, be it in the form of transactions, document verification, contracts in civil law, voting or in authenticating one's identity all of them are authorized via signature. If authenticity or verification has to be done on a regular basis, then it has to happen automatically. Hence automatic signature verification comes into picture. This fastens the speed of verification automatically. One of the best application of signature is financial application and border control. Handwritten signatures are considered as a proof of identity for legal documents such as bank cheques, legal property documents, etc. Signature is considered as one of the behavioral form of human characteristic.

Signature identification and verification are well established and is efficient. Biometrics are widely used these days as it is considered as one of the safest and most unique form of recognizing a person's identity. Machine Learning is one of the paradigm that is used in Biometrics. Biometrics is considered as one of the traditional form of an individual's authentication. Traditional methods are becoming more obsolete these days and are no longer used. ID cards and passwords are gradually taken away by Biometrics in this generation. More the metrics given for identification greater is the accuracy. In this case, if the characteristic's given for verification is multimodal, then it is almost nearly impossible to forge. It provides an identification that is unique in nature. This way of identifying humans by their traits is referred to as Biometric Authentication. In general, Biometrics are automated methods of recognizing a person based on physiological or behavioral characteristics. Biometric management systems consist of hardwares and softwares which is not complicated to analyze and install. The process of installation is not time consuming rather it can be easily installed.

For years, handwritten signatures have been a potential way of accessing one's identity. The identification is made at ease to any individual working in this approach. Handwritten signatures are proved to be more evident than any other form of Biometric. The person signing the document is called a Signer or a Signatory. Biometrics is widely used in security applications. Since technology is increasing these days, so is the threat in accessing an individual's identity is increasing. Hence the security measure in identification and verification is adopted. Ever since decades, this has been an open research problem. Various techniques and methods have been adopted to resolve this problem. Signatures are in the form of two types:

- Physiological
- Behavioral

Physiological signatures are in the form of iris, cortex, thumb impression, face recognition. The physical feature is directly captured in this form. Behavioral signatures are Biometrics in the form of handwritten signatures, voice recognition. Here two processes or two main steps has to be considered. One is the identification process and the other is the verification process. In the identification process the signatures are grouped into their particular type or class. The process of identifying and grouping is a major concern here. In the verification process, the identified signature comes down to a level of acceptance or rejection. A person's identity is based on these physiological and behavioral traits. Handwritten signatures are one of the mostly accepted form of Biometrics.

2. RELATED WORK

Handwritten signatures have been used widely hence, automated verification systems were adopted. Although nothing can replace the human eye, there is a need to improve on the performance of the automated system, when compared to the human eye. With very less features, online approach is becoming more challenging these days. In [1], to tackle the challenges and to boost their discrimination, a new method that makes use of KAZE features is adopted. The KAZE feature is purely based on Fisher Vector (FV) encoding. A visual vocabulary and statistics of a higher order is proposed. Both of these features can encode KAZE features, thus provides a better spatial distribution. The dataset used here is the MCYT-75 dataset. This method improves the performance compared to the recent Vector of Locally Aggregated Descriptors approach (VLAD). The use of component analysis to the Fisher Vector does not decrease the performance. By using the FV method, the error rate is lower when compared to the state-of-the-art verification systems. Hence handwritten signatures are meant for identity authentication because it deals with characteristics of the human traits.

In [2], the evaluation of the signature verification process had been evaluated by using a protocol wherein two scenarios that are independent in nature are considered, that is forgeries that are random and skilled forgeries. Such an approach might be necessary, this can cause misinterpretation of the analysis of the assessment process. In realistic such a huge separation between these two types of imposters is probably not real. One of the most traditional form of method is the "ink and paper" method that is the electronic way of signing. Since it is a biometric characteristic, it is a form wherein the user learns to produce. One of the main disadvantage is that, the offline signatures are pruned to forgeries than the other form of biometrics like fingerprint, face recognition, iris, etc.

The scenario or the instances that are distinguishable when skilled imposters comes into picture are firstly, the name of the user or the subject that is known. Secondly, an instance of the image alone of the signature is known. Thirdly, both the dynamic feature and the image of the signature representation is known. Here, an evaluation protocol is used wherein, if the level of knowledge used is less then, there will be no difference among the signatures made by the skilled imposters. The random scenario and the skilled scenario is treated as two independent cases. The performance of the system is independent of each other. Threshold can be treated separately for these scenarios. For each imposter class, the threshold selected can be different. Signature and its type may not be known or it is not possible for one to know its type under any circumstances, so that it can be used in the system.

As a result, the optimization of this scenario may not be specific, hence the feasibility also decreases. In [3], mimicry is a form of presentation attack in biometric form of signature verification. The skilled imposters and the presentation attacks are almost the same. Further at the sensor category the attack is generated. This is also because, the output that is expected is a hard assignment and the statistics considered is the zeroth or the first order statistics. To reduce the complexity a feature extraction process and component analysis is used to the original Fisher Vector.

For improving the authentication techniques many forms of offline signature verification process are introduced. In [4], the keyword that is described and put up is a Histogram. Histogram makes use of gradients in order to achieve the specified signature. The first stage of the literature experimentation includes the collection of signature image. The image is in the form of a binary format, represented in binary bits. The next stage is to compute the noise and the effect of noise that is being produced and evaluated. This effect is enhanced and further decreased.

The size of the signature image is the key aspect that is to be considered, therefore the image is cropped and is resized to fit the requirements of the user. In [5], one of the most natural form of signature is the handwritten signature. It is considered as an identity and a recognition of a person's well-being. The procedure that is inherent in a network can be processed with the help of a verification technique. The features that are extracted in this mechanism is of two types that is, static approach and a dynamic approach. This is mechanized in order to train the network. The topologies are therefore tested and the resulting system is categorized and an error rate is considered which sums up to a value of three-point five percent and is proved as the best case of this characteristic.

3. PROPOSED MODEL

The distinguishing feature for one's identification is signature. This was declared in the nineteenth century by the British government for legalizing documents. The main aim is to find a clear solution which is feasible to verify handwritten or offline signatures. There is more scope for offline signatures because of its static inputs and outputs. The main classifier used here is Neural Network Classifier. The security that is provided is a biometric approach. It is a process of verifying an individual's identity. Fingerprint is one of the oldest and an efficient technique that falls under biometric recognition.

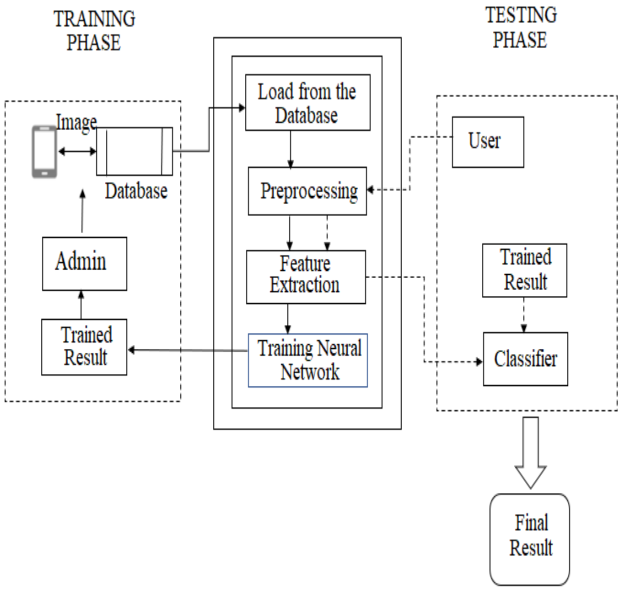


Fig -1: Proposed Architecture

The image enters the training phase at the initial portion of the module and then trespasses through the testing phase. In training phase, a fixed and a resolved set of software's are used to train the image. Trained result generated from the network is performed and processed in the training portion. The trained result is sent to the admin and the image is stored with the help of a database. Admin is responsible for the actions performed in the training unit.

Testing phase makes use of a classifier to divide the information present in the image into a set of classes. Here the result that is being trained is classified and the final result is obtained which is the key of the chapter. Testing phase performs more computations compared to the training phase. The system verification unit has five stages.

- Acquiring the data
- Preprocessing of the data
- Extracting the features from the data
- Classifying
- Verification

Database plays a key role, which is considered as the input of the model, the image is taken from the database and is loaded into the database and passed on to the next stage, called the preprocessing stage. In this stage, tools are used to process the image and the processed image is pushed to the next segment called the feature extraction section, where a set of features are extracted with the help of an algorithm. The unit comprises of the architecture that is proposed for the implementation of the project, which is the first step for starting the project. It is studied as the first module of the chart.

The second aspect that arises is the requirements of the sample. There are two types of requirements namely software requirements and hardware requirements. Hardware requirements are the necessary inputs taken from the system, whereas the software manageable requirements are tools used for the operation.

The flow diagrams used are decomposed into many levels, which is the third aspect mentioned in figure 2. Each level is segregated and placed in positions to describe the flow of data that passes from one module to another module targeting on the output, which defines the detailed description of the sub models that is taken into consideration, which gives a clear picture on both the input and output of the setup. New approaches are required to be built if the testing systems do not return the same answers. This is true when a technique called software testing is adapted which defines the test cases of any project. Further it can also respond to changes from the previous transactions.

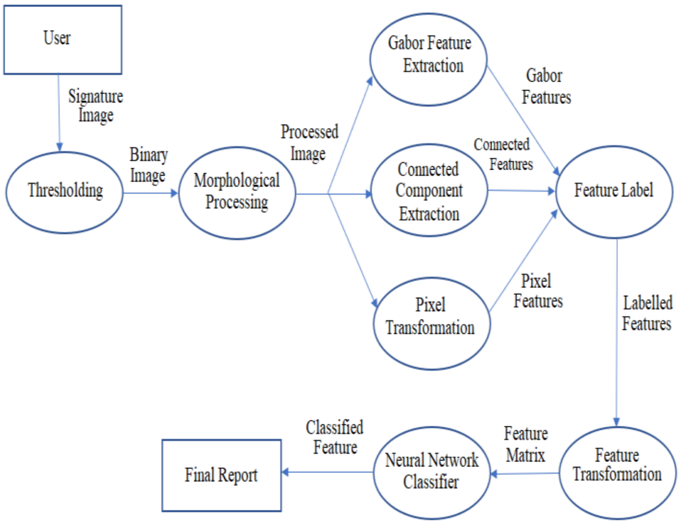


Fig -2: Flow Diagram

The previous transactions can be based on any type of scenario. Scenarios can be critical that is, once the architecture is developed for a particular test case, fresh inputs with new data has to be chosen to verify the accuracy of the output. Level of confidence in Machine Learning can be measured in statistical terms. If the model is not able to achieve its goals or to attain revenue, then defects can be easily reflected.

4. IMPLEMENTATION

After loading the signature, the signature verification system identifies whose signature it is. The next step is to identify if the signature is genuine or forged. For this, the system goes through a series of steps. The first phase is the preprocessing phase. In this, the quality of the image is improved by using the thresholding and morphological techniques. In thresholding, the image is partitioned into foreground and background. During the analysis, the image is segmented and the process of converting a grayscale image into is binary image takes place. The next phase is the feature extraction phase. The essential features of the image for verification are extracted here. It is possible by using approaches like Gabor Filter Approach, Connected Component Approach and Pixel Transformation. The extracted features are labelled and is sent to the neural classifier for classification. All the tasks are wrapped up in this section, and the features are easily examined in a straightforward approach. Irrelevant features are discarded to avoid misconception. There is no need of training or re-training the feature as long as the mechanism holds good. Therefore, it can be contributed to any aspect and used in the field of classification.

Similarities can occur between each feature used. Mechanism is hence called a genuine mechanism. Detection of the image can be called as an intermediate difficulty measure of the application. The size of the image can be defined from the schemes and setups used. The output can outperform all the state of the solutions used in the industry of images. Images are discovered at every stage of the module.

Algorithms used here are:

- Back Propagation
- Feed Forward Neural Network
- Surf Algorithm

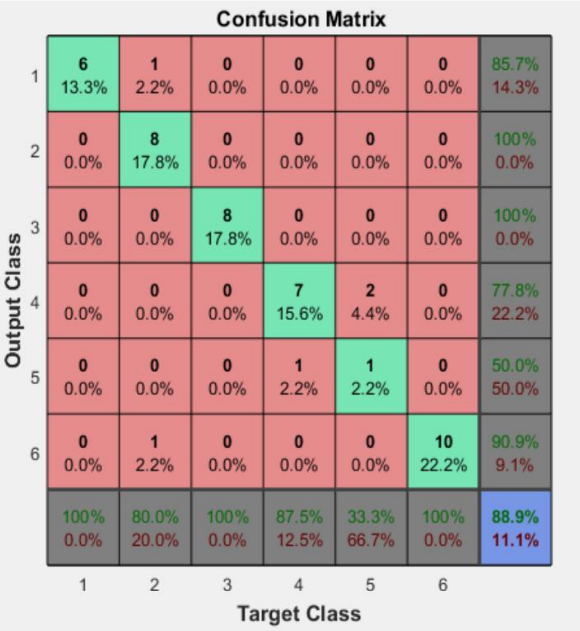
This phase utilizes an algorithm for its implementation and the algorithm used is a back-propagation algorithm, which facilitates the use of a modified network model and has the authority to perform actions and tasks that occur in any aspect of the model.

Algorithm describes the detailed information on the back-propagating mechanism, the coefficient vector is high in the algorithm. Trajectory of the algorithm goes beyond the context of data available and used in the present technology. Network that is being employed is a neural based network which provides an easy platform for computing the inner details. Algorithm can also be represented in a pictorial fashion by reducing the complexities. The framework of the algorithm can be put in five forms.

- Input unit
- Output unit
- Instance
- Hidden unit
- Weighted unit

The input unit is the initial stage of the network, whereas the output unit is the final stage of the pattern. Instances are moments that are captured during the processing of the algorithm. Preprocessing of the job can also be done by taking sufficient image tools. Image tools can vary based on the type of individual signing on a sheet of white paper. Hidden unit is the major unit which is not visible to the user that is only the admin can view the image that is being processed in the preprocessing phase. Weighted unit holds the tasks and information that is required to be processed before initialization of the textual image.

Weights are pushed into the network for its manipulation and each weight can be updated to get a proper hold in the network.



Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	
1	6 13.3%	1 2.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	85.7% 14.3%
2	0 0.0%	8 17.8%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
3	0 0.0%	0 0.0%	8 17.8%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
4	0 0.0%	0 0.0%	0 0.0%	7 15.6%	2 4.4%	0 0.0%	77.8% 22.2%
5	0 0.0%	0 0.0%	0 0.0%	1 2.2%	1 2.2%	0 0.0%	50.0% 50.0%
6	0 0.0%	1 2.2%	0 0.0%	0 0.0%	0 0.0%	10 22.2%	90.9% 9.1%
	100% 0.0%	80.0% 20.0%	100% 0.0%	87.5% 12.5%	33.3% 66.7%	100% 0.0%	88.9% 11.1%

Fig -4: Confusion Matrix

The output can also be obtained in the form of confusion matrix as shown in figure 4. The confusion matrix in Image Processing is a matrix wherein the performance of the test images can be known with true values of the test data. This table works on the classifier to obtain the output in the form of percentage. New classes can be added to the system and the number of forgeries that happen can be detected easily. A signature is the greatest asset of an individual. Signatures can also be represented in a textual manner, the complexity in a textual signature is very high and in major terms compared to a normal signature signed by the user. The length and width of the signature should also be examined to achieve the best, highest accuracy in the output. Image dataset with offline features are taken as inputs. Pre-processed Image sets of signature data of 'N' person with 'K' sets are loaded to extract offline features. Features with labels wherein each person 'N' is labeled from 0 to total number of person to obtain the feature matrix. The important step is to calculate the error content that is being produced in the network. Error content can be based on two types of errors that can occur in each portion of the networking element. Therefore, there are two errors namely

- Type I error
- Type II error

Both the errors can be constant at any given point of time, but the most used error in the algorithm is Type I error.

5. RESULTS

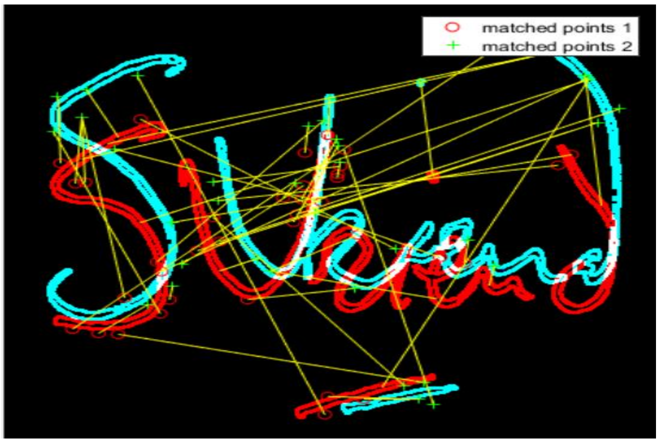


Fig -5: Matched Points of the Test Image

The above figure describes the output of the module, wherein an image is tested and a set of matched points are generated. The matched points are further computed and shown in the form of a graph.

Figure 6 defines the Performance Metrix, wherein the x-axis is counted as Epochs and y-axis is the cross entropy value.

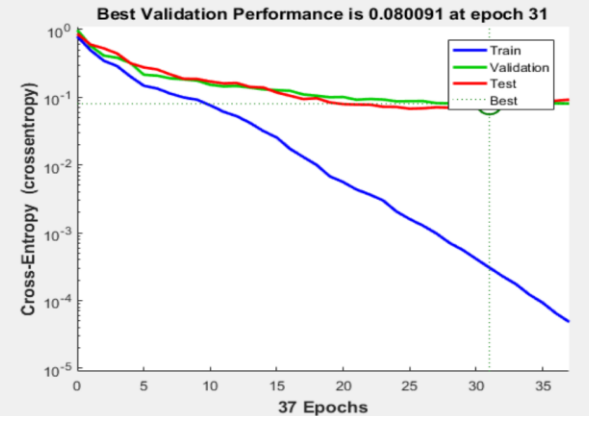


Fig -6: Performance Metrix Graph

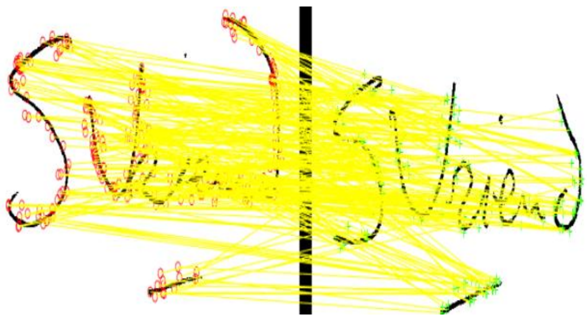


Fig -7: Result of the Surf approach

CONCLUSION AND FUTURE WORK

The reason for a Biometrics' diverse use is that, the process is non-invasive and the user is quite familiar with the use of their signatures in daily life. Features can also be sensitive based on the characteristics of the dataset. There can also be a comparison managed in a fair approach and the comparative analysis can be difficult in the future methods. Signature verification method can be more accurate, efficient and provide best results in percentage ratio compared to any state-of-the-art solution. The advantage of this approach is that every individual have their own model designed for them. Further implementation can also be done by taking different set of datasets so that it can works well with fewer number of datasets to make it more accurate.

REFERENCES

[1] Manabu Okawa "KAZE Features via Fisher Vector Encoding for Offline Signature Verification "IEEE 2017 IEEE International Joint Conference on Biometrics (IJCB).

[2] Nassim Abbas, Youcef Chibani, Bilal Hadjadji and Zayen Azzouz Omar "A DSmT Based System for Writer-Independent Handwritten Signature Verification"19th International Conference on Information Fusion Heidelberg, Germany - July 5-8, 2016.

[3] Vu Nguyen, Yumiko Kawazoe, Tetsushi Wakabayashi, Umapada Pal, Michael Blumenstein "Performance Analysis of the Gradient Feature and the Modified Direction Feature for Off-line Signature Verification "DOI 10.1109/ICFHR.2010.53.

[4] Latifa Nabila Harfiya, Agus Wahyu Widodo, Randy Cahya Wihandika, "Offline Signature Verification Based on Pyramid Histogram of Oriented Gradient Features" 2017 1st International Conference on Informatics and Computational Sciences (ICICoS).

[5] V. Nguyen, Y. Kawazoe, T. Wakabayashi, U. Pal "Performance analysis of the gradient and modified direction feature for off-line signature verification," in Proc. Int. Conf. Frontiers Handwriting Recognit., Nov. 2016, pp. 303–307.