Offline Signature Verification based on Geometric Features using Filter Method

Sunil Kumar D.S., PhD
Department of Computer Applications
Administrative Management College
Bangalore, India

ABSTRACT
This paper presents an offline signature verification using Geometric features. In this approach the acquired signatures samples undergoes pre-processing operation which includes resize, filtering, cropping, and thinning. Then Geometric features are extracted from each signature image. The extracted features are from normalized signature area**.** Experiments are conducted on publically available benchmark datasets namely CEDAR and GPDS. The best feature subsets of the data sets were selected using filter and wrapper methods. Based on the feature vector, the proposed approach will detect the forgery or genuine signature using filter matching method. Experimental results shows the performance of our proposed approach.

General Terms
Behavioral biometrics, Offline Signature Verification and Geometric features.

Keywords
Offline Signature Verification, Preprocessing, Classification, Filter Method.

1. INTRODUCTION
Hand written signature is one of the behavioral biometric trait, which is considered as one of the person authentication method in day to day transaction. These characteristics of a signature are unique from one individuals to another. Hence can be used to authenticate and represent a person. Unlike PIN or OTP number signatures can't be stolen but sometimes its pattern can be forged by other to get the benefits. In such cases there is a need of signature verification system to verify whether the written signature is genuine or forge. The signature verification systems are of two types, namely Off-line and On-line signature verification system. The procedure of signature verification is same in both types, but only difference is in signature acquiring method. In offline signature verification system signature is written on a paper by contributors then it is fed for verification process after necessary preprocessing techniques are applied. But in On-line signature verification system signatures are collected by allowing contributors to write signature on a surface of specially designed digital tablet, which is then fed for verification. Offline signatures possess static features whereas On-line signatures possess dynamic features. Both type of signature verification aims to verify whether the signature is genuine or forge. Forge signatures are of three types viz. Simple, Random and skilled. Among these type detecting skilled forgery is most challenging. We addressed this problem in our proposed approach.The order of this paper with section wise is as follows: section-2 presents a detailed literature of some of the state- of- the- art papers. The elaborated description of the proposed approach is in section-3. Section-4 illustrates the experimental results followed by the discussion. Finally, section-5 represents the conclusion part and section-6 presents references.

2.LITERATURE REVIEW
Research in the area of signature verification has brought a lot of innovations from the researchers over the last two decades. The Local histogram features approach [17] divides the signature into number of zones with the help of both the Cartesian and polar coordinate systems. For every zone Histogram of Oriented Gradients (HOG) and histogram of local binary patterns (LBP) histogram features are computed. A variant of LBP namely Blockwise Binary Pattern (BBP) [15] divides signature into 3x3 blocks to extract local features. Writer-independent [12] method emphasize on shape and texture features of a signature. This method extracts the black pixels concentrated along with candidate pixel. The proposed method employs MLP and SVM classifiers. The Inter-point Envelope Based Distance Moments [9] captures structural and temporal information by computing inter-point distances, which is the distance between reference point and other points in an extracted envelope. Shikha et al. [16] proposed an offline signature recognition system, which is based on neural network architecture. The proposed method uses Self Organizing Map (SOM) as a learning algorithm and Multi-layer perceptron (MLP) for classification of patterns. Shekar et. al., [14] proposed a grid structured morphological pattern spectrum. The proposed method divides the signature with equally sized eight grids. Then pattern spectrum is obtained for every grid. Bhattacharya et al., [3] have proposed a pixel matching technique (PMT). The proposed approach verifies by comparing each pixel of sample signature against test signature. Kruthi and Shet [8] proposed Support Vector Machine based offline signature verification system (SVM). SVM is employed for classification of signatures as objects. Yasmine et.al. [5] proposed an approach for offline signature verification system, which is based on one-class support vector machine. This technique considers only genuine signature pattern of a signer. Soleimani et.al. [16] Proposes an approach The Deep Multi task Metric Learning (DMML). The proposed approach considers the similarities as well as dissimilarities of a signature. Besides these plethora of algorithms for offline signature verification, still devising a more accurate and most efficient offline signature verification system is difficult. Therefore, we were inspired to devise a more effective techniques and model for offline signature verification system.

3. PROPOSED METHOD

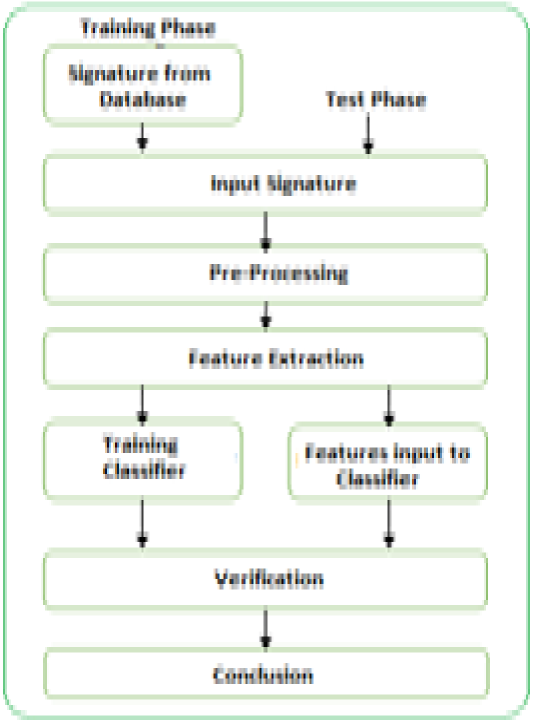

Figure 1. Proposed Signature Verification Model

3.1 Pre-processing

After capturing the signature samples, the next step is to enhance the images and make them ready for the subsequent processing. That is the scanned images need to be pre-processed before giving them to the next process. Preprocessing is done using signal processing algorithms. Preprocessing greatly helps to improve the performance of feature extraction and classification. It reduces computational cost in classification.

3.1.1Filtering

A scanned signature image may contain noise. Noise in the image deteriorates the feature extraction and its successive processes. Hence, filtering of noise is an unavoidable pre-processing step in pattern recognition. It has been observed that the scanned images are usually affected by salt-peeper noise. A median filter effectively removes such type of noise preserving the edges of the images. We applied a median filter of 3×3 window on our signature images.

3.1.2 Binarization

This is the process of converting colour image in to binary image. The binary imagecomprises of black and white pixel intensity.

3.1.3 Cropping

When scanned, signature image contains the signature and some white coloured non-signature regions. Those superfluous non-image portions are removed by cropping the image to the bounding rectangle of the signature part. Cropping is an essential pre-processing step for all types of classification techniques.

3.1.4 Thinning

In thinning, the signature image strokes are made one pixel thick. Thinning is mainly done to reduce the amount of data in the image. This helps to decrease the storage space requirement and also to reduce the computational complexities in successive stages. But during thinning, some information of the signature images such as stroke width may be lost. So, depending on the features to be extracted, thinning may or may not be required.

3.1.5 Skeletonization

Skeletonization is applied on binary images. It preserves the connectivity of the signature segments which were originally connected and removes selected foreground pixels from the image. After skeletonization, the signature image is converted into combination of some thin arcs and curves.

3.1.6 Rotation for Skew Correction

Many a times, it is seen that during scanning of the signature images, the images are not properly oriented. This angular tilt in the signature image is called ‘skew’. Skew may result poor classification (depending on the classification technique used). Therefore there may be a need for skew correction of the signature images by rotating them. After skew correction, the final image is made parallel to the horizontal axis.

3.1.7 Slant Correction

Slant is the tilt that an inclined signature makes with the vertical axis. Sometimes, the slant angle needs to be corrected before feature extraction.

3.1.8 Resizing

Signature lengths are different for different signers. Even the lengths of the signatures of a single person are also not equal. But when a grid based signature verification approach is used, the signatures are projected on the grid of same size. Hence, all the signatures must be of same size.

3.2 Feature Extraction

Feature extraction is a process of deriving some characteristic parameters or functions from the patterns (signature images). The extracted characteristic parameters or functions are called ‘features’. Function features are functions of time and these can only be derived from online signatures. Characteristic parameters are extracted from offline signatures.

We extracted the following features from the signature samples present in our datasets:

3.2.1 Normalized Signature area (with respect to bounding box)

It is the total number of signature pixels or foreground pixels in the signature image. Signature area gives information about the signature density. If the signature image is skeletonized, signature area represents a measure of the density of the signature traces.

3.2.2 Aspect Ratio (Signature width to height ratio)

This is ratio of signature width to signature height of a cropped signature. It is seen that aspect ratio of the signatures of a person fairly remains constant.

3.2.3 Horizontal and vertical centre of the signature

These two measurements indicate about the Horizontal and Vertical location of the signature image.

The horizontal center (C_x) is given by

$$C_x = \frac{\sum_{y=1}^{x_{\max}} x \sum_{x=1}^{y_{\max}} b[x,y]}{\sum_{x=1}^{x_{\max}} \sum_{y=1}^{y_{\max}} b[x,y]}$$

The vertical center (C_y) is given by

$$C_y = \frac{\sum_{y=1}^{y_{\max}} y \sum_{x=1}^{x_{\max}} b[x,y]}{\sum_{x=1}^{x_{\max}} \sum_{y=1}^{y_{\max}} b[x,y]}$$

3.2.4 Horizontal and Vertical Projections (Horizontal and Vertical Histograms)

Horizontal projection or histogram is found by counting the number of signature image pixels in each row in a signature image and plotting it horizontally with a line. The row with maximum value gives the maximum horizontal projection. Similarly, counting and plotting signature pixels in vertical direction for every column, maximum vertical projection or histogram is found.

Horizontal and Vertical Projections can be calculated as follows:

The horizontal center (C_x) is given by

$$P_h[y] = \sum_{x=1}^{n} \text{black pixel } (x,y)$$

$$P_v[x] = \sum_{y=1}^{m} \text{black pixel} (x,y)$$

3.2.5 Signature height

It is the height of a signature image, after width normalization.

3.2.6 Maximum horizontal projection (Maximum horizontal histogram)

In a horizontal histogram, the row with maximum value gives the maximum horizontal histogram.

3.2.7 Maximum vertical projection (Maximum vertical histogram)

The highest value of the projection histogram in the vertical histogram is the maximum vertical projection. Maximum Horizontal Histogram and Maximum Vertical Histograms are also termed as pure width and pure height of the signature.

Pure width: It is the maximum number of total image pixels (i.e. black pixels) among all the rows counted after cropping the signature image.

Pure height: It is the maximum number of total image pixels (i.e. black pixels) among all the columns counted after cropping the signature image.

3.2.8 Centre of Gravity or Centroid

In a binary signature image with black signature pixels, Centre of Gravity (CG) or Centroid is the average coordinate point of all black pixels.

The CG of a signature image is calculated by the following equations:

$$X = \frac{1}{N}\left(\sum_{i=1}^{n} x_i\right)$$

$$Y = \frac{1}{N}\left(\sum_{i=1}^{n} y_i\right)$$

x_i column number of ON pixels and $y_i$is the row number of ON pixels. The pixels that constitute the stroke of the signature image are called the ON pixels. In a binary image, ON pixels are the black pixels.

3.2.9 Slope of line obtained from curve fitting of centre of gravity of each column

Center of gravity of each column of the signature image is found. Then the linear polynomial curve was obtained using least square curve fitting technique. Slope of this line is the signature feature.

The general expression for curve fitting using a least square curve fitting method is:

$$err = \sum (d_i)^2 = (y_1 - f(x_1))^2 + (y_2 - f(x_2))^2 + (y_3 - f(x_3))^2 + (y_4 - f(x_4))^2$$

Here, err = The error function

(x_i, y_i) = Data point (CG value), $(x_i, f(x_i))$ = Point on the fitted line

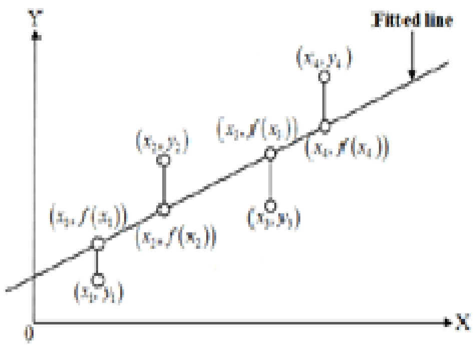

Figure 2. Curve fitting using Least Square Curve Fitting Method

3.2.10 Center of Gravities of the vertically divided images

The Signature image is vertically divided into two halves. Then for both the halves, the center of gravities CG_1 and CG_2 are found.

3.2.11 Skew Angle

If CG_1 and CG_2 are the center of gravities of the two halves of the vertically divided images, then the angle between the horizontal axis and the line joining CG_1 and CG_2 is called the 'Skew angle'.

3.2.12 Slope of Center of Gravity of two equal halves of signature image

If CGs of the first and second halves of the signature image are (x_1, y_1) and (x_2, y_2) then slope of the line joining the two CGs is

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

3.2.13 Baseline shift or Orientation of signature

To calculate this parameter, two centroids (say CG_1 and CG_2) of the vertically equally divided signature image are found. Baseline shift or Orientation of signature = $y_2 - y_1$

3.2.14 Area of the signature Bounding Box

It defines the area of the signature bounding box after cropping of the signature image.

3.2.15 Number of cross points

In a skeletonized signature image, Cross point is a signature pixel or point that has more than two 8-neighbours. (When pixels are considered on a grid, a pixel is surrounded by 8-neighbouring pixels; they are called its 8-neighbours)

3.2.16 Baseline Slant Angle

Baseline is an imaginary line assumed below the signature. The signature is assumed to be sitting on the baseline. The angle made by the baseline of a signature with the horizontal line called 'Baseline slant angle'.

3.3 Classification using Filter Method

Filter method of feature selection uses different statistical tests to find out the features that have the highest predictive power. For a chosen statistical measure, this method calculates a score for each feature, and based on the scores, features are given a rank. This method is independent of the classifier.

Filter method of feature selection uses different statistical tests (e.g. T-test, F-test, i-test, Euclidean distance, (c 2) Chi-squared test, ANOVA, Information gain, Correlation coefficient scores etc.) to find out the features that have the highest predictive power.

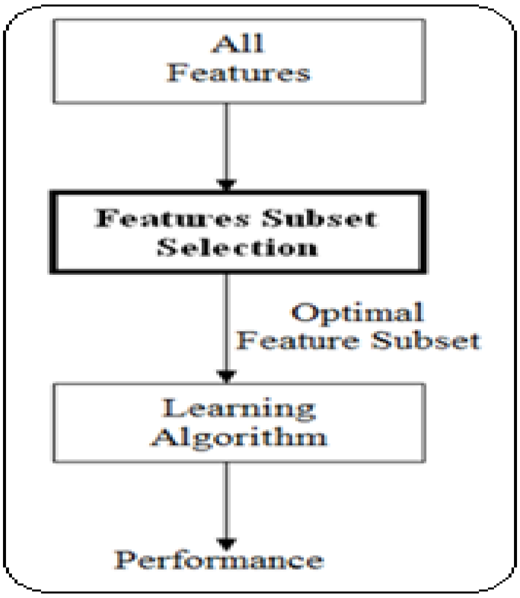

Figure 3: Operations in Proposed Filter Method

Filter Algorithm

input: $F(F_0 F_1 F_2 ...., F_{n-1})$ // a training data set with N no. of features

S_0 // a subset with which the search is started

T // a stopping criterion

output: S_{best} // the best selected subset

begin

initialize: $S_{best} = S_0$;

$\chi_{best} = eval(S_0, F, M)$; // evaluate S_0 by an independent measure M

do begin

$S = generate(F)$; // generate a subset for evaluation

$\chi = eval(S, F, M)$; // evaluate the current subset S by M

if (χ is better than χ_{best})

$\chi_{best} = \chi$;

$S_{best} = S$;

end until (T is arrive at);

return S_{best};

end;

4. EXPERIMENTAL RESULTS

The experiments are carried out in MATLAB 2016a software environment on Asus laptop with 4GB RAM, 500GB Hard Disk and i5 processor on windows 8 operating system. The following figure.3 is the snapshot of Graphical User Interface for Offline Signature Verification developed using GUID tool available in MATLAB.

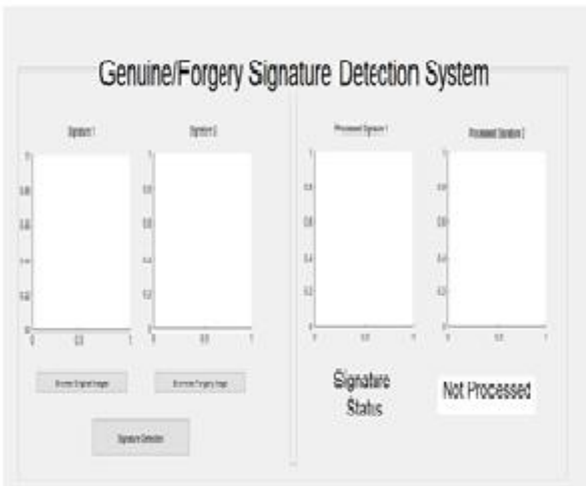

Figure 4: GUI for Genuine/Forgery signature detection system

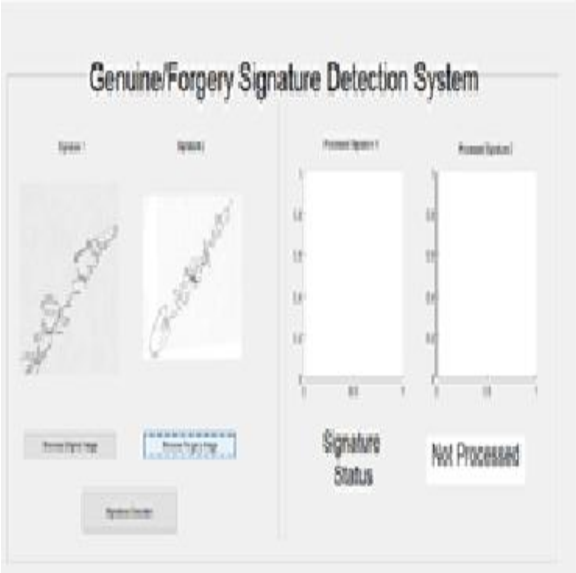

Figure 5: Reading the input forgery/genuine signature

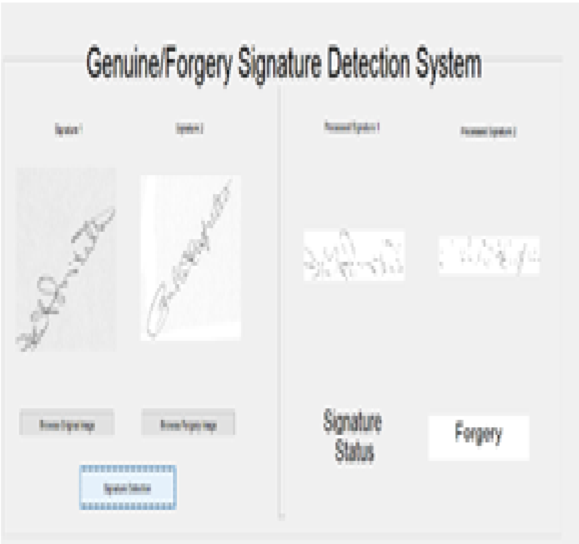

Figure 6: Feature extraction and Forgery Signature Detection Result

Experiments are conducted on publicly available benchmark datasets namely CEDAR (Center for Document Analysis and Recognition) which is developed by SUNNY Buffalo and GPDSSynthetic Signature database.Theknowledge repository contains the Edge histogram and Edge directional histogram features extracted from every signature sample of the data set including both genuine and skilled forge signature samples. For each dataset, the signature samples are considered into two groups: training sample set and testing sample set with varying number of samples. We have carried out four sets of experiments. In Set-1, first ten genuine and first ten skilled forgeries are chosen as training samples and tested against the remaining samples of the respective datasets, where as in Set-2, we have considered first 15 samples of genuine and first 15 samples of skilled forgery for training and tested with remaining samples of the dataset. In Set-3,we have randomly chosen 10 genuine and randomly chosen 10 forge samples are considered for training, and tested with the remaining samples of the dataset, and in Set-4,there are 15 samples are chosen randomly from the respective datasets for training and remaining samples are considered for testing. In order to overcome from the effect of the randomness, Set-3 and Set-4 experimentations are repeated five times and the average result is tabulated.

Table1. Datasets details

DATABASE	NO.OF USERS	NO.OF GENUINE	NO. OF FORGE	TOTAL NO.
GPDS Synthetic Signature (Offline)	4000	24	30	216000
CEDAR	55	24	24	2640

A. Experiments on CEDAR dataset:

CEDAR (Center for Document Analysis and Recognition) is a benchmark offline signature database, publicly available. We considered CEDAR database for our experiments. CEDAR database consists 24 genuine and 24 forgeries from 55 contributors. The total number of signature samples is 2640. The details of databases are tabulated in the table 1. Out of 2640 samples, we considered few samples for our experiments. We conducted experiments in 4 sets. In first and third set we considered 10 genuine samples and 10 skilled forgeries for training. And for testing we choose 14 genuine samples and 14 skilled forgeries for testing. In second and fourth set we considered 14 genuine and 14 skilled forgeries for training. And for testing 9 genuine samples and 9 skilled forgeries samples for testing. To avoid randomness set second and fourth are repeated 5 times and average is considered. In table 2 the results on CEDAR database is tabulated with FRR and FAR are performance evaluation metrics.

Table 2. Results obtained for CEDAR

Experimental Setup	FRR	FAR	Accuracy
Set-1	10	8	94
Set-2	6	6	96
Set-3	8	4	92
Set-4	7	5	92.5

From the literature, we found the experimental results of few well known approaches on CEDAR dataset. The comparative analysis presented in table 3 shows the improvements in accuracy by the proposed approach.

Table.3 Experimentation Results obtained for CEDAR Dataset - A comparison:

Proposed by	Classifier	Accuracy	FAR	FRR
Kalera et al. [12]	PDF	78.50	19.50	22.45
Chen and Shrihari [13]	DTW	83.60	16.30	16.60

Kumar et al., [14]	SVM	88.41	11.59	11.59
Shekar et al. [15]	EMD	91.06	10.63	9.4
Kumar et al. [16]	MLP	91.67	8.33	8.33
Kumar et al. [17]	SVM	92.73	6.36	8.18
Proposed approach	Filter Based	96	6	6

B. Experiments on GPDS Synthetic Offline Signature dataset:

Digital Signal Processing Group (GPDS) of the Universidad de Las Palmas de Gran Canaria, Spain had developed a large scale signature corpus; it included with many sub corpus like GPDS-960, GPDS-300 and GPDS-160. For our experiments we considered GPDS Synthetic Offline Signature dataset, which is a well-known benchmark dataset. It can be collected by communicating with the concern developer. The GPDS Synthetic offline signature dataset consist 4000 contributors, each contributor contributed 24 genuine signature samples and 30 forge signature samples. This results a total number of signatures are 216000 which is tabulated in table 1. The experiments are conducted on GPDS dataset on 4 sets. For set-1 and set-2, we considered 10 genuine and 10 forge samples for training. For testing we considered 14 genuine and 20 skilled forgery samples. For set-2 and set-4 we considered 15 genuine and 15 skilled forgery samples for training. We tested against 9 genuine and 15 skilled forgery samples. The obtained results are tabulated in table 4.

From the literature,we found the experimental results of few well known approaches onGPDS dataset. The comparative analysis presented in table 5 shows the improvements in accuracy by the proposed approach.

Table 4. Results obtained for GPDS

Experimental set-up	Accuracy	FRR	FAR
Set-1	92	8	5
Set-2	94.80	6	8
Set-3	93.52	7	5
Set-4	95	4	6

Table 5. Experimentation Results obtained for GPDS Dataset - A comparison:

Proposed by	Classifier	Accuracy	FAR	FRR
Ferrer et al., [18]	SVM HMM	86.65 -	13.12 12.60	15.41 14.10
Vargas et al., [19]	SVM+LBP	87.28	6.17	22.49
Ruiz-Del-Solar et al., [20]	Bayseian	84.70	14.20	16.40
Kumar et al., [21]	MLP	86.24	13.76	13.76
Shekar et al., [22]	EMD	91.06	10.63	9.4
Proposed Approach	Filter based	95	6	4

5 CONCLUSION:

The Proposed approach performs offline signature verification. This approach significantly improved the verification accuracy by employing image processing methods to detect the forgery or genuine signature. Once the pre-processing of images is completed then extract features from images. Filter method is applied on extracted features for the classification. As we used skilled forgeries in our experiments the results we got are quite good. Depending on the style of the signature, stability of its features varies. Rich discriminating features influences the recognition rate. Our proposed approach is a combination of feature subsets and a classifier. Obtained results demonstrates the performance of proposed approach in terms of accuracy.

6. REFERENCES

[1] Bastys, A., Kranauskas, J., &Krüger, V. (2011). Iris recognition by fusing different representations of multi-scale Taylor expansion. Computer Vision and Image Understanding, 115(6), 804-816.

[2] Bastys, A., Kranauskas, J., &Masiulis, R. (2009). Iris recognition by local extremum points of multiscale Taylor expansion. Pattern recognition, 42(9), 1869-1877.

[3] Bhattacharya, I., Ghosh, P., & Biswas, S. (2013). Offline signature verification using pixel matching technique. Procedia Technology, 10, 970-977.

[4] Chen, S., & Srihari, S. (2005, August). Use of exterior contours and shape features in off-line signature verification. In Eighth International Conference on Document Analysis and Recognition (ICDAR'05) (pp. 1280-1284). IEEE.

[5] Guerbai, Y., Chibani, Y., &Hadjadji, B. (2015). The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters. Pattern Recognition, 48(1), 103-113.

[6] Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). A practical guide to support vector classification.

[7] Kalera, M. K., Srihari, S., & Xu, A. (2004). Offline signature verification and identification using distance statistics. International Journal of Pattern Recognition and Artificial Intelligence, 18(07), 1339-1360.

[8] Kruthi, C., &Shet, D. C. (2014, January). Offline signature verification using support vector machine. In 2014 Fifth International Conference on Signal and Image Processing (pp. 3-8). IEEE.

[9] Kumar, M. M., &Puhan, N. B. (2014, July). Inter-point envelope based distance moments for offline signature verification. In 2014 International Conference on Signal Processing and Communications (SPCOM) (pp. 1-6). IEEE.

[10] Kumar, R., Kundu, L., Chanda, B., & Sharma, J. D. (2010, December). A writer-independent off-line signature verification system based on signature morphology. In Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia (pp. 261-265). ACM.

[11] Kumar, R., Sharma, J. D., &Chanda, B. (2012). Writer-independent off-line signature verification using surroundedness feature. Pattern recognition letters, 33(3), 301-308.

[12] Shekar, B. H., &Bharathi, R. K. (2011, June). Eigen-signature: A robust and an efficient offline signature verification algorithm. In 2011 International Conference on Recent Trends in Information Technology (ICRTIT) (pp. 134-138). IEEE.

[13] Shekar, B. H., Bharathi, R. K., & Pilar, B. (2013, December). Local morphological pattern spectrum based approach for off-line signature verification. In International Conference on Pattern Recognition and Machine Intelligence (pp. 335-342). Springer, Berlin, Heidelberg.

[14] Shekar, B. H., Pilar, B., & Sunil, K. D. S. (2017). Blockwise binary pattern: a robust and an efficient approach for offline signature verification. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 227.

[15] Shikha, P., &Shailja, S. (2013). Neural network based offline signature recognition and verification system. Research Journal of Engineering Sciences ISSN, 2278, 9472.

[16] Soleimani, A., Araabi, B. N., &Fouladi, K. (2016). Deep multitask metric learning for offline signature verification. Pattern Recognition Letters, 80, 84-90.

[17] Yilmaz, M. B., Yanikoglu, B., Tirkaz, C., &Kholmatov, A. (2011, October). Offline signature verification using classifier combination of HOG and LBP features. In 2011 International Joint Conference on Biometrics (IJCB) (pp. 1-7). IEEE.

[18] Parameshachari, B. D., & Kumar, D. S. (2022, October). SVM Based Brain Tumor Detection and Classification System. In *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)* (pp.1-4).IEEE.