SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I3P113

Volume 13 Issue 3, 163-174, March 2026

Original Article

Fusion of Deep Learning and Classical Learning for Offline Signature Verification

Bhavani S D[1], Bharathi R K[2]

[1,2]Department of Computer Applications, JSS Science and Technology University, Mysuru, Karnataka, India.

[1]Corresponding Author : bhavani.sd@jssstuniv.in

Received: 10 December 2025 Revised: 12 January 2026 Accepted: 15 February 2026 Published: 23 March 2026

Abstract *- As the technology is growing at a faster pace, authentication becomes an integral part of everyday activities to minimize fraudulent activities. Handwritten offline signature verification is considered one of the reliable methods for authenticating an individual's identity, widely used in financial transactions, legal documents, business contracts, and so on. Handwritten Offline Signature Verification process involves three major phases: signature acquisition, feature extraction, and Verification. Within these stages, feature extraction plays a vital role because the outcome of the verification process is reliant on the quality of features extracted from the signatures. In this study, a hybrid approach based on a convolutional neural network (CNN) is introduced, where a lightweight CNN is employed to extract detailed and discriminative features from signature images. By convention, all deep neural networks extract high-dimensional features, which may contain features that do not contribute to the verification process. To overcome this issue, a correlation coefficient-based feature selection method is used to eliminate all non-contributing features, thereby reducing the computational complexity. In the final stage, the chosen discriminative features are supplied to the voting classifier for discriminating authentic and fake signatures. The five benchmark datasets, including CEDAR, BHSig260 (Bengali and Hindi), MCYT-75, and UTSig, are utilized to evaluate the proposed model. To enhance the model's transparency, the explainable AI technique LIME (Local Interpretable Model-Agnostic Explanations) is employed to identify the features that influence the model's decisions. Additionally, Grad-CAM (Gradient-Weighted Class Activation Mapping) is employed to visualize the specific parts of the signature image from which the CNN derives its features.*

Keywords *- SSV, LIME, Grad-CAM, Voting Classifier, CNN, Correlation Coefficient.*

1. Introduction

Security is a significant concern, as fraudulent activities are increasing with the evolution of technology. One of the most secure means of authentication is biometrics. Among all the biometrics, signature is the most widely accepted one from ancient times. Handwritten signatures are commonly used in bank cheques, legal documents, business contracts, and so on. In order to prevent fraud, handwritten signatures need to be verified before proceeding further. Static or Offline Signature Verification (SSV) is more challenging because it lacks the dynamic information, such as speed, pressure, and stroke order, that is available in online signatures. [27] In general, the automatic SSV process begins with acquiring signatures, extracting features, and classification. The significance of feature extraction lies in the fact that the model's effectiveness is directly affected by the characteristics of the features selected for training. In Conventional Machine Learning (ML), domain knowledge is required to extract proper features from the signatures. Some well-known, conventional methods rely on manually designed descriptors to extract texture, shape, and statistical properties from the given data. Although these techniques are computationally efficient and interpretable, they might struggle when there are variations in the signature, which is quite common in the case of handwritten signatures. Conventional methods often fail to capture complex intra-personal variations and inter-personal similarities, which leads to poor performance. Moreover, handcrafted features such as texture, shape, or gradient are easy to capture with less computational requirement. However, a significant challenge here is to choose a proper descriptor that generalizes well on a variety of signatures. In contrast, Deep Learning (DL) approaches remove the need for manual feature engineering, as they automatically learn discriminative representations from raw input data. Among these methods, Convolutional Neural Networks (CNNs) and Transformer architectures have been particularly impactful, as they learn feature representations automatically directly from data. In many computer vision and pattern recognition problems, DL models are successfully used to learn feature representations without prior domain knowledge. [6, 8, 9, 10, 23]. Deep neural Networks like CNNs are designed to capture intricate features from data, which often results in a high-dimensional feature vector, leading to overfitting and computational complexity. One possible solution to this issue is to use a shallow CNN made up of fewer hidden layers, making them particularly suitable for SSV tasks where the availability of training samples is restricted. Although shallow CNNs resolve the issue of computational complexity, they suffer from high-dimensional features that contain non-contributing and redundant features. In such cases, applying feature selection can improve generalisation, reduce overfitting, and enhance computational efficiency. Correlation coefficient-based feature selection is a simple yet efficient method that identifies features that show a prominent degree of association with the class label. The correlation coefficient method reduces the feature dimension, thereby reducing computation complexity. Finally, the filtered features are used for discriminating between authentic and forged signatures. A voting classifier reduces the risk of relying on a single classifier for prediction. The use of a shallow CNN to learn discriminative representations with reduced complexity, followed by a correlation coefficient Feature Selection Method(FSM) to retain the features that contribute to the classification, leading to improved efficiency, and finally a voting classifier that combines predictions from multiple learners, is a perfect fusion of DL and conventional ML methods used in this paper.

To add transparency to the model, two explainable AI approaches are employed. A Local Interpretable Model-Agnostic Explanations (LIME) is used to examine the contribution of each feature towards the prediction of the model. A Gradient-weighted Class Activation Mapping (Grad-CAM) helps in displaying the areas of the signature images that are considered significant by the CNN.

2. Literature Survey

In the area of biometric and pattern recognition systems, the effectiveness of the model largely depends on the right choice of features. Hence, the derivation of features plays a significant role in modifying the unprocessed input data into discriminative features that are apt for classification. Researchers have vigorously exploited both conventional handcrafted approaches (e.g., HOG, LBP, GLCM, SIFT, SURF) and DL methods (e.g., CNNs, Transformers). [3, 5, 7, 11, 13, 18, 19, 22, 26] A significant limitation of handcrafted features lies in the challenge of selecting an appropriate feature descriptor. To address this issue, recent research has increasingly shifted towards DL approaches, which enable the automatic learning of discriminative representations directly from unprocessed data, thereby removing the reliance on handcrafted feature engineering. Moreover, deep architectures can capture both global structures and fine-grained local details, enabling more robust handling of intra-writer variations and inter-writer similarities.

The authors Muhtar et al. [16], collected ethnic people's signatures as there was no static handwritten signature dataset for traditional people, consisting of 38,400 signature pictures. They combined ResNet18 and the CBAM, resulting in a more reliable and flexible method capable of handling differences. The authors proposed a novel architecture to learn both global and regional representations. A new loss function (Co-tuplet) was used to address intra-writer variability and inter-writer similarity. The authors Parcham et al. [17] They used the CNN and Capsule Neural Network(CapsNet) together. The application of CapsNet enabled a reduction in the count of layers, thereby decreasing the structural sophistication of the network.

The extraction of a large set of features from signature images often leads to increased computational complexity, while also introducing noise that can degrade model performance. Consequently, eliminating redundant features that do not contribute to class discrimination becomes essential. To address this, numerous approaches have been proposed in the literature for identifying and removing non-informative features. The authors Abdulhussien et al. [1], combined texture (LBP, EDMs), interest point (SURF), and curvelet transform features to capture both global and local characteristics. This hybrid approach increases robustness against intra-class variations (same signer) and inter-class similarities (skilled forgeries). Feature fusion is done via Early Serial Concatenation Fusion that Integrates multiscale features into a unified representation.

The GA is used to select the most discriminative features from the fused feature space. To tackle the issue of limited forged samples, One-Class SVM is used, which is well-suited for real-world SSV. The authors Banerjee et al. [4], employed a comprehensive feature extraction strategy (statistical, shape, similarity, and frequency features) based on SVD-transformed signals. A novel Binary Red Deer Algorithm (BRDA) is introduced for wrapper-based feature selection, which effectively reduces dimensionality while enhancing classification accuracy compared to existing meta-heuristics. Furthermore, it is validated under both signer-dependent and signer-independent scenarios, and results indicate superior performance compared to deep feature baselines. The authors Batool et al. [5], proposed the use of 22 Grey Level Co-occurrence Matrix (GLCM) characteristics and 8 structural characteristics for SSV.

A new method called High-Priority Index Feature (HPFI) parallel fusion approach is introduced to join these two feature sets. The authors propose a Skewness-Kurtosis controlled PCA (SKcPCA) technique to select the most discriminative features. This method leverages higher-order statistical measures to improve feature relevance and reduce redundancy, thereby enhancing classification efficiency. The system employs a SVM using the Radial Basis Function kernel for final classification between authentic and fake signatures. The authors Sivaiah et al. [21], introduced a novel feature extraction framework that integrates Convolutional Neural Network (CNN) features with Histogram of Oriented Gradients (HOG) descriptors.

A Decision Tree-based Recursive Feature Elimination (RFE) method is used for selecting the most discriminative features, thereby reducing redundancy and improving classification performance. Three classifiers, namely "Support Vector Machine (SVM), K-Nearest Neighbours (KNN), and Long Short-Term Memory (LSTM)", are employed for classification. Additionally, a Voting Classifier joining Random Forest (RF) and Decision Tree (DT) is tested, demonstrating superior robustness. The authors in the paper [28], fused deep features extracted using SigNet and handcrafted features (GLCM, HOG, LBP, SIFT). The combined features are sent to SVM for verification.

2.1. Research Gap

From the literature survey, it is evident that CNNs can extract more intricate features from signature images compared to conventional feature descriptors like GLCM, LBP, HOG, etc. However, when CNNs are used for both feature extraction and classification, their performance declines, mainly when a smaller number of hidden layers is used. Moreover, applying feature selection methods to determine contributing features from the CNN-generated feature vector remains unexplored in the literature. Similarly, the leverage of a voting classifier in place of a solo classifier has received less focus from the researchers. Based on the author's current knowledge, no prior research has combined DL (for feature extraction) and conventional ML techniques (feature selection and classification), aiming to achieve efficient verification results.

3. Proposed Methodology

To address the tricky challenge of SSV, this study introduces a hybrid method built on a Convolutional Neural Network (CNN) framework. CNNs serve as the feature extractor, capturing high-level and informative patterns from signature samples. Since CNN models generate high-dimensional feature vectors, there is a chance of the presence of redundant and non-contributing features, which results in an overfitting issue. Consequently, the informative features that drive the model's decisions must be retained through a correlation-coefficient feature selection approach.

In the verification phase, an ensemble voting classifier is used, which integrates SVM, KNN, and RF to improve decision reliability. By integrating multiple classifiers, the ensemble benefits from their combined strengths, ultimately boosting verification accuracy. To further ensure interpretability, the framework employs two XAI methods, LIME and Grad-CAM, to offer insight into the decision procedure of the model. The LIME is used to analyze the contribution of selected features, while Grad-CAM reveals the areas of signature images that the CNN prioritizes during feature extraction. Figure 1 depicts the proposed methodology.

3.1. Feature Extraction

Extracting discriminative features is the crucial stage in the SSV problem. DL techniques offer automatic feature extraction methods, such as the Convolutional Neural Network (CNN). These CNNs automatically learns intricate features directly from the raw signature images. In a lightweight CNN, three convolutional layers are used first, and then pooling layers are applied; at the end, a dense layer is used. Initially, all the images are converted to (128x128) size images and fed to the input layer. In the beginning, convolutional layers with 16 filters of size 3x3 are utilised to capture low-level features, edges, or blobs, and the ReLU is employed to introduce non-linearity to the model. This is followed by the first max pooling layer, which reduces the spatial dimension by half to improve computational efficiency. In the second layer, 32 filters having the same size and activation function as the previous layer are applied to extract mid-level representations. The third layer then expands to 64 filters, again matching the earlier filter size and activation function, to learn more abstract, high-level features. Since adding more convolutional layers did not yield any performance gains, the three convolutional layers are sufficient to capture intricate features from the signature image. At the end, a fully connected layer having 128 neurons and a Sigmoid activation function is included in order to learn the higher-level representations produced by the preceding layers. The dense layer is named to extract features after training the CNN model. In total, 128 features are extracted from each signature image.

3.2. Feature Selection

CNNs learns discriminative features, but the resultant feature vector is high-dimensional, which may contain redundant information. Hence, feature selection methods can highlight the most discriminative features, improving classifier stability. In this work, a correlation coefficient–based FSM is used to remove redundant features and those that exhibit strong inter-feature correlations.

Correlation Coefficient Feature Selection (CFS) is a technique that assesses the strength of association between individual features and the target variable, allowing the detection and selection of the predominant informative characteristics for an ML model. CFS is built upon the principle that good feature subsets contain features that: 1) are highly correlated with the target variable (i.e., they provide helpful information for predicting the target). 2) have low intercorrelation with each other (i.e., they provide unique information). The core concept involves calculating the correlation of each feature with the target variable and retaining those features that exhibit high absolute correlation values. If a feature is highly correlated with the target variable, then those features are retained, and the rest of the features are filtered out. Because signature images exhibit non-linear characteristics, "Spearman's Rank Correlation Coefficient" is employed to identify the most significant features. This coefficient is a non-parametric measure of rank correlation. This is used to assess the association between two variables. Spearman's correlation coefficient (ρ) is computed as follows:

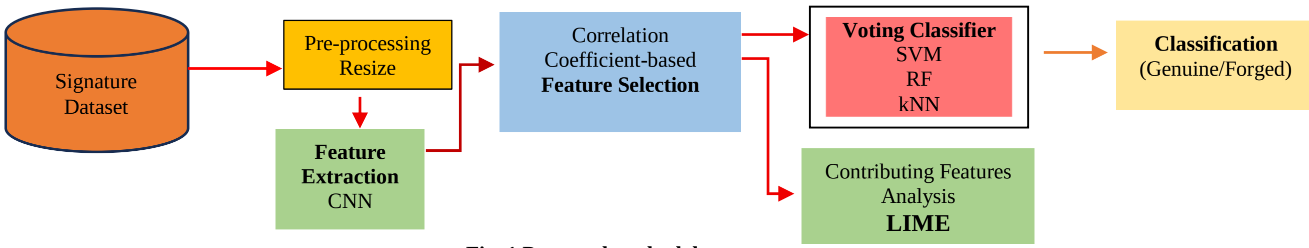

Fig. 1 Proposed methodology

$$\rho = 1 - 6\Sigma d_i^2 / n(n^2 - 1) \tag{1}$$

where d_i corresponds to the difference in ranks assigned to each pair of observations, and n indicates the total count of observations. The correlation coefficient ρ takes values in the interval $[-1,1]$. A coefficient value of 1 denotes complete positive monotonicity, implying that an increase in one variable corresponds to an increase in the other. A coefficient of -1 denotes complete negative monotonicity, indicating an inverse ordering between the two variables. A coefficient value of 0 denotes a lack of monotonic dependence between the variables.

3.3. Explainable AI

3.3.1. Local Interpretable Model-Agnostic Explanations (LIME)

LIME is an Explainable AI (XAI) that helps us understand the decisions made by any ML model. Here, LIME is used to understand the predictions of the Voting Classifier. It helps us to understand why Voting Classifier has classified a single signature as authentic or fake. It provides the importance of features around a particular prediction.

LIME helps us to understand why a particular signature is predicted as "Authentic" or "Fake". LIME shows which features are the top ones that pushed the prediction toward "Forged" or "Genuine", and how much each feature contributed to the decision for a randomly picked signature.

3.3.2. Gradient-Weighted Class Activation Mapping (Grad-CAM)

Grad-CAM highlights the areas of a signature picture that the CNN deems important during feature extraction. During the forward pass, the input image is propagated through the network to produce a class score(e.g., *genuine* vs. *forged*). In the backpropagation stage, gradients corresponding to the target class score are evaluated with respect to the feature maps in the last convolutional layer. These gradients are subsequently averaged over the spatial dimensions to derive weights for each feature map. The averaged gradients are then multiplied by their respective feature maps to generate a class-discriminative visualization. The regions that support the prediction are retained, and the weighted map is unsampled and overlaid on the original image.

3.4. Voting Classifier

This approach uses an ensemble voting classifier because it takes the best parts of several learning algorithms and puts them together. This makes them more general and less likely to be biased than a single model. The reason for choosing the voting classifier is that the chosen classifiers work well together. Support Vector Machines (SVMs) are highly effective for classification tasks involving high-dimensional data and discover an ideal separation hyperplane, making it resilient for feature-rich representations such as those retrieved by CNNs. K-Nearest Neighbors (KNN) is a lazy learning algorithm, which is simple yet powerful in capturing local similarity patterns, which is particularly useful in discriminating minute variations in handwritten signatures. Random Forest (RF) is an ensemble of multiple decision trees, which contributes robustness through a method called bagging, to reduce variance and prevent overfitting. These three distinct classifiers are integrated via a majority voting mechanism. The voting classifier exploits the strength of each of these three classifiers, resulting in enhanced robustness, reliability, and classification accuracy in contrast to dependence on a single classifier.

3.5. Cross-Validation

Training and testing the model on a single split does not generalize well to unseen data and may lead to biased results due to overfitting issues. Hence, the cross-validation technique is used to assess how well a voting classifier generalises on unseen data.

To ensure a balanced evaluation, stratified cross-validation with five folds is utilized to assess model performance. The entire dataset is split into five equally sized folds. While splitting the dataset, it is ensured that the same proportion of genuine and forged signature samples is maintained. During each run, the four folds are deployed for training, and the remaining one fold is deployed for testing. The experiment is repeated across five runs, ensuring that every fold is evaluated as a test set. The performance metrics of the model are computed by taking the average of the metrics computed across all five runs to assess the generalizing ability of the model.

4. Experimental Results and Discussion

4.1. Description of the Signature Dataset

Evaluation of the proposed framework is conducted using the CEDAR, MCYT-75, BHSig260-Bengali, BHSig260-Hindi, and UTSig benchmark offline signature datasets, which include genuine samples and skilled forgeries.

4.1.1. CEDAR

This dataset was created for research purposes at the CEDAR research centre. The signers were asked to sign within a grid measuring 2 by 2 inches. All signature samples were scanned at three hundred dots per inch and stored in PNG format. Each signer has contributed 24 authentic and 24 forged signatures, resulting in 2640 signatures from 55 signers.

4.1.2. MCYT-75

The dataset comprises scanned images of signatures from 75 different individuals. Each individual typically provides 15 genuine signature samples and a set of 15 skilled forged signatures. The signatures are digitized at 600 dpi and saved in Bitmap (BMP) format.

4.1.3. BHSig260-Bengali

This dataset comprises signatures from 100 different individuals written in the Bengali language. Each individual typically provides 24 genuine samples as well as a set of 30 skilled forged signatures. The signature samples are digitized at three hundred dots per three-hundred-dots-per-inch resolution and stored in Tagged Image File Format (TIF).

4.1.4. BHSig260-Hindi

This dataset comprises signatures from 160 different individuals written in the Hindi language. Each individual typically provides 24 genuine samples as well as a set of 30 skilled forged signatures. The signatures are digitized at 600 dots per inch resolution and stored in a Bitmap (BMP) file format.

4.1.5. UTSig

The UTSig dataset, a publicly accessible Persian offline signature database, contains 8,280 images across 115 classes. For each class, the dataset provides twenty seven authentic signatures, three opposite-hand signatures, 6 skilled fakes, and 36 simple fakes.

4.2. Experimental Setup and Results

This paper focuses on developing a writer-independent SSV system that combines DL and conventional ML techniques. To evaluate the system, experiments are conducted on five benchmark datasets: CEDAR, MCYT-75, BHSig260-Bengali, BHSig260–Hindi, and UTSig. Instead of relying on a single data split, stratified five-fold cross-validation is utilized.

In this section, the experimental setup and the results obtained are described. Model evaluation is carried out using widely adopted performance measures, including FAR, FRR, AER, VAC, Precision, Recall, and F1-score. While extracting the features, a batch size of sixteen, Adam optimizer, and different epochs for different datasets are used. While selecting the features using the CFS, the threshold value was set to 0.2. The optimal results are summarized in Table 1.

Table 1. Performance results of the proposed model

Dataset	CEDAR	BHSig-B	BHSig-H	MCYT-75	UTSig
FVD*	34	14	39	20	43
VCA**	100	99.98	99.75	99.06	99.39
FAR	0.00	0.04	0.49	1.24	0.19
FRR	0.00	0.00	0.06	0.62	2.46
AER	0.00	0.02	0.27	0.93	1.32
Precision	100	99.96	99.61	0.99	99.11
Recall	100	100	99.94	0.99	97.53
F1 score	100	99.98	99.71	0.99	98.32
Time in seconds	23.11	104.49	267.63	38.96	53.69
Epochs	20	20	40	50	40

**FVD- Feature Vector Dimension, **VCA – voting classifier accuracy*

Experimental results showed that there is no fixed epochs for all datasets; it is determined by extensive experiments, and in Table 1, it can be seen that the model reaches peak performance at different training epochs for each dataset. The BHSig-B datasets received the best results for the least FVD, i.e, 14, whereas the highest FVD is selected for the UTSig dataset. Even the feature vector dimension varies across datasets. The model has achieved 100% accuracy for the CEDAR dataset, and the highest error rate for the MCYT-75 dataset. The possible reason for achieving the highest results on the CADAR dataset could be less intra-person variation and high inter-person variation. Table 2 shows the Feature Vector Dimension (FVD) and feature indices selected across various datasets.

The effectiveness of the verification system depends on the selection of features employed during training. Hence, a thorough study was conducted to determine the optimal feature set for which the model's performance is best. It is to be noted from Table 2 that the selected features are different for different datasets.

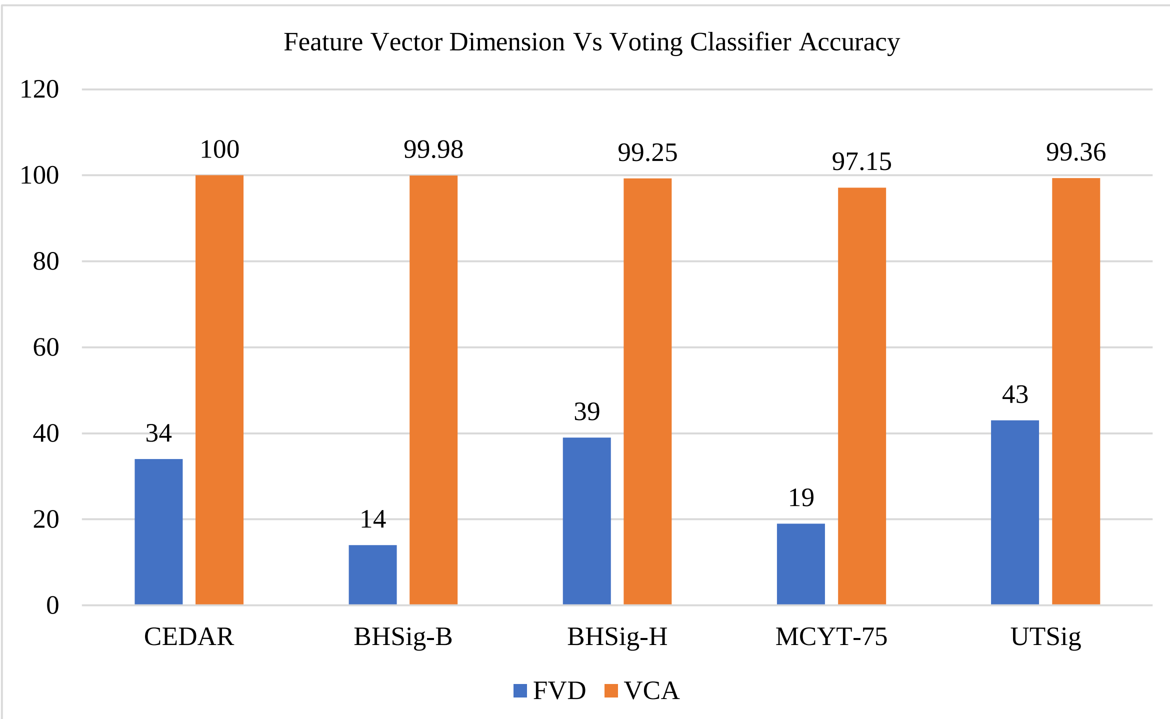

Fig. 2 Feature vector dimension Vs Voting classifier accuracy

Table 2. FVD and selected features of each dataset

Dataset	FVD	Selected Features
CEDAR	34	['0', '3', '4', '5', '7', '8', '16', '17', '29', '37', '49', '52', '58', '63', '69', '75', '76', '78', '79', '83', '84', '85', '86', '89', '92', '93', '95', '97', '99', '100', '108', '110', '117', '125']
BHSig-B	14	['2', '3', '23', '25', '52', '65', '66', '74', '76', '82', '99', '108', '120', '127']
BHSig-H	39	['3', '7', '16', '17', '19', '23', '26', '32', '38', '39', '40', '53', '55', '57', '61', '62', '63', '65', '68', '73', '74', '75', '77', '78', '79', '83', '86', '88', '89', '91', '92', '95', '102', '108', '109', '110', '118', '120', '125']
MCYT-75	20	['5', '6', '7', '9', '17', '32', '34', '41', '51', '55', '66', '76', '77', '78', '81', '84', '100', '114', '118', '125']
UTSig	43	['1', '12', '14', '15', '18', '20', '22', '24', '25', '26', '35', '38', '40', '41', '42', '47', '48', '50', '54', '56', '57', '58', '60', '62', '63', '70', '71', '74', '77', '94', '101', '104', '105', '107', '108', '110', '112', '113', '120', '121', '122', '124', '126']

This may be due to a difference in image resolution, signature style, pen thickness, device used for acquisition, and so on. Figure 2 depicts the FVD and VCA for benchmark datasets.

4.3. Comparison of CNN-Based Hybrid Approach

The proposed model, a CNN-based hybrid approach, has achieved superior results because non-contributing features are identified and removed using the correlation coefficient feature selection method. An experiment is conducted where all 128 features extracted by CNN are fed to a voting classifier, and the results obtained are compared with the proposed model, where all 128 features extracted by CNN undergo a feature selection process before being fed to the voting classifier. The results of models before feature selection and the proposed model are noted in Table 3.

Table 3. Results of the CNN-based hybrid approach (proposed method) compared with the results of the model before feature selection

Dataset		CEDAR	BHSig-B	BHSig-H	MCYT-75	UTSig
Before feature selection	FVD	128	128	128	128	128
	Time (sec)	26.38	188.59	319.71	19.97	44.89
	VCA	95.30	99.44	97.80	81.68	90.61
CNN-based hybrid approach	FVD	34	14	39	19	43
	Time (sec)	23.11	104.49	286.73	15.68	41.94
	VCA	100	99.98	99.25	97.15	99.36
Percentage decrease in FVD		73.43	89.06	69.53	85.15	66.40
Decline in training time units		3.27	84.1	32.98	4.29	2.95
Improvement in VCA		4.7	0.54	1.45	15.47	8.75

The effectiveness of the proposed model is assessed by comparing its performance before and after feature selection. From the results obtained, it is observed that CNN extracted features contain some non-contributing features, which have deteriorated the performance of the model. The correlation coefficient FSM has successfully eliminated non-contributing features; consequently, the proposed approach yields remarkably improved results. For instance, the model has achieved only 95.30% accuracy on the CEDAR dataset when all 128 CNN extracted features were used for classification; in contrast, after applying feature selection, with only 34 contributing features, it was able to achieve 100% accuracy. One can also observe a percentage decrease in FVD, a decline in training time, and an improvement in VCA in Table 3.

4.4. Interpreting Model Predictions with the Local Interpretable Model-Agnostic Explanations (LIME) Technique

Figure 3 shows the output of an explainable AI technique, LIME—the output of LIME shows which features have influenced the decision the most. For demonstration purposes, an authentic and a fake signature have been randomly chosen from the BHSig-B dataset. The 14 contributing features were selected by CFS for the BHSig-B dataset (refer to Table 2). In Figure 3, each bar corresponds to a feature index used by the classifier for verification. The orange bars push the prediction toward Genuine. The blue bars push the prediction toward Forged. The length of each bar indicates the strength of that feature's contribution. In Figure 3, the value just above the bar represents the feature indices, and the value next to the feature indices represents the standardised feature value. The standardised feature value with a positive sign indicates above the mean, and the negative sign indicates below the mean. In Figure 3 (a), the predicted class is genuine, the prediction probability for a genuine signature is 1.00, and the forged signature is 0.00. LIME explains the prediction by listing which features (from the CNN-extracted feature vector) contributed to the decision and in which direction, i.e., "Genuine Class" or "Forged class". The features supporting genuine prediction are indicated by the orange color and fake signatures by the blue color. The features which pushed the predictions toward "Genuine" are 23, 120, 65, 3, 66, 2, 127, 82, 74, 108, 25, 76, 99 and only feature 52 tried to push the prediction towards forge but its contribution was not strong enough to counteract the large set of genuinely supporting features (i.e., 23, 120, 65, 3, 66, 2, 127, 82, 74, 108, 25, 76, 99). These features were highly influential and influenced the model toward classifying the signature as genuine.

Hence, the signature is predicted as genuine. Similarly, Figure 3 (b) is another instance from the same dataset that shows the influence of features 3, 23, 66, etc, towards class "genuine" prediction. Figure 3 (c) and (d) show the influencing features for forged class prediction for a randomly chosen signature sample from the BHSig-B dataset.

4.5. Visual Interpretation using Gradient-Weighted Class Activation Mapping (Grad-CAM)

The output of Grad-CAM is depicted in Figure 4, which visualises the vital parts that support the prediction. The CNN considered yellow/red regions as most influential for the decision. The Blue/Purple areas are the regions with low contribution to the decision. In Figure 4 (a) and (c) are the genuine signature images showing that the model has focused on the central stroke region of the signature (yellow zone). This indicates the core structural parts of the signature, like consistent stroke thickness, smooth curves, and alignment, are identified as discriminative by the CNN. In Figures 4 (b) and (d), the forged signature images show that the attention is more spread out and uneven compared to the genuine one.

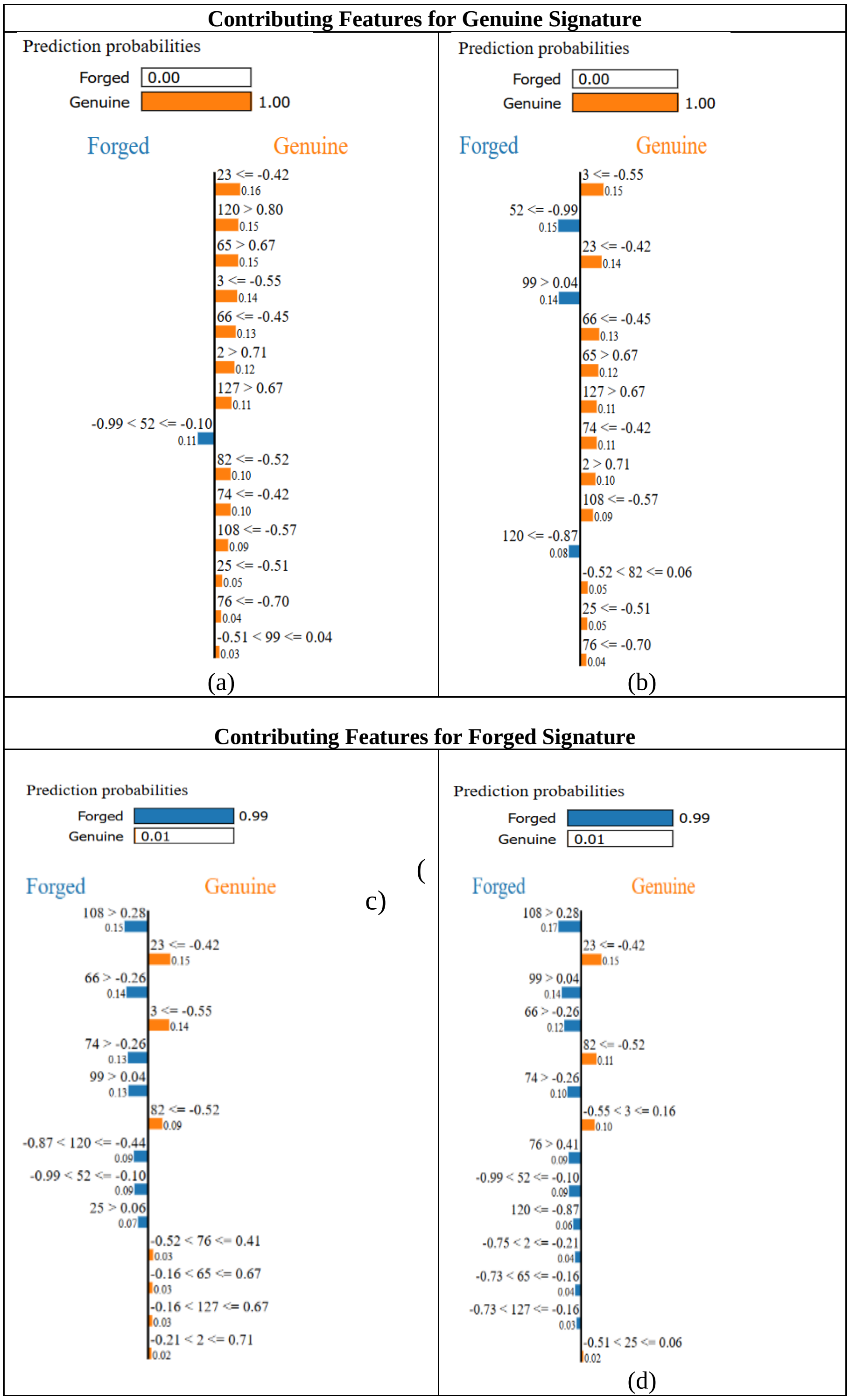

Fig. 3 Illustration of LIME for both genuine and forged signatures

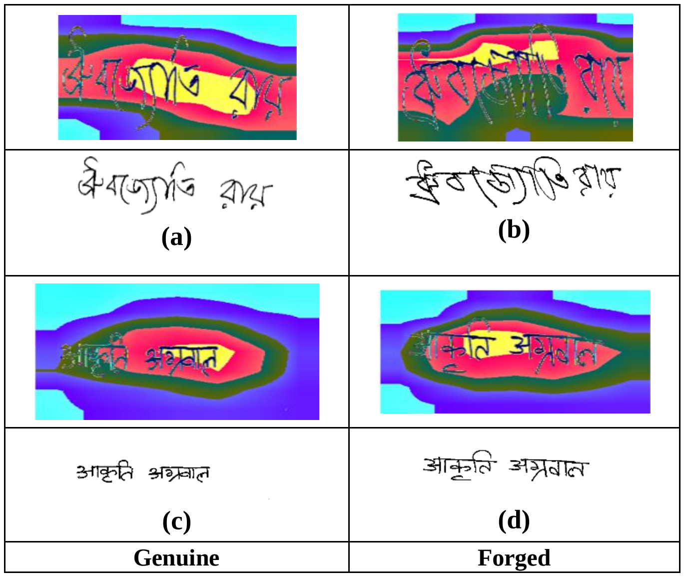

Fig. 4 Important regions visualisation through grad-CAM

The CNN is highlighting multiple disconnected stroke regions (scattered yellow zones), which may indicate irregularities. This often happens because forged signatures have inconsistent pressure, shaky lines, or distortions, so the model tries to "look harder" across different parts.

4.6. Comparison with State-of-the-Art Works

The empirical findings clearly reveal that the CNN is capable of extracting highly discriminative features that effectively separate genuine signatures from forged ones.

Table 4. Performance comparison of the proposed model with existing recently reported methods

Reference	Methods Used	Results
(Alsuhimat & Mohamad, 2023)	HOG + CNN	UTSig : 95.4 CEDAR: 94.1
(Longjam et al., 2023)	CNN capsule Network	CEDAR: 100 BHSig-B:94.3 BHSig-H:100
(Shekar et al., 2022)	CNN+SVM	CEDAR: 93.63
(Muhtar et al., 2023b)	ResNet	CEDAR:96.21 BHSig-B:98.42 BHSig-H:97.28
(Yu-Jie et al., 2024)	Swin Transformer	CEDAR:100 BHSig-B:96.43 BHSig-H:95.26
(Zheng et al., 2025)	Hybrid Transformer	CEDAR:100 BHSig-B:98.63 BHSig-H:97.81 UTSig: 66.19
(Li et al., 2024)	Vision Transformer	BHSig-B:90.10 BHSig-H:96.62
(Zhao et al., 2025)	SigNet + Handcrafted features	CEDAR: 1.04 EER BHSig-B: 2.65 EER BHSig-H: 2.99 EER UTSig : 7.64 EER
Proposed Model		CEDAR:100 BHSig-B:99.98 BHSig-H:99.25 MCYT-75:97.15 UTSig:99.36

To further validate the strength of the presented method, its performance is compared with existing recently reported methods, as summarized in Table 4. The comparison indicates that the proposed model consistently surpasses these advanced methods.

The presented work has outperformed existing recent works. The CNNs have been widely used to solve many pattern recognition and computer vision problems. In the presented work, a shallow CNN has been used to extract discriminative characteristics from the signature pictures with a varying number of epochs for each dataset, as documented in Table 1. These feature vectors, containing 128 features, are subjected to further refinement using the correlation coefficient FSM to remove all non-contributing characteristics. The outcome of FSM contains pure contributing features that are passed to the Voting classifier. Here, the voting classifier plays a vital role in integrating the strengths of the three classifiers. Figure 5 illustrates the accuracy of the voting classifier against individual classifiers. This combination of deep features, FSM, and Voting classifier has outperformed all the existing works.

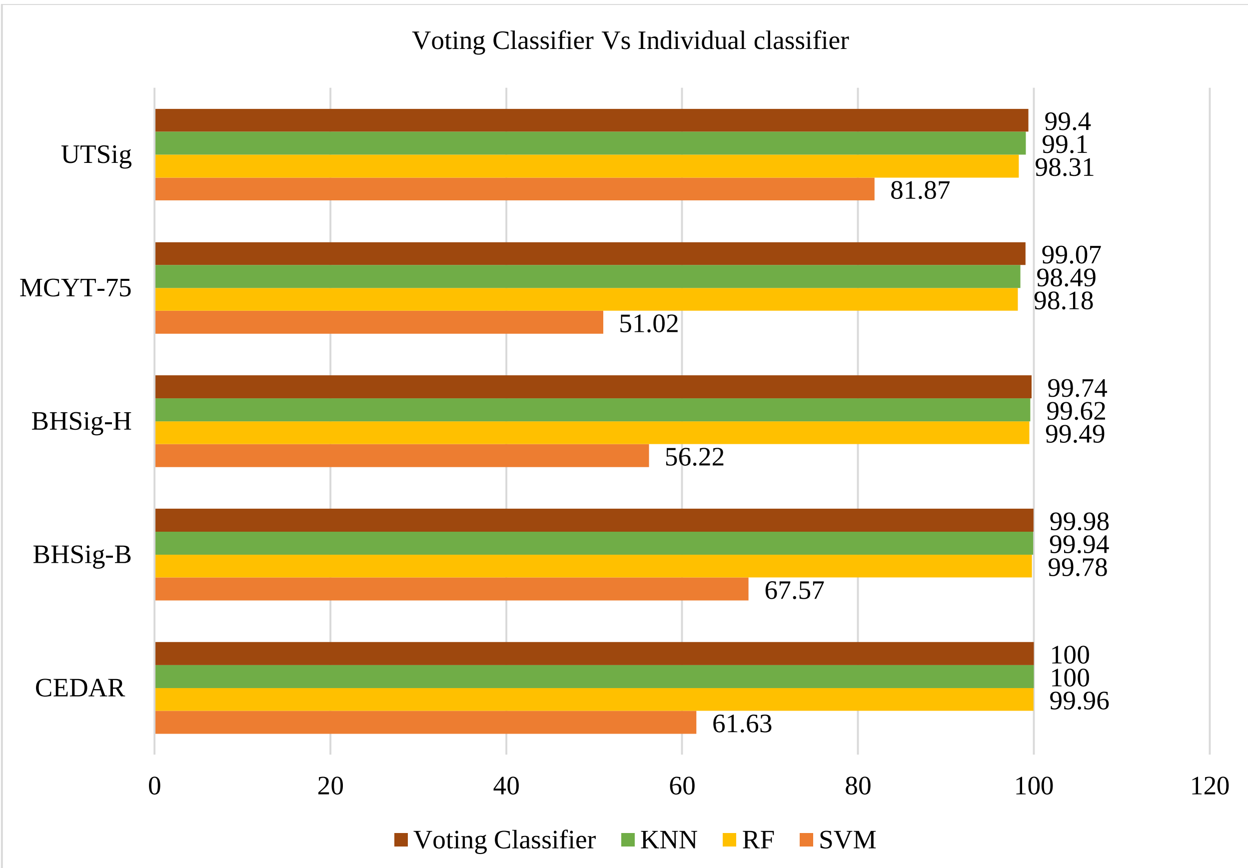

Fig. 5 Illustrates the results of the individual classifiers and the voting classifier

In all datasets, RF and KNN have performed well, but SVM has consistently shown poor performance. The voting classifier has remarkably outperformed by bringing out the strengths of RF and KNN, and at the same time, neglected the weakness of SVM.

5. Conclusion

This paper proposes an SSV model that integrates DL and ML techniques. A lightweight CNN is used to extract superior features, and then non-performing characteristics are eliminated using a correlation coefficient-based feature selection method. The selected discriminative features that contribute to the verification process are used to train the voting classifier. The model has achieved 100% on CEDAR, 99.98% on BHSig-260 Bengali, 99.74% on BHSig-260 Hindi, 99.07% on MCYT-75, and 99.40% on UTSig datasets.

Moreover, an explainable AI, LIME, is used to show which features contributed more to the decision of the model. Additionally, Grad-CAM is employed to illustrate the parts of the signature picture that the CNN identifies as important during classification. This study also compares the efficacy of the presented model with that of the model prior to feature selection, highlighting the impact of removing non-contributory features on overall system performance. The findings indicate that the CNN-generated feature set includes certain non-contributing features, and removing these features leads to a noticeable improvement in the outcome of the model. Prospective studies will concentrate on further exploring advanced characteristic extraction and selection techniques to determine stable and discriminative characteristics essential for robust Static Signature Verification (SSV).

Conflicts of Interest

All authors declare that they have no competing interests related to this work. The authors declare that this manuscript is original, has not been published before, and is not currently being considered for publication elsewhere.

References

[1] Ansam A. Abdulhussien et al., "A Genetic Algorithm Based One Class Support Vector Machine Model for Arabic Skilled Forgery Signature Verification," *Journal of Imaging*, vol. 9, no. 4, pp. 1-26, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Fadi Mohammad Alsuhimat, and Fatma Susilawati Mohamad, "A Hybrid Method of Feature Extraction for Signatures Verification Using CNN and HOG a Multi-Classification Approach," *IEEE Access*, vol. 11, pp. 21873-21882, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Samy Bakheet, and Ayoub Al-Hamadi, "Automatic Detection of COVID-19 using Pruned GLCM-Based Texture Features and LDCRF Classification," *Computers in Biology and Medicine*, vol. 137, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Debanshu Banerjee et al., "A New Wrapper Feature Selection Method for Language-Invariant Offline Signature Verification," *Expert Systems with Applications*, vol. 186, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Faiza Eba Batool et al., "Offline Signature Verification System: A Novel Technique of Fusion of GLCM and Geometric Features using SVM," *Multimedia Tools and Applications*, vol. 83, pp. 14959-14978, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Aman Singla, and Ajay Mittal, "Exploring Offline Signature Verification Techniques: A Survey Based on Methods and Future Directions," *Multimedia Tools and Applications*, vol. 84, pp. 2835-2875, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[7] Subhash Chandra, and Sushila Maheshkar, "Verification of Static Signature Pattern based on Random Subspace, REP Tree and Bagging," *Multimedia Tools and Applications*, vol. 76, pp. 19139-19171, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Moises Diaz et al., "A Perspective Analysis of Handwritten Signature Technology," *ACM Computing Surveys*, vol. 51, no. 6, pp. 1-39, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Luiz G. Hafemann, Robert Sabourin, and Luiz S. Oliveira, "Offline Handwritten Signature Verification—Literature Review," *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Montreal, QC, Canada, pp. 1-8, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[10] M. Muzaffar Hameed et al., "Machine Learning-Based Offline Signature Verification Systems: A Systematic Review," *Signal Processing: Image Communication*, vol. 93, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Maryam Houtinezhad, and Hamid Reza Ghaffary, "Writer-Independent Signature Verification based on Feature Extraction Fusion," *Multimedia Tools and Applications*, vol. 79, pp. 6759-6779, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Fu-Hsien Huang, and Hsin-Min Lu, "Multiscale Feature Learning Using Co-Tuplet Loss for Offline Handwritten Signature Verification," *arXiv preprint*, pp. 1-16, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[13] Anamika Jain, Satish Kumar Singh, and Krishna Pratap Singh, "Signature Verification using Geometrical Features and Artificial Neural Network Classifier," *Neural Computing and Applications*, vol. 33, pp. 6999-7010, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Huan Li et al., "TransOSV: Offline Signature Verification with Transformers," *Pattern Recognition*, vol. 145, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Teressa Longjam, Dakshina Ranjan Kisku, and Phalguni Gupta, "Multi-Scripted Writer Independent Off-line Signature Verification using Convolutional Neural Network," *Multimedia Tools and Applications*, vol. 82, pp. 5839-5856, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Yusnur Muhtar et al., "FC-ResNet: A Multilingual Handwritten Signature Verification Model Using an Improved ResNet with CBAM," *Applied Sciences*, vol. 13, no. 14, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[17] Ebrahim Parcham, Mahdi Ilbeygi, and Mohammad Amini, "CBCapsNet: A Novel Writer-Independent Offline Signature Verification Model using a CNN-based Architecture and Capsule Neural Networks," *Expert Systems with Applications*, vol. 185, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Sourodeep Roy et al., "Offline Signature Verification System: A Graph Neural Network based Approach," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 8219-8229, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Muhammad Sharif et al., "A Framework for Offline Signature Verification System: Best Features Selection Approach," *Pattern Recognition Letters*, vol. 139, pp. 50-59, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] B.H. Shekar, Wincy Abraham, and Bharathi Pilar, "Offline Signature Verification using CNN and SVM Classifier," *2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, Mangalore, India, pp. 304-307, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] B. Venkata Sivaiah et al., "Signatures Verification using CNN and HOG including Voting Classifier," *Proceedings of the International Conference on Computational Innovations and Emerging Trends*, vol. 112, pp. 598-608, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yizhen Wang, Jianbin Zheng, and Yiwen Zhou, "An Efficient Offline Signature Verification Method Based on Improved Feature Extraction," *2022 2nd International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI)*, Nanjing, China, pp. 609-612, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Zhaoya Wang et al., "Advances in Offline Handwritten Signature Recognition Research: A Review," *IEEE Access*, vol. 11, pp. 120222-120236, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Xiong Yu-Jie et al., "Past: Pairwise Attention Swin Transformer for Offline Signature Verification," *SSRN*, pp. 1-14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Lidong Zheng et al., "HTCSigNet: A Hybrid Transformer and Convolution Signature Network for offline Signature Verification," *Pattern Recognition*, vol. 159, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[26] Yiwen Zhou et al., "Handwritten Signature Verification Method Based on Improved Combined Features," *Applied Sciences*, vol. 11, no. 13, pp. 1-14, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] R.S.A. Zneit, "On-Line Handwriting Signature Verification Based on Using Extreme Points Extraction," *Engineering, Technology & Applied Science Research*, vol. 6, no. 4, pp. 1084-1088, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[28] Xingbiao Zhao et al., "Fusing Deep and Hand-Crafted Features by Deep Canonically Correlated Contractive Autoencoder for Offline Signature Verification," *Pattern Recognition*, vol. 168, 2025. [CrossRef] [Google Scholar] [Publisher Link]