Global Journal of Computer Science and Technology: G
Interdisciplinary
Volume 19 Issue 2 Version 1.0 Year 2019
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

An Efficient Automated Attendance Entering System by Eliminating Counterfeit Signatures using Kolmogorov Smirnov Test

By K.M.L.P. Weerasinghe & B.H. Sudantha

University of Moratuwa

Abstract- Maintaining the attendance database of thousands of students has become a tedious task in the universities in Sri Lanka. This paper comprises of 3 phases: signature extraction, signature recognition, and signature verification to automate the process. We applied necessary image processing techniques, and extracted useful features from each signature. Support Vector Machine (SVM), multiclass Support Vector Machine and Kolmogorov Smirnov test is used to signature classification, recognition, and verification respectively. The described method in this report represents an effective and accurate approach to automatic signature recognition and verification. It is capable of matching, classifying, and verifying the test signatures with the database of 83.33%, 100%, and 100% accuracy respectively.

Keywords: image processing, kolmogorov smirnov test, machine learning, support vector machine.

GJCST-G Classification: K.3.0

ANEFFICIENTAUTOMATEDATTENDANCEENTERINGSYSTEMBYELIMINATINGCOUNTERFEITSIGNATURESUSINGKOLMOGOROVSMIRNOVTEST

An Efficient Automated Attendance Entering System by Eliminating Counterfeit Signatures using Kolmogorov Smirnov Test

K.M.L.P. Weerasinghe [α] & B.H. Sudantha [σ]

Abstract- Maintaining the attendance database of thousands of students has become a tedious task in the universities in Sri Lanka. This paper comprises of 3 phases: signature extraction, signature recognition, and signature verification to automate the process. We applied necessary image processing techniques, and extracted useful features from each signature. Support Vector Machine (SVM), multiclass Support Vector Machine and Kolmogorov Smirnov test is used to signature classification, recognition, and verification respectively. The described method in this report represents an effective and accurate approach to automatic signature recognition and verification. It is capable of matching, classifying, and verifying the test signatures with the database of 83.33%, 100%, and 100% accuracy respectively.

Keywords: image processing, kolmogorov smirnov test, machine learning, support vector machine.

I. Introduction

Attendance records are very essential in the academic activities of universities. Almost all the universities in Sri Lanka, signatures of candidates are taken in lectures, practical sessions, during examinations, etc. to verify the presence of the real candidate.

The paper-based attendance sheet is passed in each session to take the signature of each student.

For later evaluations, attendance records should enter to the excel sheets and entering student's attendance into the excel sheets for each of the subjects which are very crucial, time-consuming process.

Automated student's attendance entering system can be used to simplify this task. Many attempts are made to automate this process with success to a certain extent. Many of these systems make use of sophisticated biometric equipment while some others use Barcodes and Radio Frequency Identity Cards [5]. Handwritten signatures considered the most natural method of authenticating a person's identity. However they are handled as images and recognized using computer vision and machine learning techniques. With modern computers, it is needed to develop fast algorithms for signature extraction, recognition, and verification. But even today the majorly used system is to take the signature of present candidates and then manually enter these records into the computer. In this study, the process has automated by developing a system which uses image processing techniques to update the attendance records automatically. To build up such a system signature extraction, recognition and verification are essential. Another task is the identification of counterfeit signatures. If we count the number of signatures and the number of heads during a lecture, practical session or in examinations, they should be the same. But sometimes some students sign for their colleagues or replaced by other students. Therefore, identification of counterfeit signatures is very much essential in this type of situations.

The handwritten signature is a prevalent way of authenticity. Despite its known weaknesses, and development of cryptographic and biometric techniques, it is still the most commonly used way of authentication when dealing with paper documents and forms. In this thesis, we focus on the application of biometric recognition for automatic student authentication, in particular making use of handwritten signatures, which are one of the most socially accepted biometric traits.

In education, signatures are used for attendance control, either to lectures or exams, but not for authentication. With the rapid deployment of dynamic signature extraction, recognition and verification automated students attendance entering system has been used for student authentication. Also, the use of this technology can be extended to different administrative services within the education system to add a higher security level to the traditional procedures of authentication.

Signature matching has been used in areas such as extraction [4], recognition [16] and verification [7]. While signature extraction aims to find document images that contain signatures [4], and signature recognition tries to find the corresponding signer of a test sample given a database of signature exemplars from different signers [2], signature verification deals with confirming the authenticity of a signature, i.e., decides whether a sample signature is genuine or forgery by comparing it with stored reference signatures. From the viewpoint of automating the attendance entering system, it involves machine learning from a

Authorα: Department of Information Technology Sri Lanka Institute of Information Technology, Malabe, Sri Lanka. e-mail: lokesha.w@sliit.lk
Authorσ: Department of Information Technology University of Moratuwa, Sri Lanka. e-mail: sudanthbh@uom.lk

population of signatures. In this study person, dependent learning is used in signature verification phase, so that there are only genuine signatures in the database. This technique is called special-learning. In special-learning, a person's signature is learnt from multiple samples of only that person's signature, where within-person similarities are learnt to identify the signature is genuine or counterfeit.

The rest of this article is organized as follows. Section II mentioned the current achievements in this domain. Section III gives a general description of the proposed method. In section IV, discuss the results and discussions. In section V, provides the conclusions.

II. Literature Review

Ritesh Banka [1] has presented a new approach for the extraction of signature and handwritten regions from official binary document images. He proposed a new two-level scale invariant classification technique to extract the gray-scale handwritten area from the scanned document.

Ogul and coworkers [4] described a discriminative framework to extract the signature from a bank service application document. This is based on the classification of segmented image regions using a set of representative features. The segmentation is done using a two-phase connected component labeling approach. Then evaluate solely and combined effects of several feature representation schemes in distinguishing signature and non-signature segments over a Support Vector Machine classifier.

Gupta [6] has done a cursive signature extraction and verification. In his research, he presented a new approach, based on connected component analysis and geometric properties of labeled regions.

Manesh [15] proposed a method to automatically identify the signature in the scanned document images using a simple region growing algorithm.

Offline handwritten signature verification using ANN [10][13][17] was another concern on this research paper. Sisodia [17] implemented a Static Signature Verification System with four stages such as image pre-processing, feature extraction, classification and decision making. Classifier used an ANN with Error Back Propagation algorithm to attain the result. The relevant features used by the classification are centroid, length and width of the signature in the 200×100 pixels' image box, quadrant areas, one dimensional first and second derivatives of the image and global slant angle. Menu Bhatia [15] was used maximum horizontal and vertical histogram, the center of mass, normalized are of signature, aspect ratio, tri surface feature, six-fold surface feature and transition feature as the extracted features from the candidate signature.

In contrast to the previous research, some have also used HMM and Graphometric features [8][9] and conjunction with neural network and support vector machines [12]. Abdullah [3] proposes a new method for signature recognition using Delaunay triangulation.

Rupali Mehra and coworkers [13] present Surf features and neural-fuzzy techniques based recognition of offline signatures system that is trained with low-resolution scanned signature images. Gautam [17] has used SIFT and Delaunay triangulation for image matching in their research.

Woods [5] considered image area, vertical center, and the horizontal center of the signature, maximum vertical projection, maximum horizontal projection, vertical projection peaks, horizontal projection peaks, number of edge points, number of cross-points and Hough transform for feature extraction of each signature. Extracted values of each signature images from the database of 150 are given to the feed forward neural network (trained using back propagation gradient descent learning).

Gulzar and coworkers [10] present neural network based recognition of offline handwritten signature system that is trained with low- resolution scanned signature images. And also Prashanth C.R. [21] presents DWT based offline signature verification using angular features (DOSVAF). The signature is resized, and Discrete Wavelet Transform (DWT) is applied on the blocks to extract the features.

Vahid Kiani [11] proposes a new method for signature verification using local Radon Transform. The proposed method uses Radon Transform locally as feature extractor and Support Vector Machine (SVM) as the classifier. The main idea is using Radon Transform locally for line segments detection and feature extraction, against using it globally. The advantages of the proposed method are robustness to noise, size invariance and shift invariance. Having used a dataset of 600 signatures from 20 Persian writers, and another dataset of 924 signatures from 22 English writers, their system achieved good results.

In paper [8] a system is introduced that uses only global features. A discrete random transform which is a sinograph is calculated for each binary signature image at the range of $0 - 360$, which is a function of the total pixel in the image and the intensity per given pixel calculated using non-overlapping beams per angle for X number of angles. Due to this periodicity, it is a shift, rotation, and scale invariant. An 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.

Support Vector Machines (SVMs) are machine learning algorithms that use a high dimensional feature space and estimate differences between classes of given data to generalize unseen data. The system in [15] 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 eight signatures thus a total of 320 signatures for training. For initial testing, the approach uses eight original signatures and eight forgeries and achieves FRR 2% and FAR 11% [15].

III. Methodology

The main steps of this research consist of

- Signature extraction from attendance sheets based on morphological operations
- Separate signature, non- signature area using binary Support Vector Machine
- Signature recognition by training Error-Correcting Output Codes multiclass model using SVM
- Signature verification using Kolmogorov Smirnov test

A software package, Matlab2016b is used for this procedure.

a) Signature Extraction from the Scanned Attendance Sheets

An important task in the automated processing of scanned attendance sheets is to extract the signatures. Here both signature and non-signature area are extracted and classify them to separate signature area. The main steps in the signature extraction process are represented by Fig.1.

The main steps are:

- Image Pre-processing
- Classification using binary SVM

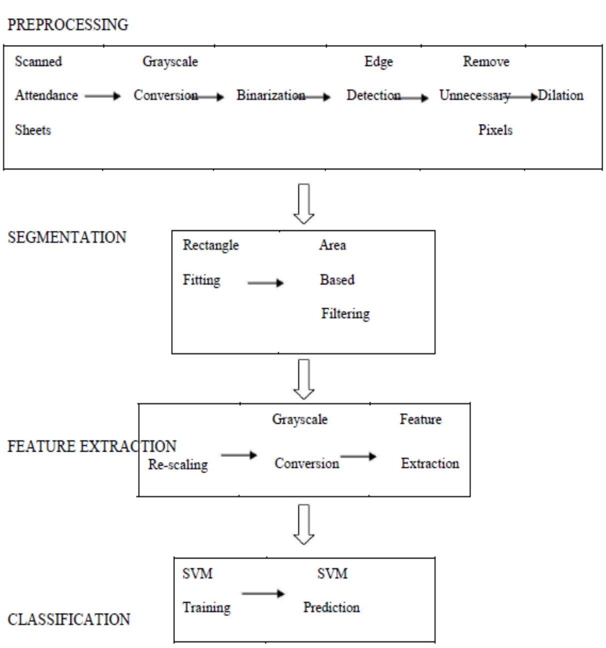

Fig. 1: A brief outline of signature extraction

b) Image Pre-processing

In the pre-processing phase gray conversion, image binarization, edge detection, remove unnecessary pixels, morphological dilation (close, thicken, bridge), image segmentation and cropping are performed.

Fig. 2: Apply Pre-processing Techniques for Scanned Attendance Sheets

c) Classification using Binary SVM

After extracting the images from attendance sheets, the segmented images should be classified using a set of representative features. Here used features of segmented images in distinguishing signature and non-signature segments over a binary SVM classifier.

d) Signature Recognition

Signature recognition is a writer identification problem, whose objective is to find the author of a test signature given a database of signature exemplars from different signers. They are composed of special characters and flourishes, and therefore most of the time they can be unreadable. Also, intrapersonal variations and interpersonal differences make it necessary to analyze them as complete images and not as letters and words put together [15]. The signature acquisition, Pre-processing of signatures, Feature extraction, Train and test an Error Correcting Output Codes (ECOC) multiclass model using SVM are the main steps of signature recognition.

e) Signature Acquisition

Handwritten signatures are taken from 103 students who followed Statistical Inference-I course module in semester-I of academic year 2014/2015, in Faculty of Applied Sciences, Wayamba University of Sri Lanka. Seven signatures are taken from each student so that 721 are used in signature recognition phase.

Fig. 3: Sample Signature Database

f) Pre-processing of Signatures

Gray conversion, Image binarization, Remove unnecessary pixels, Thinning and Data area cropping are performed in the pre-processing phase:

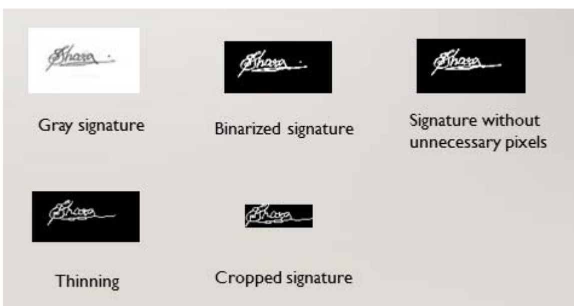

Fig. 4: Apply Pre-processing Techniques for Signatures

g) Feature Extraction of Signatures

Before the feature extraction process to increase the accuracy of the system signature image is partitioned into four equal parts and extract features from each part. So, that the number of features which can be used to train the model has been increased.

Fig. 5: Partitioned Signature into 4 Parts

The choice of the features that provided to the classifiers of the system is essential. In this work, global and local features are used. Pure width, pure height, baseline shift, kurtosis, skewness, maximum vertical projection, maximum horizontal projection, vertical center of mass and horizontal center of mass, Hough transform, etc. are used as global features.

To increase the accuracy of the system grid based features are also extracted from the handwritten signatures as local features. Here Histogram Orient Gradient (HOG) features are extracted as grid features and combine them with global features in the recognition process. The total number of extracted HOG features is 2592.

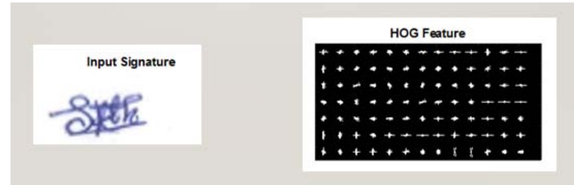

Fig. 6: HOG features extracted from one signature

Local and global features of images are given to *error-correcting output code multiclass model* (ECOC) to train it, so that particular signature image is recognized. ECOC classification requires a coding design, which determines the classes that the binary learners train on, and a decoding scheme, which ensures how the results (predictions) of the binary classifiers are aggregated. Suppose that there are three classes, the coding design is one vs. one, the decoding scheme uses loss *g*, and the learners are SVMs. To build the classification model, ECOC follows following steps.

A one vs. one coding design is:

	Class 1	Class 2	Class 3
Learner 1	1	1	0
Learner 2	-1	0	1
Learner 3	0	-1	-1

Learner 1 trains on observations having Class 1 and Class 2, and treats Class 1 as the positive class, and Class 2 as the negative class. The other learners are trained similarly. Let *M* be the coding design matrix with elements m_{kl}, and s_l be the predicted classification score for the positive class of learner *l*.

A new observation is assigned to the class () that minimizes the aggregation of the losses for the *L* binary learners. That is,

$$\hat{k} = \underset{k}{argmin} \frac{\sum_{l=1}^{L} |m_{kl}| \, g(m_{kl}, s_l)}{\sum_{l=1}^{L} |m_{kl}|}$$

ECOC models can be used to improve the classification accuracy, even compared to other multiclass models.

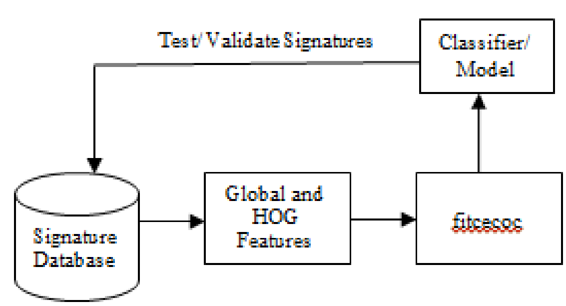

Fig. 7: Signature Recognition using Multiclass SVM

h) Signature Verification

The performance task of signature verification process is one of determining whether a questioned signature is genuine or not. The image of a questioned signature is matched against multiple images of known signatures. Visual signature verification is naturally formulated as a machine learning task. The machine learning tasks can be stated as general learning (which is person-independent) or special learning (which is person-dependent), paralleling the learning tasks of the human questioned document examiner. In the case of general learning the goal is to learn from a large population of genuine and forged signature samples. The focus is on differentiating between genuine-genuine differences and genuine-forgery differences.

Special learning focuses on learning from genuine samples of a particular person. The focus is on learning the differences between members of the class of genuine. The verification task is a one-class problem of determining whether the questioned signature belongs to that class or not.

Using Kolmogorov Smirnov test correctly classified signatures are used to confirm the genuineness.

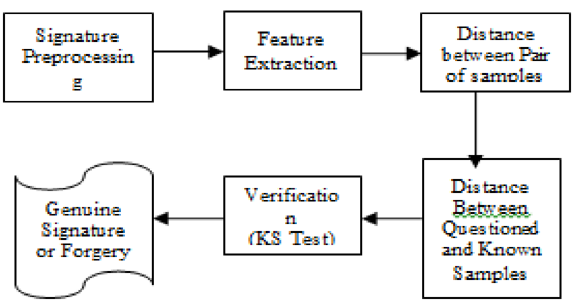

Fig. 8: Signature Verification using Kolmogorov Smirnov Test

If a given person has N samples, $\binom{N}{2}$defined as N! /N! (N-r)! pairs of samples can be compared as shown in Fig.8. Let N be the total number of samples and $N_{WD} = \binom{N}{2}$ be the total number of comparisons that can be made which also equals the length of the within-person distribution vector. The within-person distribution can be written as

$$D_W = (d_1, d_2, \ldots, d_{NWD})^T$$

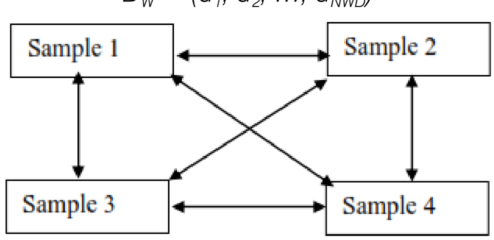

Fig. 9: Comparing All Possible Genuine-Genuine Pairs

Analogous to this, the questioned sample signatures can be compared with every one of the N knowns similarly to obtain the questioned vs. known distribution. The questioned vs. the known distribution is given by

$$D_{QK} = (d_1, d_2, \ldots, d_N)^T$$

where d_j is the distance between the questioned sample and the j^{th} known sample, $j \in \{1, \ldots, N\}$.

i) Performance Evaluation

For unmatched signatures distance statistics is large and for matched signatures distance statistic is small. Since .004 is less than .20, the null hypothesis has been accepted. That is distributions are approximately same for matched signatures. In this study 0.01 has taken as the significance level.

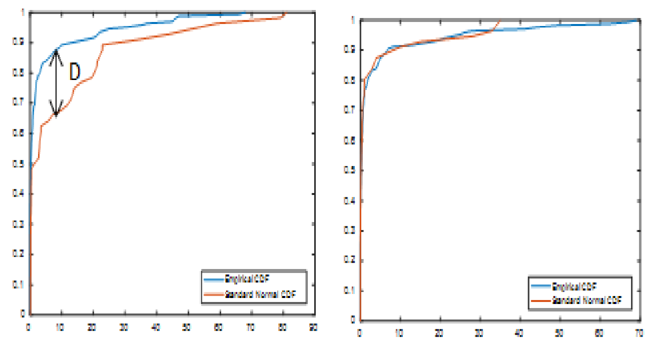

Fig. 10: Test Statistic for Unmatched and Matched Signatures

ALGORITHM:

Input = Signature image

Output = Conformation from system whether the signature is genuine or counterfeit.

Step 1: Acquire matched signature images from the signature recognition process

Step 2: Enhanced the signature images by pre-processing

Step 3: Create a feature vector by combining extracted features from the pre-processed signature images.

Step 4: Obtain the distances of features between every seven samples of the known signature in the database. (Results gave 21 × 8 matrices)

Step 5: Same has been done between the known sample and questioned sample. (Results gave 7 × 8)

Step 6: Apply KS test for two distributions and obtain the probability of similarity.

Step 7: Repeat step 1-7 to test all the signatures recognized by the system.

Step 8: If the probability is less than 0.01 the signature is identified as “Forge”, otherwise as “Genuine”.

Table 1: Sample Distance Distribution of Known Signature

1	2	3	4	5	6	7
0.44	0.68	0.52	0.87	1.01	0.81	0.93
1.67	2.44	1.93	2.95	3.27	2.80	3.09
122	98	139	67	38	69	43
0.27	0.54	0.67	0.08	0.30	0.40	0.28
0.048	0.07	0.05	0.1	0.12	0.09	0.10
0.10	0.16	0.12	0.21	0.24	0.19	0.22
5	7	9	7	8	5	5
0.03	0.06	0.04	0.08	0.09	0.07	0.08

Table 2: Sample Distance Distribution of Known vs. Questioned

1	2	3	4		19	20	21
0.24	0.08	0.43	0.57	…	0.19	0.07	0.11
0.77	0.26	1.27	1.59	…	0.46	0.18	0.28
24	17	55	84	…	31	5	26
0.26	0.39	0.18	0.03	…	0.09	0.01	0.11
0.03	0.00	0.05	0.07	…	0.03	0.01	0.01
0.06	0.02	0.11	0.14	…	0.05	0.02	0.03
2	4	2	3	…	3	3	0
0.02	0.00	0.04	0.05	…	0.01	0.00	0.01

j) Enter Data to Excel Sheets

After identifying whether a particular signature is genuine or forge the attendance records has been entered to the excel sheets. If the KS test identify the signature as genuine in verification process, '1' has been entered in front of the relevant student index number in the excel sheet.

IV. Results & Discussion

When extract the signatures from scanned attendance sheets in some situations there are some discontinuities in the signatures. In those situations, whole signature is not including in the bounded region as following figure (one signature is separated into parts).

Fig. 11: Signature Image with Multiple Bounded Regions due to Discontinuity

To overcome that problem edge detection, morphological dilation, thicken and bridge has been used in pre-processing stage.

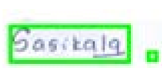

Fig. 12: Signature Image after Removing the Discontinuity

The errors we are trying to minimize in recognition and verification are: classifying one person's signature as belonging to another one and acceptance of a fake signature. Some signatures are misclassified by another student's signature due to some similarities between two signatures in the recognition process.

Fig. 13: Misclassification of a Signature

V. Conclusion

Today information technology has proved that there is a need to retrieve, search, query and store large amount of electronic information efficiently and accurately.

In signature extraction phase, the whole part of the signature not extracted due to some reasons. Those are: excessive dusty noise, logos, figures, printed and handwritten text etc., large ink- blobs joining disjoint characters or components, degradation of printed text due to poor quality of paper and ink, text overlapping the signature. By increasing the space between text and signatures, it could be avoided the overlapping of signatures with text. Proposed system is extracted signatures with 100% accuracy.

In recognition process the combination of global and local features are used to train the ECOC multiclass model using SVM. The accuracy is 83.33% and it suggest that the use of gradient-based feature sets with global features can serve the most reliable way of detecting signatures in signature recognition process.

A machine learning approach is used in signature verification process, because only the genuine signatures are in the registered student database. All signatures are identified as genuine with 100% accuracy. Finally, we can conclude that this system can be used in a university educational environment for automatic student authentication. Eventually, based on the methodologies employed in this report, it provides a promising stage for the development of an automated attendance entering system.

References Références Referencias

1. Banka, R., Nourbakhsh, F., 2010. Extraction of Signature and Handwritten Regions from Official Binary document Images.
2. Bhattacharyya, D., Kim, T., 2010. Design of artificial neural network for handwritten signature recognition. Int. J. Comput. Commun. 4, 59–66.
3. Coetzer, J., 2005. Off-line signature verification. Stellenbosch: University of Stellenbosch.

4. Cüceloğlu, İ., Oğul, H., 2014. Detecting handwritten signatures in scanned documents, in: Proceedings of the 19th Computer Vision Winter Workshop. pp. 89–94.
5. Gonzalez, R.C., Woods, R.E., 2002. Digital image processing. Pearson Education, Delhi, India.
6. Gupta, C.S., Dixit, U.D., 2015. A Review on Signature Detection and Signature Based Document Image Retrieval 4.
7. Hafemann, L.G., Sabourin, R., Oliveira, L.S., 2015a. Offline handwritten signature verification-literature review.
8. Kalera, M.K., Srihari, S., Xu, A., 2004a. Offline signature verification and identification using distance statistics. Int. J. Pattern Recognit. Artif. Intell. 18, 1339–1360.
9. Karouni, A., Daya, B., Bahlak, S., 2011. Offline signature recognition using neural networks approach. Procedia Comput. Sci. 3, 155–161. 2010.12.027
10. Khuwaja, G.A., Laghari, M.S., 2011. Offline handwritten signature recognition. World Acad. Sci. Eng. Technol. 59, 1300–1303.
11. Kiani, V., Pourreza, R., Pourreza, H.R., 2009. Offline signature verification using local radon transform and support vector machines. Int. J. Image Process. 3, 184–194.
12. Kumar, P., Singh, S., Garg, A., Prabhat, N., 2013a. Hand written signature recognition & verification using neural network. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 3.
13. Mehra, R., Gangwar, R.C., 2014. A Survey: Enhanced Offline Signature Recognition Using Neuro-Fuzzy and SURF Features Techniques Vol 5(3), 4350–4353.
14. Nguyen, V., Blumenstein, M., Muthukkumarasamy, V., Leedham, G., 2007. Off-line signature verification using enhanced modified direction features in conjunction with neural classifiers and support vector machines, in: Ninth International Conference on Document Analysis and Recognition (ICDAR 2007). IEEE, pp. 734–738.
15. Özgündüz, E., Şentürk, T., Karslıgil, M.E., 2005. Off-line signature verification and recognition by support vector machine, in: Signal Processing Conference, 2005 13th European. IEEE, pp. 1–4.
16. Piyush Shanker, A., Rajagopalan, A.N., 2007. Off-line signature verification using DTW. Pattern Recognit. Lett. 28, 1407–1414.
17. Pushpalatha, N., Gautam, A., 2014. Offline signature Verification using spatial domain feature sets and support vector machine. Int. J. Emerg. Technol. Adv. Eng. 4, 544.
18. Shirdhonkar, M.S., Kokare, M.B., 2010. Discrimination between printed and handwritten text in documents. IJCA Spec. Issue on 2010.
19. Singh, N., Kaushal, S., 2015. International Journal of Engineering Sciences & Research Technology Off-Line Signature Recognition Using Modified Neural Networks Approach.
20. Sisodia, K., Anand, S.M., 2009. Off-line handwritten signature verification using artificial neural network classifier. Int. J. Recent Trends Eng. 2, 205–207.