

ForgeXplain: An Explainable Machine Learning System for Offline Signature Forgery Detection

Madiha Kaunain¹, M. Nalini², Noor Ayesha², R. Saaranya²

¹Department of Computer Science and Engineering, Rajiv Gandhi Institute of Technology, Bengaluru, India

²Rajiv Gandhi Institute of Technology, Bengaluru, India

Published October 2026

Abstract

Achieving accurate signature verification is critical in many contexts including financial transactions and legal contracts, for reasons of security and fraud risk. The authors describe an offline signature verification system called ForgeXplain that uses hand engineered image features in conjunction with machine learning algorithms and explainable AI. From pre-processed signature images the system extracts 23 hand designed geometric and intensity features that in turn feed into support vector machine (SVM) and random forest (RF) models in order to predict fake from real signatures, with SHAP values explaining the decision on a feature-by-feature basis. A test of 1,320 real and 1,320 fake signatures from CEDAR offline found that Random Forest was correct 91.5 percent of the time and achieved a ROC-AUC score of 0.966, better than SVM. This shows that ForgeXplain can tackle off-line signature forgery detection effectively and faithfully.

Keywords: signature verification; forgery detection; explainable AI; SHAP; random forest; support vector machine; CEDAR dataset

1 Introduction

Handwritten signatures are commonly used to verify identity in banking, legal agreements, academic records, and other sensitive transactions. Since signatures can be forged, fraudulent use may result in financial loss, impersonation, or legal disputes. Traditionally, verification is performed by comparing a questioned signature with known genuine samples. However, this manual process is time-consuming and difficult to automate because handwriting varies between individuals and examiner judgements may be subjective [1, 2, 13, 14].

Signature verification is divided into online and offline approaches. Online methods use dynamic information recorded during the signing process, whereas offline methods examine a static signature image obtained through scanning or photography. Therefore, offline verification relies heavily on the visual features extracted from the image, making feature selection an important factor in system performance [1, 2, 11, 13, 14].

Several studies have applied handcrafted geometric, structural, texture, and directional features with traditional machine-learning classifiers, including Random Forest and Support Vector Machine, to distinguish genuine signatures from forged ones [3–6, 9, 11, 13]. Deep-learning approaches have also been investigated for automatically learning discriminative representations from signature images [7, 8, 12, 14]. However, many automated systems provide only a classification result without explaining the features responsible for the prediction. This limitation may reduce user confidence and make it more difficult to assess incorrect decisions.

To address this issue, ForgeXplain is proposed as an explainable offline signature verification system. The framework combines image preprocessing, handcrafted feature extraction, machine-learning classification, and explainable artificial intelligence. It extracts 23 features related to geometry, shape, density, contour, and intensity, and evaluates Support Vector Machine (SVM) and Random Forest (RF) classifiers for distinguishing genuine and forged signatures. SHAP-based explanations are used to identify the contribution of individual features to each classification decision.

The proposed system is evaluated using the CEDAR offline signature dataset, which contains 1,320 genuine and 1,320 forged signature samples. The evaluation considers both classification performance and the interpretability of the model predictions.

2 Related work

Research into offline handwritten signature verification has explored the use of image processing, handcrafted features, conventional machine learning, and deep learning. Khalifa and colleagues [1] provide a comprehensive review of major offline signature verification and forgery detection techniques, encompassing preprocessing, feature extraction, and classification. Ghanim and Nabil [2] demonstrated the application of image-processing and machine-learning techniques for distinguishing between genuine and forged signatures, where Thenuwara and Nagahamulla [3] highlighted the effectiveness of Random Forest when paired with appropriate signature features. Larkins and Mayo [4] introduced adaptive feature thresholding to enhance the discrimination between genuine and forged items. Approaches based on texture and shape have also been investigated. Wajid and colleagues [5] investigated Local Binary Pattern features for capturing local texture information, whereas Ooi et al. [6] employed the Radon transformation to isolate directional features. With the advancement of deep learning, Hafemann et al. [7] demonstrated the ability of convolutional neural networks to learn discriminative signature features, and Dey et al. [8] introduced SigNet, a convolutional Siamese network designed for writer-independent verification. Guerbai et al. [9] further investigated the application of one-class SVM for writer-independent verification. To enhance model interpretability, Lundberg and Lee [10] introduced SHAP, a method that assigns contribution values to input features, thereby enabling explanations of machine-learning predictions at feature level. Other studies such as those by Kiran et al. [11], Poddar et al. [12], Hashim et al. [13], Wang et al. [14] further investigated image processing, neural networks, machine-learning classifiers and the shift towards deep-learning-based signature verification.

Research gap: While existing studies have demonstrated the promising verification performance, they primarily focus on classification accuracy and offer limited insight into the reasoning behind individual predictions. While SHAP facilitates interpretable machine-learning decisions [10], its application in offline signature forgery detection is still limited. This drives the development of ForgeXplain, a method that integrates signature forgery detection with feature-level explanations to deliver predictions that are both accurate and interpretable.

3 System architecture and methodology

ForgeXplain is built as a multi-page Streamlit web application that leverages a scikit-learn-based pipeline for both training and inference. The overall workflow is depicted in Figure 1.

The system handles an uploaded signature by sequentially executing preprocessing, feature extraction, classification, and explainability stages. The final output includes the predicted class, confidence score, and feature-level contributions, all displayed via the application interface.

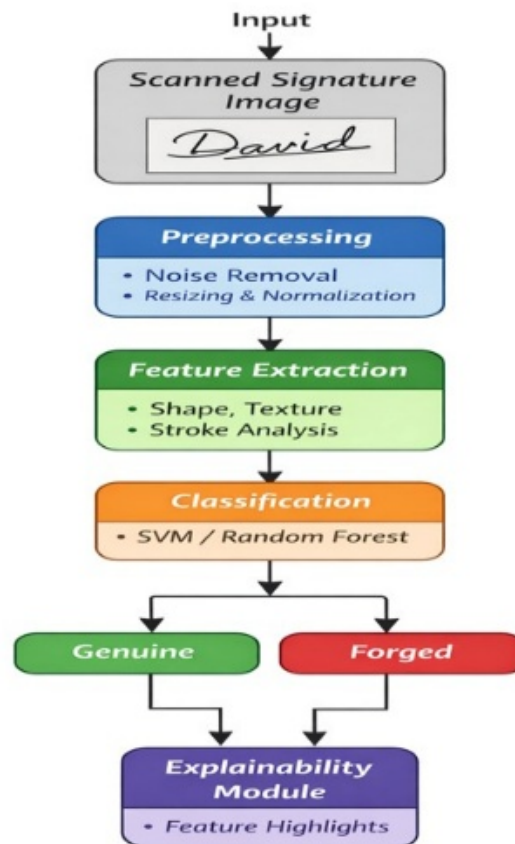


Figure 1. ForgeXplain system architecture for signature forgery detection.

3.1 Preprocessing and segmentation

The signatures are resized, converted to grayscale, and denoised using Gaussian blur. The image is first binarized using adaptive thresholding, and then morphological operations are applied to remove noise while preserving the strokes. The signature is cropped and re-centered to ensure consistency. For documents, the system automatically detects a candidate region, which the user then confirms to filter out irrelevant content.

3.2 Feature extraction

ForgeXplain extracts a fixed 23-dimensional feature vector from each preprocessed signature. These features capture complementary geometric, shape, density, contour, and intensity characteristics while maintaining interpretability for subsequent explanations. The features are grouped into five categories:

- **Geometric features (5):** contour area, perimeter, aspect ratio, bounding-box width, and height capture the signature's size and geometric proportions.
- **Shape features (7):** seven log-scaled Hu moment invariants capture the overall signature shape while providing invariance to translation, scaling, and rotation.

- **Density features (2):** the foreground pixel density and the black-to-white pixel ratio serve as indicators of ink coverage within the signature area.
- **Contour features (3):** the count of unique contours, mean contour area, and solidity reflect continuity, fragmentation, and structural properties.
- **Intensity histogram features (6):** a six-bin normalized grayscale histogram captures the distribution of pixel intensities.

This feature set strikes a balance between discriminative power and interpretability. In contrast to high-dimensional representations, which can be challenging to connect with visible signature characteristics, these features possess direct geometric or visual interpretations that can be seamlessly integrated into the explainability layer.

3.3 Classification

The extracted feature vectors are standardized using a scaler fitted solely on the training data and subsequently fed into two independently trained classifiers: a Support Vector Machine (SVM) and a Random Forest (RF).

The SVM employs an RBF kernel with a C value of 2.0 and has probability estimation enabled. The Random Forest model comprises 300 decision trees. Since both classifiers employ the identical feature representation and train-test split, their performance can be compared consistently.

3.4 Explainability layer

ForgeXplain incorporates SHAP-based explainability directly into the prediction pipeline. For the Random Forest classifier, TreeSHAP leverages the structure of the decision-tree ensemble to efficiently compute feature-level Shapley values.

For every analyzed signature, the explainability layer pinpoints the features driving the predicted class and quantifies both the direction and magnitude of their impact on distinguishing between genuine and forged signatures. These explanations are displayed within the application interface alongside the classification results, thereby embedding interpretability directly into the prediction workflow.

The result interface also presents the processed signature image, enabling users to review the analyzed image alongside its corresponding feature-level explanation.

4 Experimental results and analysis

4.1 Dataset

ForgeXplain is evaluated on the CEDAR offline signature dataset, which comprises 2,640 signature images in total—1,320 genuine and 1,320 forged signatures collected from 55 writers. The dataset includes 24 genuine and 24 forged samples for each writer. The dataset is employed to assess the performance of the proposed signature verification system as well as the two selected classification models.

4.2 Experimental setup

The dataset is partitioned into training and testing sets via an 80:20 stratified random split performed at the image level. This yields 2,112 training images and 528 test images, each subset

preserving an equal balance of genuine and forged signatures. Both the SVM and Random Forest classifiers were trained using the identical 23-dimensional feature representation detailed in Section III-B, along with the configurations specified in Section III-C, and all performance metrics were evaluated on the held-out test set comprising 528 images.

4.3 Classification results

Table 1 summarizes the classification performance of the SVM and Random Forest models, reporting metrics for accuracy, precision, recall, F1-score, and ROC-AUC.

Table 1. Classification performance on the CEDAR test set.

Classifier	Accuracy	Precision	Recall	F1-score	ROC-AUC
SVM (RBF)	86.4%	86.1%	86.7%	86.4%	0.939
Random Forest (300 trees)	91.5%	91.3%	91.7%	91.5%	0.966

The SVM misclassified 37 forged signatures as genuine and 35 genuine signatures as forged. These results suggest that both classifiers achieved relatively balanced performance, although the Random Forest model provided better overall accuracy for the selected feature representation.

4.4 Feature importance and explainability

The Random Forest analysis showed that intensity-histogram features and the first Hu moment had the greatest influence on classification. As shown in Table 2, hist_bin_2 was the most important feature, contributing 23.0% to the model's decisions. The first Hu moment was the most influential non-histogram feature, with an importance of 6.3%.

Table 2. Top six Random Forest feature importances.

Rank	Feature	Relative importance
1	hist_bin_2 (ink-density histogram)	23.0%
2	hist_bin_3 (ink-density histogram)	9.2%
3	hist_bin_5 (ink-density histogram)	6.8%
4	hist_bin_4 (ink-density histogram)	6.7%
5	hu_moment_1 (overall shape)	6.3%
6	hist_bin_6 (ink-density histogram)	5.4%

These results demonstrate the value of ForgeXplain's explainability module. In addition to predicting whether a signature is genuine or forged, the system identifies the features that influence each decision, thereby supporting more transparent verification and user review.

These findings demonstrate the usefulness of the explainability component in identifying the characteristics that influence the classifier. Rather than providing only the final classification, ForgeXplain allows the contribution of individual features to be examined for a given prediction. This provides additional information for users reviewing the output of the signature verification system. An example of the local SHAP explanation is shown in Figure 2.

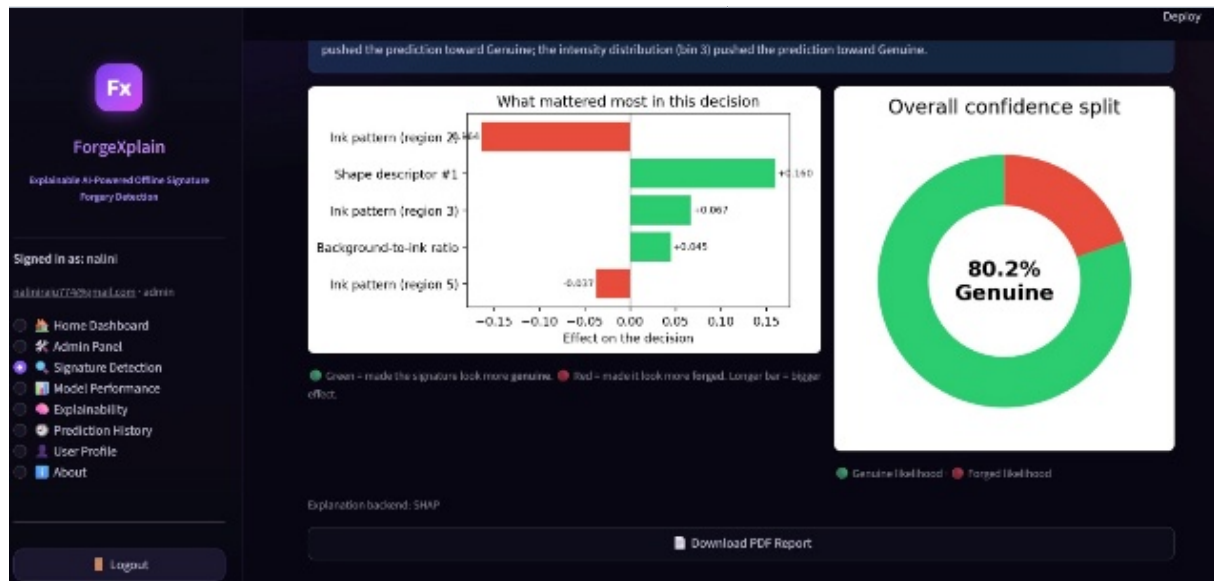


Figure 2. Local SHAP explanation interface illustrating individual feature impacts (ink patterns and region-based shape descriptors) on model decision outputs.

4.5 Comparison with prior work

Table 3 compares ForgeXplain's accuracy with that of selected offline signature-verification systems. This comparison is intended only to provide context, since the studies used different datasets, feature representations, training methods, and evaluation protocols. Consequently, the reported accuracy values should not be considered directly equivalent.

Table 3. Comparison with reported offline signature verification results.

System	Dataset	Features	Best accuracy
Ghanim and Nabil [2]	UTSig	HOG + SVM	94.0%
Larkins and Mayo [4]	In-house dataset (2,106 signatures)	Modified direction feature + RBF-NN	91.2%
Thenuwara and Nagahamulla [3]	ICDAR	Random Forest	67.0%
ForgeXplain (this work)	CEDAR	Geometric/shape/density/histogram + Random Forest	91.5%

Ghanim and Nabil [2] evaluated an offline signature verification system using the UTSig dataset and various feature–classifier combinations. Their highest accuracy was 94%, achieved using an SVM with Histogram of Gradients features.

Larkins and Mayo [4] introduced Adaptive Feature Thresholding (AFT) and tested it on the CEDAR and GPDS datasets. Their CEDAR results were 92% for the manual version and 90% for the automatic version; therefore, the 92% result is included in Table 3 for comparison.

Thenuwara and Nagahamulla [3] also assessed several classifiers, with their Random Forest model achieving 60% accuracy.

ForgeXplain's Random Forest model achieved 91.5% accuracy on the CEDAR test set, which is comparable to earlier results. However, these values should be interpreted cautiously because the studies

use different datasets, features, training methods, and evaluation procedures. The CEDAR-based result reported by Larkins and Mayo offers the most relevant comparison. ForgeXplain focuses not only on classification accuracy but also on interpretable features and SHAP-based explanations, which link feature characteristics to individual predictions. Thus, its main contribution is the integration of signature verification with interpretability rather than accuracy optimization alone.

5 Practical deployment considerations

ForgeXplain is designed for practical use as a client-facing web application, with emphasis on response time, reliable preprocessing, and secure system management.

5.1 Inference latency

The trained models and SHAP Tree Explainer are loaded once and retained in memory for subsequent requests. This avoids repeated disk access and reduces processing overhead. Under the tested environment, post-initialization model retrieval and inference were completed within the sub-millisecond range, although the first request incurred additional cold-start time.

5.2 Human-assisted signature cropping

For documents containing signatures within larger regions, the system detects a candidate signature area for user confirmation or adjustment. This step helps exclude unrelated elements such as text, stamps, and borders, improving the quality of the image supplied to the feature-extraction stage (Figure 3).

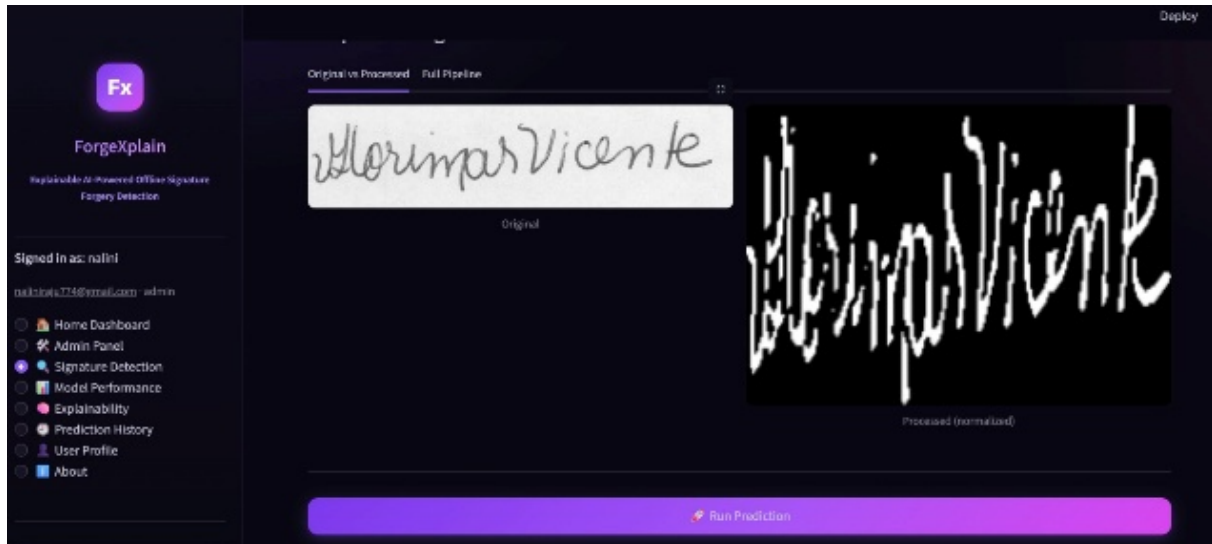


Figure 3. Interactive Streamlit application illustrating user-assisted signature upload, candidate region cropping, and prediction confidence scoring.

5.3 Access control

ForgeXplain uses role-based access control to separate ordinary verification tasks from administrative functions. General users can verify signatures, while administrators can access restricted features such as retraining, diagnostics, and model management. This separation reduces the risk of unauthorized changes to the models and system configuration. A generated analysis report is shown in Figure 4.

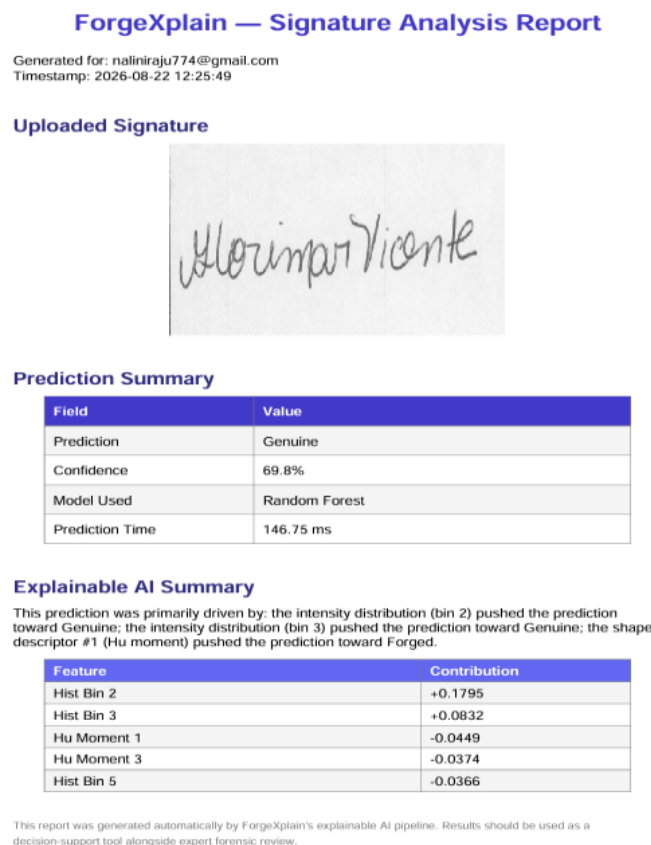


Figure 4. Generated ForgeXplain signature analysis report detailing verification outcome, model metrics, and feature-level explainable AI (XAI) contributions.

6 Discussion

6.1 Interpretation of results

Random Forest performed better than SVM with the proposed 23-feature representation. It achieved 91.5% accuracy, 91.3% precision, 91.7% recall, 91.5% F1-score, and a ROC-AUC of 0.966, compared with 86.4% accuracy and a ROC-AUC of 0.939 for SVM. The balanced precision and recall indicate reliable identification of both genuine and forged signatures, confirming the usefulness of the extracted geometric, shape, density, contour, and intensity features on the CEDAR dataset.

6.2 Accuracy-interpretability trade-off

ForgeXplain combines competitive classification performance with interpretable features. Instead of relying on complex, high-dimensional representations, it uses 23 meaningful features that can be directly examined by the explanation module. The Random Forest classifier achieved 91.5% accuracy while retaining greater interpretability than typical black-box approaches.

6.3 SHAP-based explainability

SHAP was used to explain individual predictions by measuring the positive or negative contribution of each feature. The results identified hist_bin_2, other intensity-histogram features, and the first Hu moment as major contributors to the model's decisions. These explanations support the review of

automated verification results. However, they remain dependent on the trained model and evaluation conditions.

6.4 Practical implications and limitations

The Streamlit implementation integrates signature classification, user-assisted cropping, SHAP explanations, and role-based access control into a single workflow. Nevertheless, the evaluation used only the CEDAR dataset and an image-level 80:20 train-test split. Therefore, the results may not fully represent performance under different writers, documents, or acquisition conditions. Validation using multiple datasets and writer- and document-independent protocols is required.

7 Conclusion and future work

ForgeXplain is an explainable framework for offline signature verification that combines feature extraction, machine learning, SHAP-based explanations, and web-based deployment. Using 23 extracted features, the study evaluated SVM and Random Forest classifiers on the CEDAR dataset. Random Forest achieved the best results, with 91.5% accuracy, 91.3% precision, 91.7% recall, 91.5% F1-score, and a ROC-AUC of 0.966. The Streamlit application further integrates signature cropping, instance-level explanations, verification, and role-based access control into a unified system.

Future work will evaluate the framework on diverse datasets covering different writing styles, image qualities, and document conditions. Further improvements may include feature-selection methods, alternative classifiers, model-agnostic explanation techniques, and enhanced cropping for pages containing multiple signatures, seals, and text regions.

References

- [1] Khalifa OO, Hashim AHA, Samad SA. Offline signature verification and forgery detection: A review. *International Journal of Computer Science and Network Security*. 2011;11(3):170–176.
- [2] Ghanim TM, Nabil AM. Offline signature verification and forgery detection approach. In: *2018 13th International Conference on Computer Engineering and Systems (ICCES)*; 2018. p. 293–298. <https://doi.org/10.1109/ICCES.2018.8639420>
- [3] Thenuwara M, Nagahamulla HRK. Offline handwritten signature verification system using Random Forest classifier. In: *17th International Conference on Advances in ICT for Emerging Regions (ICTER)*; 2017. p. 191–196. <https://doi.org/10.1109/ICTER.2017.8257828>
- [4] Larkins RL, Mayo M. Adaptive feature thresholding for off-line signature verification. In: *23rd International Conference Image and Vision Computing New Zealand (IVCNZ)*; 2008. p. 1–6. <https://doi.org/10.1109/IVCNZ.2008.4762072>
- [5] Wajid R, Malik H, Malik MI. Offline signature verification using local binary pattern features. *International Journal of Pattern Recognition and Artificial Intelligence*. 2014;28(7):1456003. <https://doi.org/10.1142/S0218001414560036>
- [6] Ooi SY, Teoh ABJ, Pang YH. Radon transform based approach for offline signature verification. *Expert Systems with Applications*. 2013;40(17):6838–6846. <https://doi.org/10.1016/j.eswa.2013.06.042>
- [7] Hafemann LG, Sabourin R, Oliveira LS. Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognition*. 2017;70:163–176. <https://doi.org/10.1016/j.patcog.2017.05.012>
- [8] Dey S, Dutta A, Toledo JI, Ghosh SK, Lladós J, Pal U. SigNet: Convolutional Siamese network for writer independent offline signature verification. *Pattern Recognition Letters*. 2019;128:182–189. <https://doi.org/10.1016/j.patrec.2017.07.024>

-
- [9] Guerbai A, Chibani Y, Hadjadji B. The effective use of one-class SVM classifier for handwritten signature verification based on writer-independent parameters. *Pattern Recognition*. 2015;48(1):103–113. <https://doi.org/10.1016/j.patcog.2014.07.015>
 - [10] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*. 2017;30:4765–4774.
 - [11] Kiran P, Parameshachari BD, Yashwanth J, Bharath KN. Offline signature recognition using image processing techniques and back propagation neuron network system. *SN Computer Science*. 2021;2(3):196. <https://doi.org/10.1007/s42979-021-00591-y>
 - [12] Poddar J, Parikh V, Bharti SK. Offline signature recognition and forgery detection using deep learning. *Procedia Computer Science*. 2020;170:610–617. <https://doi.org/10.1016/j.procs.2020.03.133>
 - [13] Hashim Z, Ahmed HM, Alkhayyat AH. A comparative study among handwritten signature verification methods using machine learning techniques. *Scientific Programming*. 2022;2022:8170424. <https://doi.org/10.1155/2022/8170424>
 - [14] Wang Z, Muhammad M, Yadikar N, Aysa A, Ubul K. Advances in offline handwritten signature recognition research: A review. *IEEE Access*. 2023;11:120222–120236. <https://doi.org/10.1109/ACCESS.2023.3326471>