https://doi.org/10.48047/AFJBS.6.13.2024.6562-6573

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

A MULTI-CLASSIFICATION ADVANCED FEATURE EXTRACTION HYBRID METHOD USING CONVOLUTIONAL NEURAL NETWORK (CNN) AND HISTOGRAM OF ORIENTED GRADIENTS (HOG)

Manoj Tallapragada, Prasanth Yalla

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur 522302, AP, India

Volume 6, Issue 13, Aug 2024

Received: 15 June 2024

Accepted: 25 July 2024

Published: 15 Aug 2024

doi: 10.48047/AFJBS.6.13.2024.6562-6573

Abstract: How effectively these systems can tell the difference between real and fake marks depends on how many factors are taken out and how they are changed. This is why the part of a detached mark check system that extracts parts is so important to how well the system works in general. The study shows a mixed technique for offline signature verification systems that pull-out features from pictures of signatures. The plan uses both the Convolutional Neural Network (CNN) and Histogram of Oriented Gradients (HOG) methods to find important features. The next step is to use a decision tree feature selection method. Two sets of data (UTSig and CEDAR) and three models (KNN, SVM, and LSTM) are operated to exam the mixed method again. The findings of the tests showed that it was very accurate at telling the variation among real and fake signatures, even for expert frauds.

Index Terms *– Offline signature verification, CNN, HOG, SVM,deep learning.*

1. INTRODUCTION

Biometrics is the best main expertise for identifying people and judging their strength based on their unique body and behavior. The scribbled signature is one of the fingerprint identification methods that is most widely used around the world. The signatures written by hand are used as unique behavioral fingerprints in bank papers, credit cards, IDs, check processing, and financial documents. It's tough to check these signs, especially if they aren't clear. A method that can tell the difference between real and fake signatures is needed to stop theft and scams. A plenty of research has been done here in the last few years. From ML calculations to DL calculations and from well-qualified results to standard confirmation, we've come a long way. Even after a lot of research, making a separate signature check system work is still a lot of work. Based on the information shared, this study tries to figure out how to use both ML and DL models to tell the difference between real and fake marks in pictures of marks. It should also be possible for this mixed method to make different algorithms work better. We'll use this method to check signatures when we're not online.

This study suggests using hybrid feature extraction to check offline signatures and tries it. CNN and HOG are put together, and features are chosen by a Decision Tree method. The goal is to make many models better at telling the difference between real and fake signatures on two different datasets (UTSig and CEDAR), specifically when it derives to expert forgeries.

This work tries to solve the problem of why we need solid offline ways for checking signatures. The way signing systems work now often fails horribly to tell the difference between real and fake signatures, especially when professional copies are used. By creating a mixed feature extraction method and testing how well it works. The objective of this analysis is to make the system work better by using different algorithms on different datasets.

Offline signature verification reduces hacking and criminality by checking for false signatures [14]. By automatically validating the question's signature, signature analysis systems may also distinguish legitimate from false signatures. Multiple offline signatures checking techniques don't always function due to varying amounts of bogus signatures. They had to tell the difference between three types of fake signatures: random, easy, and skilled [15]. As for the first two structures, the fake signatures were made without any plan or knowledge of the real signatures' name or style. If someone makes a skilled fake, people think that person is an expert at copying the mark's shape and style and knows how the real one looks. Without moving elements [16], it is harder to spot fakes that are made by professionals.

2. LITERATURE REVIEW

2.1 Signature Forgery Detection Using Machine Learning:

Signatures are vital to contemporary society, yet they may be falsified. Examples include passports, bank checks, and licenses. This causes identity theft, hacking, and fake identities. To focus this problem, our analysis aims to employ CNN and DL to recognize authentic and fake signatures from a collection. CNN and DL are used because signatures fluctuate over time owing to behavioral changes like aging, mental health, and physical health. An algorithm that learns from several training datasets is needed to increase detection accuracy. Online and offline signature authentication are used. The offline signature forgery detection method underpins our proposal. Handwritten signatures on paperwork need a photo. Thus, image processing should be considered for this project. We want to improve accuracy by developing offline methods. Several studies use DL models to identify online and offline signature fraud.

2.2 Handwritten Signatures Forgery Detection Using Pre-Trained Deep Learning Methods:

Handwritten signature recognition (HSR) is crucial for document verification, authentication, financial transactions, banking transactions, and legal agreements. Common signature fraud threatens the integrity and security of various authentication mechanisms. The purpose of signature forgery detection (SFD) systems is to differentiate between legitimate and fraudulent signatures. This is tough, especially offline when dynamic signing process information is missing, and scanned signature pictures are required for signature identification. Recently, pre-trained deep learning (DL) models exhibit show off prevalent precisely to their excellent accuracy and short training time and computational resource needs. Images are processed using these models. Developers may save time by leveraging pre-trained models instead of starting from scratch. Thus, this paper compares pre-trained DL models for SFD. These methods provide great SFD accuracy. The MobileNet model is very accurate at 98.44%. Its tiny model and rapid training time are significant advantages. MobileNet is suitable for embedded systems and mobile devices because of its features.

2.3 An integrated approach on verification of signatures using multiple classifiers (SVM and Decision Tree): A multi-classification approach:

A handwritten signature is often used to identify and authenticate the writer. Automated identity verification is needed. In essence, the signature has dynamic elements and static aspects that fluctuate with place and time. Several academics have found ways to raise the signature verification system function extraction point. Digital, manual, or other signature verification methods are compared in the research. The signature attributes were identified using the best ML methods (DT and SVM). Features were also presented after impact assessment. Research examined numerous language databases. The feature improved accuracy.

2.4 Recent developments in pretreatment technologies on lignocellulosic biomass: Effect of key parameters, technological improvements, and challenges:

Lignocellulosic biomass is a cheap renewable resource for biofuels and bioproducts. Due of biomass's resistance, enzymes and bacteria have trouble reaching polysaccharides. Several pretreatment methods may turn lignocellulosic biomass into valuable products. However, these pretreatment methods yield several microbe- and enzyme-inhibiting secondary compounds. As mentioned in the review, an effective, process-optimized pretreatment method may reduce inhibitory compound synthesis and boost fermentable sugar and biochemical output. Additionally, genetic engineering and evolution are being used to boost microbe inhibitor resistance. Pretreatment and detoxification may boost lignocellulose biorefinery output. We describe the newest inhibitor removal and lignocellulosic biomass pretreatment methods in this research.

2.5 Offline Handwritten Signature Verification Using Deep Neural Networks:

Students' handwritten signatures on attendance sheets were the main way to authenticate their presence in class before computerized methods. For short courses or areas lacking alternate processes, the technique is used. However, handwritten signature verification is tedious. This research describes two ways to verify attendance sheet signatures. One visual spot recognition method simply checks for a signature. The alternative technique uses a multiclass convolutional neural network exceptional by AlexNet to recognize the signature originator with over 85% accuracy and recall after training on a little actual training data. A higher number of legitimate signatures and data augmentation improves signature confirmation accuracy.

3. METHODOLOGY

Existing methods check signatures without being online by using a set of geometric properties with simple shapes. For example, the Baseline Slant Angle (BSA), Aspect Ratio (AR), Normalized Area (NA), and the Center of Gravity and Slope of the line linking the two picture pieces' Centers of Gravity make up a signature. From the start, a list of people's grades that the system is hypothetical to check is managed to set up the system. This means that a mark is used as a guide for verification after being assessed to a reliable test signature. To measure how similar two fingerprints are in the component space, the Euclidean distance is used. If the Euclidean distance is less than a certain cutoff, which means it is close to the lowest level of similarity that is allowed, then the test signature is proven to belong to the stated person. It is known to be a fake if it isn’t. This research gives details on the traits that were stated, as well as the pre-processing, execution, and results.

3.1 Drawbacks:

- The method described in the works only uses a few geometry properties. There's a chance that these traits won't be able to tell the difference between people, or they could lose some significant information that are vital for accurate and reliable signature verification. Adding more complex or texture-based features could make the system work better.
- The signature preparation stage, which splits parts and gets rid of noise, has a big effect on how well the signature verification system works. If the preprocessing step isn't strong enough or doesn't take into account all the different types of variation or noise, it could affect how accurate and reliable the feature extraction process is.
- An authentication method based on prototypes could not be competent to simplify closely whilst challenged with new signatures or signatures from different people.

The present study presents a novel approach to feature extraction from signature photos using a hybrid method that integrates CNN and HOG approaches. The relevant features are then identified through a feature selection algorithm that employs decision trees. To make offline signature verification systems more accurate and useful, the objective is to find significant contrasts that can be utilized to differentiate between real and fake marks.

LSTM, SVM , and KNN were the three models used to test the new technique on two datasets: UTSig and CEDAR.

3.2 Benefits:

1. There are a lot of helpful parts in the combination model, and it might work better with a low-complexity prediction.
2. Using three models from DL and ML will help show that the mixed method used to find highlights works.
3. The evaluation of multiple classifiers and datasets makes the suggested method more useful and durable.
4. Effective Extraction of Features: HOG features save local gradient data, but CNNs are eminent for existing able to instantly learn hierarchical features from raw image data. This mix makes feature images better by including both global and local traits that are typical.
5. The investigation is greatly about utilizing feature selection algorithms by decision trees to discover the most useful traits for classifying things. This approach might help you hit the target more accurately.

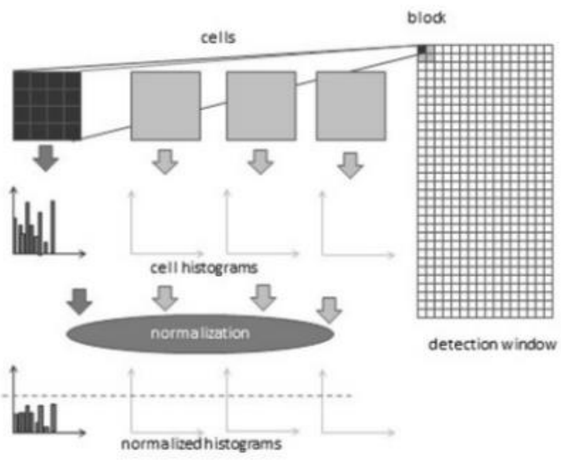

Figure 1: System Architecture

4. MODULES:

We created the following modules for the project.

- Data exploration: Data will be imported using this module.
- Processing: This module reads and processes data.
- Splitting data into train & test: This unit will split the data into train and test.
- Model generation: Model building -CNN, Feature Extraction using HOG, Feature Extraction using CNN and HOG with Feature Selection using DT with RFE, SVM, KNN, LSTM, Voting Classifier (RF + DT)
- User signup & login: You may register and log in using this module.
- User input: This module aids forecasting.

- Prediction: Last forecast

Note: To make things even better, we used a CNN model and a CNN with HOG feature, which is a HOG-based feature extraction method. They also said to use SVM, KNN, and LSTM models for feature extraction and data analysis. We got 95,95.5, and 91.3% accuracy, respectively. We also used a Voting Classifier to investigate the dataset, and feature extraction based on the Voting Classifier gave us 100% accuracy.

5. IMPLEMENTATION

The equations that go with this job are being used.

CNN: CNN says that a convolutional neural network is a type of DL that is meant to look into visual and spatial input. It uses convolutional layers to naturally gain and focus different leveled features from input data by identifying patterns and spatial relationships in the data. This lets it do tasks like object recognition, picture recognition, and arrangement with great accuracy.

Feature Extraction using HOG: Using HOG for feature extraction is a way of using computer vision to measure local picture gradient directions in order to describe and record information about images' structure and shape. It gives a short summary of visual data that can be used in many situations, mainly in image processing and computer vision jobs. In many situations, it is employed to retrieve and identify objects.

RFE: When CNN and HOG are used together in feature extraction, they pick up on rich visual features. DT with RFE is used to choose the features. Then, a Recursive Feature Elimination (RFE) method based on Decision Trees is used to choose which features to use. This mixed method lowers the number of dimensions and improves pattern recognition, which makes image-based tasks like recognizing items and checking signatures more effective and efficient.

Support Vector Machine (SVM): supervised machine learning method SVM can be employed for both regression and classification. It finds a hyperplane in a high-dimensional space and divides data points into classes in the best way possible. SVM works well with difficult data that doesn't have straight lines between the values because kernel functions change the data into extreme magnitude that make separation easier.

KNN: The math behind K-Nearest Neighbors. The KNN method, which is also written as KNN or k-NN, is a non-parametric supervised learning predictor that groups news stories together based on how similar they are in succession to group or predict them.

LSTM: Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) design that is used in DL. Because it is prepared to describe sequential data, it can remember information across long sequences and catch long-term relationships. LSTMs are frequently used for jobs as natural language processing and time series analysis.

Classifier for Voting (RF + DT): A Voting Classifier takes the results from several ML models and puts them all together to decide. It takes the predictions from a Decision Tree (DT) and a Random Forest (RF) algorithm and puts them together here. When the Voting Classifier picks the class that most of the component models predict, it often makes the total estimate more stable and accurate.

6. RESULTS &CONCLUSION

The study ends with a CNN-HOG hybrid feature extraction method and a feature selection algorithm for systems that check signatures offline. Three classifiers—LSTM, SVM, and KNN—were used for the study. The tests showed that our suggested model worked great in conditions of momentum and its capacity to guess what might happen right away. It was also very good at telling the difference between a real and a fake signature, even for pretty good fakes. The UTSig dataset and the CEDAR dataset were used to get this done with a good level of precision. The study stresses how important it is to identify features in offline signature verification systems and says that more research in this area could improve performance and forecasts.

Here are charts that display how the CEDAR and UTSig datasets compare.

6.1 Accuracy Comparison graph of CEDAR Dataset's-

The X-axis shows the accuracy score, and the Y-axis shows the algorithms.

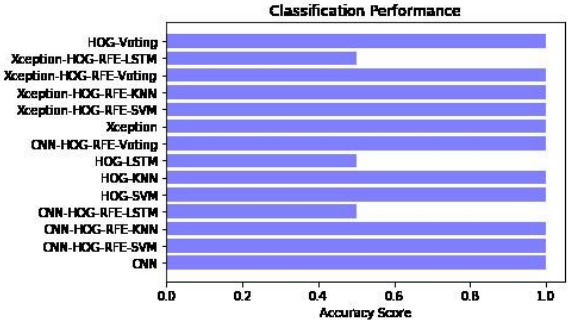

Figure 2: Classification Performance-1

6.2 Precision Comparison graph of CEDAR Dataset's-

The X-axis shows the precision score, and the Y-axis shows the algorithms.

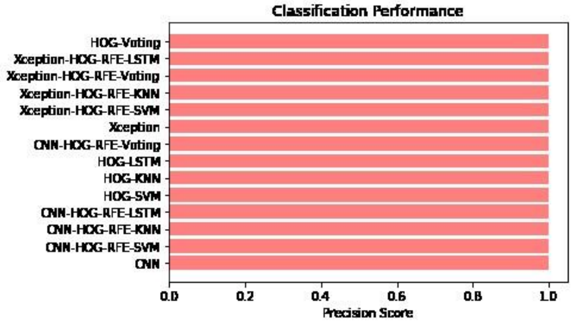

Figure 3: Classification Performance-2

Recall Comparison graph of CEDAR Dataset's-

The X-axis shows the recall score, and the Y-axis shows the algorithms.

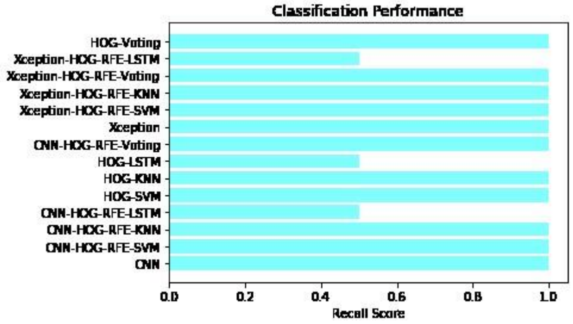

Figure 4: Classification Performance-3

F1 Comparison graph of CEDAR Dataset's-

X-Axis representsF1 score, and Y-Axis represents Algorithms.

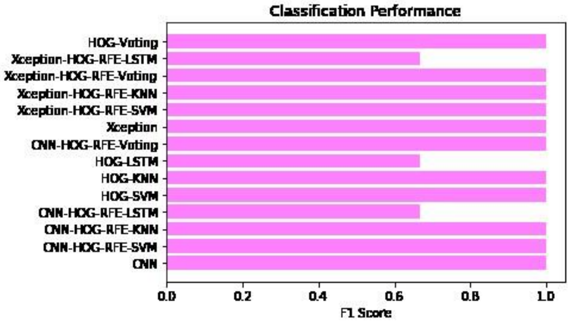

Figure 5: Classification Performance-4

Accuracy Comparison graph of UTSigDataset's-

X-Axis represents Accuracy score, and Y-Axis represents Algorithms.

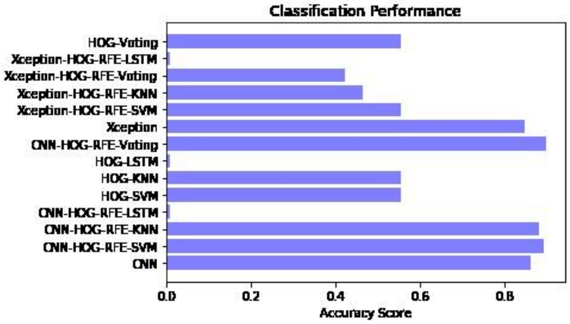

Figure 6: Classification Performance-5

Precision Comparison graph of UTSigDataset's-

X-Axis represents Precision score, and Y-Axis represents Algorithms.

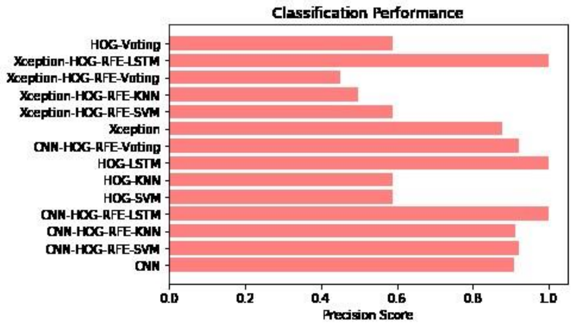

Figure 7: Classification Performance-6

Recall Comparison graph of UTSigDataset's-

X-Axis representsRecall score, and Y-Axis represents Algorithms.

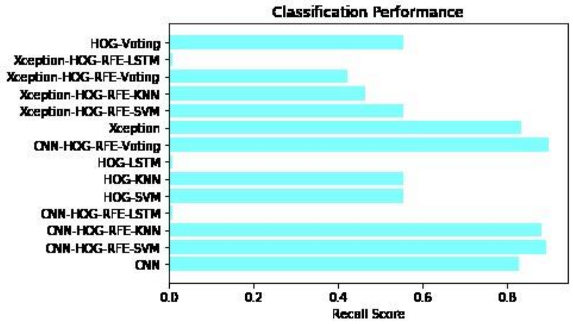

Figure 8: Classification Performance-7

F1 Comparison graph of UTSigDataset's-

X-Axis representsF1 score, and Y-Axis represents Algorithms.

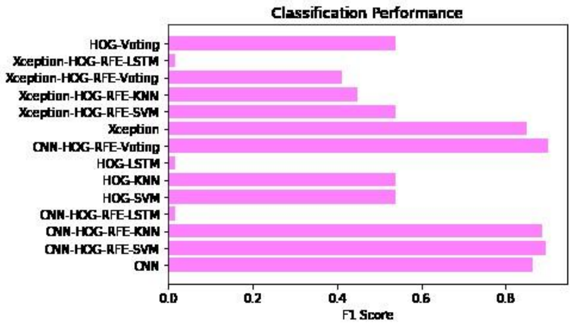

Figure 9: Classification Performance-8

REFERENCES

[1] F. M. Alsuhimat and F. S. Mohamad, ''Offline signature verification using long short-term memory and histogram orientation gradient,'' Bull. Elect. Eng. Inform., vol. 12, no. 1, pp. 283–292, 2023.

[2] M. Ajij, S. Pratihar, S. R. Nayak, T. Hanne, and D. S. Roy, ''Off-line signature verification using elementary combinations of directional codes from boundary pixels,'' Neural Comput. Appl., vol. 35, pp. 4939–4956, Mar. 2021, doi: 10.1007/s00521-021-05854-6.

[3] A. Q. Ansari, M. Hanmandlu, J. Kour, and A. K. Singh, ''Online signature verification using segment-level fuzzy modelling,'' IET Biometrics, vol. 3, no. 3, pp. 113–127, 2014.

[4] K. Cpałka and M. Zalasiński, ''On-line signature verification using vertical signature partitioning,'' Expert Syst. Appl., vol. 41, no. 9, pp. 4170–4180, 2014.

[5] K. Cpałka, M. Zalasiński, and L. Rutkowski, ''A new algorithm for identity verification based on the analysis of a handwritten dynamic signature,'' Appl. Soft Comput., vol. 43, no. 1, pp. 47–56, Jun. 2016.

[6] E. Griechisch, M. I. Malik, and M. Liwicki, ''Online signature verification based on Kolmogorov–Smirnov distribution distance,'' in Proc. 14th Int. Conf. Frontiers Handwriting Recognit., Sep. 2014, pp. 738–742.

[7] N. Sae-Bae and N. Memon, ''Online signature verification on mobile devices,'' IEEE Trans. Inf. Forensics Security, vol. 9, no. 6, pp. 933–947, Jun. 2014.

[8] S. Chen and S. Srihari, ''A new off-line signature verification method based on graph matching,'' in Proc. Int. Conf. Pattern Recognit. (ICPR), 2006, pp. 869–872.

[9] M. A. Ferrer, J. B. Alonso, and C. M. Travieso, ''Offline geometric parameters for automatic signature verification using fixed-point arithmetic,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 6, pp. 993–997, Jun. 2005.

[10] Y. Guerbai, Y. Chibani, and B. Hadjadji, ''The effective use of the oneclass SVM classifier for handwritten signature verification based on writerindependent parameters,'' Pattern Recognit., vol. 48, no. 1, pp. 103–113, 2015.

[11] R. Larkins and M. Mayo, ''Adaptive feature thresholding for off-line signature verification,'' in Proc. 23rd Int. Conf. Image Vis. Comput. New Zealand, Nov. 2008, pp. 1–6.

[12] H. Lv, W. Wang, C. Wang, and Q. Zhuo, ''Off-line Chinese signature verification based on support vector machines,'' Pattern Recognit. Lett., vol. 26, no. 15, pp. 2390–2399, Nov. 2005.

[13] Y. Serdouk, H. Nemmour, and Y. Chibani, ''New off-line handwritten signature verification method based on artificial immune recognition system,'' Expert Syst. Appl., vol. 51, pp. 186–194, Jun. 2016.

[14] F. E. Batool, M. Attique, M. Sharif, K. Javed, M. Nazir, A. A. Abbasi, Z. Iqbal, and N. Riaz, ''Offline signature verification system: A novel technique of fusion of GLCM and geometric features using SVM,'' Multimedia Tools Appl., pp. 1–20, Apr. 2020, doi: 10.1007/s11042-020-08851-4.

[15] F. M. Alsuhimat and F. S. Mohamad, ''Histogram orientation gradient for offline signature verification via multiple classifiers,'' Nveo-Natural Volatiles Essential OILS J., vol. 8, no. 6, pp. 3895–3903, 2021.

[16] N. M. Tahir, N. Adam, U. I. Bature, K. A. Abubakar, and I. Gambo, ''Offline handwritten signature verification system: Artificial neural network approach,'' Int. J. Intell. Syst. Appl., vol. 1, no. 1, pp. 45–57, 2021.

[17] A. B. Jagtap, D. D. Sawat, R. S. Hegadi, and R. S. Hegadi, ''Verification of genuine and forged offline signatures using Siamese neural network (SNN),'' Multimedia Tools Appl., vol. 79, nos. 47–48, pp. 35109–35123, Dec. 2020.

[18] B. Kiran, S. Naz, and A. Rehman, ''Biometric signature authentication using machine learning techniques: Current trends, challenges and opportunities,'' Multimedia Tools Appl., vol. 79, no. 1, pp. 289–340, 2020.

[19] M. Sharif, M. A. Khan, M. Faisal, M. Yasmin, and S. L. Fernandes, ''A framework for offline signature verification system: Best features selection approach,'' Pattern Recognit. Lett., vol. 139, pp. 50–59, Nov. 2020.

[20] N. Sharma, S. Gupta, and P. Metha, ''A comprehensive study on offline signature verification,'' in Proc. J. Phys., Conf., 2021, Art. no. 012044, doi: 10.1088/1742-6596/1969/1/012044.

[21] H. H. Kao and C. Y. Wen, ''An offline signature verification and forgery detection method based on a single known sample and an explainable deep learning approach,'' Appl. Sci., vol. 10, no. 1, p. 3716, 2020.

[22] M. K. Kalera, S. Srihari, and A. Xu, ''Offline signature verification and identification using distance statistics,'' Int. J. Pattern Recognit. Artif. Intell., vol. 18, no. 7, pp. 1339–1360, 2004.

[23] 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.

[24] T.-A. Pham, H.-H. Le, and N.-T. Do, ''Offline handwritten signature verification using local and global features,'' Ann. Math. Artif. Intell., vol. 75, nos. 1–2, pp. 231–247, Oct. 2015.

[25] Z. ZulNarnain, M. S. Rahim, N. F. Ismail, and M. M. Arsad, ''Triangular geometric feature for offline signature verification,'' Int. J. Comput. Inf. Eng., vol. 10, no. 3, pp. 485–488, 2016.

[26] R. K. Bharathi and B. H. Shekar, ''Off-line signature verification based on chain code histogram and support vector machine,'' in Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI), Aug. 2013, pp. 2063–2068.

[27] V. Nguyen, Y. Kawazoe, T. Wakabayashi, U. Pal, and M. Blumenstein, ''Performance analysis of the gradient feature and the modified direction feature for off-line signature verification,'' in Proc. 12th Int. Conf. Frontiers Handwriting Recognit., Nov. 2010, pp. 303–307.

[28] 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, Feb. 2012.

[29] M. Hanmandlu, M. H. M. Yusof, and V. K. Madasu, ''Off-line signature verification and forgery detection using fuzzy modeling,'' Pattern Recognit., vol. 38, no. 3, pp. 341–356, 2005.

[30] N. Jiang, J. Xu, W. Yu, and S. Goto, ''Gradient local binary patterns for human detection,'' in Proc. IEEE Int. Symp. Circuits Syst. (ISCAS), May 2013, pp. 978–981.