International Journal on Advanced Science Engineering Information Technology

Vol.8 (2018) No. 3
ISSN: 2088-5334

Component-connected Feature for Signature Identification

Naeli Umniati[#], A. Benny Mutiara[#], Tb. Maulana Kusuma[#], Suryarini Widodo[#]

[#]*Faculty of Computer Science and Information Technology, Gunadarma University, Depok, 16424, Indonesia*
E-mail: naeli@staff.gunadarma.ac.id, amutiara@staff.gunadarma.ac.id, mkusuma@staff.gunadarma.ac.id, srini@staff.gunadarma.ac.id

***Abstract*—A signature is the oldest security techniques to verify the identification of a person. This is due to every person has a different signature, and each signature has the characteristic physiological and behavior. There are two kinds of signature such as offline and online signatures used to verify someone identity. Offline signatures were used in this study because offline signature does not have dynamic features such as an online signature. This study proposed an identification system of offline signature by using k-NN based on the features that were stored in the database. The proposed identification system consists of preprocessing, feature extraction and verification stages. We collected the data samples from 10 persons. Each person wrote ten signatures. Total data was 100 signatures. The first stage used in this study was preprocessing such as noise removal, binarization, skeleton, and cropping. The second stage was feature extraction. Feature extraction had some vital information such as height-width ratio, the ratio of the density of signatures, edge distance ratio, the ratio of the number and proximity of the column, and the number of connected components in the signature. That information was stored in a separate database. We separated ten signatures of each person into six signatures as data sample and four signature as test data. We verified 40 signatures of test data from 10 persons using k-NN. It is shown that from 40 signatures used in our test data, 28 signatures were correctly identified and 12 signatures belong to others.**

***Keywords*— offline signature; identification systems; feature extraction; k-NN**

I. INTRODUCTION

Biometric technology is currently used in a wide range of security applications. Such technology aims to be able to recognize or identify a person based on physiological characteristics or behavior. At first, this identification is based on the measurement of biological characteristics such as fingerprint, face, iris, and others. In the next development, the identification is done by observing behavioral traits such as speech and signature.

Someone's signature is essential biometric characteristics that are usually applied to personnel or verification of identity verification documents. Verification of a person's identity through the signature is based on the person's biometric measurements. The use of a signature as proof of verification is indicated by the many financial and business transactions approved by signature [1].

The purpose of the signature identification system is to distinguish between two classes: genuine and forgery, about the variability of Intra and Interpersonal. Intrapersonal variation is a variation of the signature of the same person, whereas the variation between genuine and forgery called interpersonal variation.

It was generally known that no two original signature of someone who is the same. Some experts signatures know that the results of the signature written by the same person in a row will have a resemblance, both regarding size and direction of a signature [2].

Based on the method of data acquisition, there are two categories of systems for verifying the identity (ID) someone through signature, namely: dynamic system (online) and a static system (offline) [3]. In the online system, signature data obtained from an electronic tablet to obtain dynamic information about writing activities such as write speed, pen pressure when writing, and the number of strokes [4]. While the system is offline, the signature is written on paper and then converted to digital image format with the help of a camera or scanner [5].

Most of the identification system existing signature follows the general structure of five main stages consisting of data acquisition, pre-processing, feature extraction, identification/verification and performance evaluation [6]. In the data acquisition phase, the data is processed signature is stored as a result of scanning the image. The results of the scan image usually contain much noise so that the necessary pre-processing to produce a clean image before feature extraction stage. Pre-processing is an essential stage in a system identification signatures mainly on offline systems. The purpose of this stage is setting the standard for image feature extraction stage. Typically, the acquired signature image has a different format and resolution that needs to be processed to generate an accurate extraction feature. This stage will affect the accuracy and reduces the computation

time offline signature identification system. Pre-processing typically involves two steps, namely: enhancement and segmentation [7],

- Enhancement: Consists of the elimination of noise, convert the image into a binary image and skeletonization using appropriate algorithms [8].
- Segmentation: consists of extracting a signature to extract the smallest box that contains the signature data, determine the height and width of the signature, the signature cropping, and the standardization of the size of the signature [9].

Feature extraction is the process in which digital information is modified, simplified, combined so that necessary information may be classified. To be successful, should feature extraction technique using rules governing the formation of a pattern class. The feature extraction is used as an input to the process of learning and decision-making process. Therefore the feature extraction technique is critical to the success of the whole process automatic pattern identification [10]. A useful feature is a feature that allows the system to identify the pattern class with the least amount of mistakes. Features elections must be suitable for application and approaches.

Feature extraction signature identification system can be classified into two types, namely the global features and local features.

1) Global Features: The global feature describes the overall signature such as length, width, density, dot the edge of the signature, and wavelet transformation. These features are less sensitive to noise and variations of the signature, so it will not provide high accuracy for skilled forgeries. The global feature would be suitable for this type of counterfeiting random and better when combined with other types of features.

2) Local Features: Local features describe a small area of the image signature and extract more detailed information, more accurate than global features but have high computational time. Based on the level of detail is considered, local features can be subdivided into: a feature-oriented components (ie the higher the ratio of the width of the stroke, the relative positions of the stroke, the orientation of the stroke, and others), and oriented pixels (ie based information grid, pixel density, gray-level intensity, texture, etc.). Phase identification/verification is the process of deciding whether the signature is genuine or fake. Several methods can be used in verifying a signature, such as Hidden Markov Model (HMM), Neural Networks, Support Vector Machine (SVM), K-Nearest Neighbors (k-NN). The most famous distance approach is Dynamic Time Warping (DTW), which is useful when the signature has a different length. Meanwhile, Hidden Markov Model (HMM) has long been used in the model-based approach [11]. k-Nearest Neighbors (k-NN) algorithm is also widely used for signature verification because the k-NN performance was excellent in pattern recognition systems. Excess k-NN because the method of decision-making reflects the way humans tend to only by the size of the spacing between samples designed by the researcher. K-NN also does not involve many parameters like other verification methods [9].

System performance identification/verification of the signature can be evaluated base on False Rejection Rate (FRR) and False Acceptance Rate (FAR). FRR measures some original signatures are considered as false, while FAR evaluate the amount of counterfeiting that is classified as genuine. The average value of the FAR and FFR is called the Average Error Rate (AER). FAR should be avoided in practical applications while the FRR must be tolerated. To deal with the FAR, the system must be tested against various classes of counterfeiting [12].

Research has been conducted by combining global features: height ratio of the width, density ratio and the ratio of the distance the edges, with local features: the number of pixels "0" per column and the number of a spatial symbol [13]. This research used k-NN method for verification.

Previous research also presented a series of new features based on the nature of the signature (the image in binary form) for off-line signature verification [14]. The proposed feature set illustrates the signature form regarding the spatial distribution of black pixels around the candidate pixel (on signature). This feature also provides texture size through correlation between signature pixels around the pixel of the candidate. Thus, the proposed feature set is unique in that it contains form and texture properties unlike most previously proposed features for off-line signature verification. The proposed feature is based on the idea of inherent problems and therefore feature evaluation by various feature selection techniques has also been attempted to obtain a series of compact features. To test the effectiveness of the proposed feature, two popular classifiers, multilayer perceptron and supporting vector engine are implemented and tested on two publicly available databases, ie GPDS300 corpus, and CEDAR signature databases. Test results showed 13.76% false rejection rate.

Another study performed the signature verification research with the aim to reduce fraud in financial transactions, security across international borders and boarding airplanes [15]. The study used a signature database of 50 Punjabis with 200 signatures for training and testing. Features extracted using Gabor filters and matching are done using SURF feature and critical point matching. This classification is based on HMM classification, and the experimental results show 97% accuracy verification rate.

Zuraidasahana Zulkarnain and colleagues presented a feature based on geometric concepts [16]. To achieve the goal, triangle attributes are utilized to design new features because triangles have orientation, angle, and transformation that will improve accuracy. The proposed feature uses triangulation of geometric arrangement consisting of the side, angle and triangular perimeter coming from the center of gravity of the signature image. For classification, Euclidean classifier along with voting based classification is used to verify the signature forgery tendencies. This classification process experimented using triangular geometric features and selected global features. Based on experiments that validated using the signature database Groupo de Senales 960 (GPDS-960), features a geometric triangle proposed reaches the level of the Average Error Rate (AER) with a percentage of 34% compared with 43% of the overall global selected.

II. Material and Method

Offline signature identification system requires signatures of data as an input. Subsequently, the signature is compared with the signature contained in the database to ascertain whether the signature is genuine. Offline signature verification system generally consists of three stages: data acquisition and pre-processing, feature extraction and verification. Offline signature verification stages of research can be seen in Fig. 1.

Fig. 1 Stages research offline signature verification

A. Data Acquisition and pre-Processing

This study used the signatures of 10 writers which writers affix respectively ten signatures on a piece of paper that had been provided. Those signatures were scanned with a resolution of 200 dpi and saved in .png format. The examples of those signatures are shown in Fig. 2.

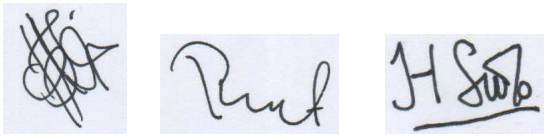


Fig. 2 The original signatures

The pre-processing stage was a necessary step to improve the accuracy of the feature extraction stage. Preprocessing was done to do noise removal, binarization, skeleton, and cropping.

1) Elimination of Noise, this process aimed to reduce noise in digital images. A median filter was used in this study. The results of this step are shown in Fig. 3.

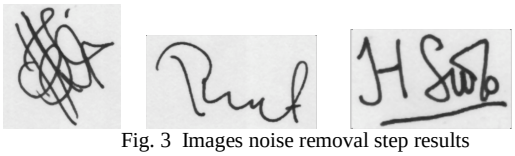


Fig. 3 Images noise removal step results

2) Binarization, the results of image noise removal step was converted into a binary image. The image in Fig. 4 shows the results of binarization of the image after noise removal and binarization procedure.

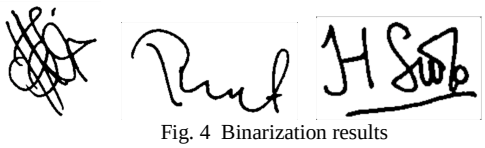


Fig. 4 Binarization results

3) Skeleton, changing the thickness of the results of binarization image into the image in the form of a framework to facilitate the calculation in the next step. Images of the skeleton are shown in Fig. 5.

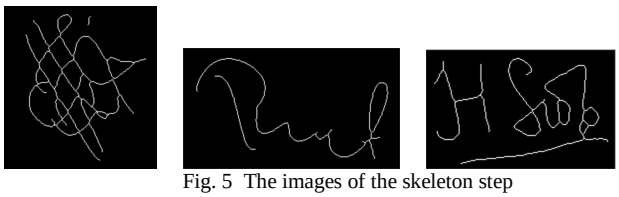


Fig. 5 The images of the skeleton step

4) Cropping, focus area pixel signatures from the edge of the left, right, top and bottom of the signatures. The results of this process are shown in Fig. 6.

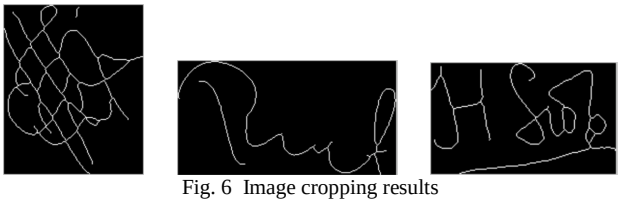


Fig. 6 Image cropping results

B. Feature Extraction

The output of the pre-processing stage will be used as input to the feature extraction stage. Feature extraction is the process of extracting information to represent the form by the classification. As well as the problems of other pattern recognition, feature extraction is a crucial step that significantly affects the performance of a signature verification system. Feature extraction technique must be able to tolerate differences in input variation signature form.

We used five features to distinguish the characteristics of each signature. We used a combination of global and local features, such as:

1) Height-width Ratio of Signatures: This ratio is obtained from the quotient of height to the width of the signature in pixels. High derived from the maximum number of pixels in one column of the signature image that has been in the crop. While the width is obtained from the maximum number of pixels in a row. It is shown in Fig. 7.

$$F1 = Height/Width \quad (1)$$

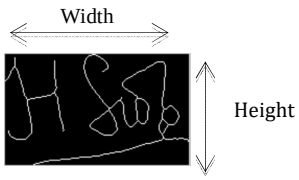


Fig. 7 The height and width signature

2) The Ratio of the Density of Signatures: This ratio is calculated from the number of pixels that are part of the signature only (J) divided by the total pixels of the image of the signature of the whole (T).

$$F2 = J/T \quad (2)$$

3) Edge Distance Ratio Signatures: After the crop, the border of signatures (left, right, top and bottom) can be

calculated the distance from the left edge. Pixel signature from the top leftmost pixel is L1, and the distance pixels signature of the lower left pixel is L2. It is shown in Fig. 8. The ratio of the distance the edge of the signature is:

$$F3 = L1 / L2 \quad (3)$$

Fig. 8 Distance edges L1 and L2 signatures

4) Accumulation and Proximity Ratio Column

This feature is obtained by accumulating pixel 0 to pixel 1 nearest per column from the top (A) and bottom of the signature (B), multiplied by the ratio of the density of the signature (F2).

$$F4 = A / B * F2 \quad (4)$$

5) The Number of Connected Components of Signatures: A signature results from pre-processing will have several components that are interconnected. Connectedness is seen from an 8-connected neighborhood.

The basic step in finding the connected components are:

1. Search for pixels that do not have a label, name p.
2. Use the flood-fill algorithm to label all pixels on a connected component containing p.
3. Repeat steps 1 and 2 until all the pixels are labeled.

C. Verification

The next stage after verification feature extraction stage. In this study, the authors used the k-NN method for classifying signatures. K-NN algorithm is a method to classify new objects based (*k*) nearest neighbors. Steps k-NN algorithm:

1. Determine the parameter *k* (the number of nearest neighbors). The optimum value *k* obtained by experiment
2. Calculate the Euclidean distance of each object on the data given sample.

$$d(q,x) = \sqrt{\sum_{i=1}^{n}(x_i - q_i)^2} \quad (5)$$

3. Sort the results of the calculation in step two of the smallest value (having the smallest Euclidean distance)
4. Classify classes based on the number of majority nearest *k*

III. Results and Discussion

The results of this research are discussed in this section. We collected 100 signatures from 10 persons which each person wrote 10 signatures. We used 6 signatures of each person as samples and 4 signatures of each person as test data.

Table 1 shows data of the extracted features from sample data. It shows that each person has differences of the height-width ratio of signature (F1), the ratio of the density of signatures (F2), edge distance ratio signatures (F3), accumulation and proximity ratio column (F4), and the number of connected components of signatures (F5) for each signature. It means that each person has a variety of signatures. For example, Person 1 has different values of F1, F2, F3, F4, and F5 from each of his signatures.

TABLE I
Feature Extraction of Samples

Person	F1	F2	F3	F4	F5
1	1.1837	0.0483	0.2584	0.0327	2
	1.0709	0.0526	0.3902	0.0389	2
	1.2654	0.0404	0.5177	0.0356	2
	1.1203	0.0457	0.5714	0.0452	2
	1.3241	0.0489	0.2069	0.0307	1
	1.2014	0.0524	0.3441	0.0368	2
2	0.5261	0.0247	0.1620	0.0386	2
	0.5023	0.0265	0.0941	0.0467	2
	0.6119	0.0223	0.0952	0.0476	2
	0.5370	0.0265	0.0313	0.0369	1
	0.6351	0.0218	0.0980	0.0398	1
	0.3971	0.0340	0.0051	0.0814	1
3	0.5922	0.0461	0.9643	0.1072	2
	0.6977	0.0426	1.3182	0.0687	5
	0.5337	0.0490	1.0526	0.1554	5
	0.6230	0.0422	2.1389	0.0791	3
	0.5440	0.0453	1.1026	0.1383	3
	0.5659	0.0453	2.7241	0.2077	3
4	1.2628	0.0393	1.6667	0.0947	1
	1.4074	0.0452	190	0.1400	1
	1.5616	0.0372	2.1429	0.1311	2
	1.4405	0.0341	1.9910	0.0821	1
	2.0684	0.0482	2.4318	0.1165	1
	1.5988	0.0347	2.1240	0.0708	1
5	0.5372	0.0340	0.9710	0.0247	3
	0.5561	0.0362	1.2031	0.0448	2
	0.6684	0.0368	1.0781	0.0469	1
	0.6851	0.0262	1.0127	0.0259	3
	0.8367	0.0346	0.5822	0.0597	1
	0.6816	0.0303	0.6207	0.0350	3
6	1.1942	0.0450	1.0513	0.0725	2
	1.1190	0.0563	0.3913	0.1384	2
	1.2674	0.0501	0.7273	0.1164	3
	1.3026	0.0588	0.3171	0.1360	2
	1.3182	0.0563	0.8235	0.1061	2
	1.8857	0.0519	6.7857	0.1078	2
7	1.1635	0.0399	1.3028	0.0370	2
	0.8704	0.0356	0.5892	0.0308	2
	1.2293	0.0343	1.3846	0.0108	2
	1.1443	0.0317	9.3333	0.0135	4
	1.0432	0.0425	2.5094	0.0251	2
	1.4000	0.0431	1.5244	0.0311	3
8	0.5754	0.0220	1.3971	0.0369	4
	0.4254	0.0240	5.3529	0.0421	4
	0.6107	0.0172	8.5385	0.0276	5
	0.4792	0.0162	0.8982	0.0305	4
	0.6677	0.0167	5.8800	0.0290	5
	0.7500	0.0172	6.5714	0.0223	6
9	0.9012	0.0259	2.7447	0.0314	2
	1.0214	0.0247	0.8068	0.0294	3
	1.0308	0.0302	1.7308	0.0375	2
	1.2865	0.0292	3.0233	0.0386	1
	0.9343	0.0306	2.6486	0.0320	1
	1.1667	0.0275	31.4290	0.0335	2
10	0.3919	0.0194	1.1205	0.0184	6
	0.3180	0.0228	1.1370	0.0231	8
	0.4386	0.0217	9.2500	0.0168	7
	0.3963	0.0213	1.0333	0.0176	6
	0.3700	0.0241	1.0779	0.0192	6
	0.2648	0.0234	0.7790	0.0211	6

We performed the same steps to get the feature extraction value from data testing. The results of feature extraction of data testing can be seen in Table 2.

TABLE II
FEATURE EXTRACTION OF TEST DATA

Person	F1	F2	F3	F4	F5
1	1.2632	0.0478	0.4886	0.0497	2
	1.1242	0.0496	0.3810	0.0280	1
	1.4286	0.0504	0.4362	0.0523	2
	1.2681	0.0512	0.3177	0.0432	2
2	0.6119	0.0231	0.0708	0.0409	1
	0.5854	0.0295	0.0066	0.0603	1
	0.6477	0.0241	0.1461	0.0480	1
	0.5809	0.0241	0.1150	0.0281	1
3	0.6552	0.0437	0.8125	0.1131	3
	0.6380	0.0493	1.0909	0.1508	3
	0.6168	0.0473	1.2432	0.1358	2
	0.6629	0.0420	1.3889	0.0759	4
4	1.88	0.0469	2.5542	0.1375	1
	1.6471	0.0374	2.3789	0.0901	1
	1.3526	0.0392	1.9490	0.1175	1
	1.4266	0.0445	2.1860	0.1248	2
5	0.7763	0.0300	1.1605	0.0323	2
	0.6414	0.0282	1.0923	0.0364	2
	0.7548	0.0331	1.4416	0.0365	2
	0.8522	0.0306	1.2651	0.0380	2
6	1.2019	0.0435	0.7347	0.0842	2
	1.4286	0.0628	0.8000	0.1421	2
	1.4021	0.0561	0.4237	0.1664	2
	1.6667	0.0572	1.2105	0.1354	2
7	1.7281	0.0436	1.7692	0.0124	3
	1.4126	0.0381	1.2525	0.0194	2
	1.1613	0.0403	3.2326	0.0224	2
	1.2143	0.0351	2.7500	0.0168	3
8	0.7003	0.0158	4.7667	0.0275	5
	0.6771	0.0144	12,700	0.0187	4
	0.5714	0.0153	4.9630	0.0240	5
	0.7893	0.0181	4.7419	0.0413	4
9	1.4025	0.0339	1.8654	0.0407	3
	1.1615	0.0305	1.6163	0.0551	2
	1.0000	0.0249	2.1538	0.0274	2
	1.1684	0.0250	1.3824	0.0563	2
10	0.4023	0.0182	4.0968	0.0125	6
	0.3352	0.0188	1.5938	0.0154	7
	0.3491	0.0211	0.9667	0.0205	7
	0.2914	0.0205	5.3889	0.0212	5

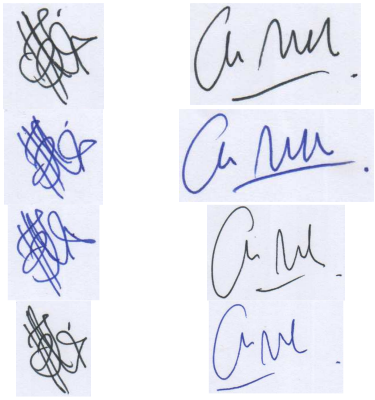
Euclidian Distance (ED) is used to calculate the distance between Table 2 and Table 1 which based on formula (4). ED is calculated from all feature extraction of all sample. For example, one of the ED between Table 2 and Table 1 is 0.2441.

TABLE III
VERIFICATION RESULT

Person	Data Testing	Identified as Person
1	1	1
	2	1
	3	1
	4	1
2	1	2
	2	2
	3	2
	4	2
3	1	5
	2	3
	3	5
	4	8
4	1	4
	2	4
	3	4
	4	4
5	1	5
	2	5
	3	5
	4	5
6	1	6
	2	6
	3	6
	4	7
7	1	7
	2	7
	3	9
	4	3
8	1	8
	2	7
	3	8
	4	8
9	1	7
	2	9
	3	7
	4	7
10	1	8
	2	10
	3	10
	4	8

All ED of all signature is sorted in ascending order. The smallest ED result is used to verify who the owner of the signature. Verification results are shown in Table 3. It shows that 28 signatures were correctly identified. The highlighted column in Table 3 shows falsely identified.

Table 3 shows that all signatures of 4 persons are verified correctly. Two persons have only one signature verified correctly. This is likely due to the inconsistency of the person in writing the signature. Examples of accepted and rejected signatures are shown in Fig. 8.



(A) (B)

Fig. 8 (A) Signature of Person 1, (B) Signature of Person 3

IV. Conclusion

This research conducted using k-NN to verify the owner of the signature. The results showed that the proposed identification system could identify 28 signatures of test data correctly. This research can be developed by adding another extraction features to increase the accuracy of signature verification.

References

[1] G. Sulong, A. Y. Ebrahim, and M. Jehanzeb, "Offline Handwritten Signature Identification Using Adaptive Window," *Signal Image Process. An Int. J.*, vol. 5, no. 3, pp. 13–24, 2014.

[2] E. J. R. Justino, F. Bortolozzi, and R. Sabourin, "The Interpersonal and Intrapersonal Variability Influences on Offline Signature Verification Using HMM," in *Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI '02)*, 2002, no. 64, pp. 197–202.

[3] D. Impedovo and G. Pirlo, "Automatic Signature Verification: The State of the Art," *IEEE Trans. Syst. MAN, Cybern. C Appl. Rev.*, vol. 38, no. 5, pp. 609–635, 2008.

[4] A. Kholmatov, "Biometric Identity Verification Using On-line & Off-line Signature Verification," Sabanci University, 2003.

[5] O. O-khalifa, M. K. Alam, and A. H. Abdalla, "An Evaluation on Offline Signature Verification using Artificial Neural Network Approach," in *2013 International Conference On Computing, Electrical and Electronic Engineering (ICCEEE)*, 2013, pp. 368–371.

[6] R. Jana, R. Saha, and D. Datta, "Offline Signature Verification using Euclidian Distance," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 1, pp. 707–710, 2014.

[7] Y. M. Al-omari, S. N. H. Sheikh Abdullah, and K. Omar, "State-of-the-Art in Offline Signature Verification System," in *2011 International Conference on Pattern Analysis and Intelligent Robotics*, 2011, no. June, pp. 59–64.

[8] M. R. Deore and S. M. Handore, "A Survey on Offline Signature Recognition and Verification Schemes," in *2015 International Conference on Industrial Instrumentation and Control (ICIC)*, 2015, pp. 165–169.

[9] A. N. Azmi, D. Nasien, and F. S. Omar, "Biometric Signature Verification System Based on Freeman Chain Code and k-Nearest Neighbor," *Multimed. Tools Appl.*, vol. 76, no. 14, pp. 15341–15355, 2016.

[10] S. Pal, U. Pal, and M. Blumenstein, "Signature-based Biometric Authentication," *Comput. Intell. Digit. Forensics Forensic Investig. Appl.*, vol. 555, pp. 1–32, 2014.

[11] R. A. Mohammed, R. M. Nabi, S. M. R. Mahmood, and R. M. Nabi, "State-of-the-Art in Handwritten Signature Verification System," in *2015 International Conference on Computational Science and Computational Intelligence, CSCI 2015*, 2015, pp. 519–525.

[12] B. Kovari and H. Charaf, "A Study on The Consistency and Significance of Local Features in Off-line Signature Verification," *Pattern Recognit. Lett.*, vol. 34, no. 3, pp. 247–255, 2013.

[13] S. Biswas, T. Kim, and D. Bhattacharyya, "Features Extraction and Verification of Signature Image using Clustering Technique," *Int. J. Smart Home*, vol. 4, no. 3, pp. 43–56, 2010.

[14] R. Kumar, J. D. Sharma, and B. Chanda, "Writer-Independent Off-line Signature Verification using Surroundedness Feature," *Pattern Recognit. Lett.*, vol. 33, no. 3, pp. 301–308, 2012.

[15] R. Kaur and P. Choudhary, "Handwritten Signature Verification Based on SURF Feature Using HMM," *Int. J. Comput. Sci. Trends Technol.*, vol. 3, no. 1, pp. 187–195, 2015.

[16] Z. Zulkarnain, M. S. Mohd Rahim, N. A. F. Ismail, and M. A. M. Arsad, "Triangular Geometric Feature for Offline Signature Verification," *Int. J. Comput. Electr. Autom. Control Inf. Eng.*, vol. 10, no. 3, pp. 543–546, 2016.