SIGNATURE VERIFICATION USING IMAGE PROCESSING

SOMU SRIKANTH REDDY
Computer Science and Engineering
Dhanalakshmi Srinivasan University
Samyapuram, Trichy
8309669374
srik84397@gmail.com

UPPALAPATI SRIKANTH
Computer Science and Engineering
Dhanalakshmi Srinivasan University
Samyapuram, Trichy
9391778198
Srikanthuppalapati95@gmail.com

SUNKARI RAKESH
Computer Science and Engineering
Dhanalakshmi Srinivasan University
Samyapuram, Trichy
8008980126
sunkariraki@gmail.com

Mr.A.T.BARANI VIJAYA KUMAR
Assistant Professor
Computer Science and Engineering
Dhanalakshmi Srinivasan University
Samyapuram, Trichy
8220351548
baranivijayakumarat.set@dsuniversity.ac.in

Abstract*—Signature verification is crucial for authenticating documents in banking, legal, and educational sectors. The paper outlines an automated system that uses machine learning and image processing to identify genuine signatures and detect forgeries. By analyzing unique signature features, the system ensures accurate verification, enhancing security and efficiency in document authentication processes. The proposed method employs Histogram of Oriented Gradients (HOG) to capture distinctive features reflecting the shape and orientation of elements in signature images. A Random Forest (RF) classifier, utilizing an ensemble approach, is trained on these features to improve accuracy and mitigate overfitting. Comprehensive preprocessing is conducted, including converting images to grayscale and resizing them for consistency. Hyperparameter tuning via GridSearchCV optimizes the model's performance, yielding accuracies of 99.70% on the CEDAR dataset, 99.59% on the BHSig-B dataset, and 98.94% on the BHSig-H dataset. Precision, recall, and F1-scores surpass 94% for CEDAR and BHSigB, and 98% for BHSig-H. The results highlight the model's effectiveness in identifying skilled forgeries, with strong generalization and computational efficiency across these datasets. This approach is chosen for its reduced computational demand compared to prior methods, driven by HOG's efficient feature extraction and RF's optimized ensemble framework, which together minimize processing time while delivering high accuracy.

***Keywords*—Signature verification, machine learning, image processing, Histogram of Oriented Gradients (HOG), Random Forest, GridSearchCV, biometric authentication, forgery detection, digital security**

I. INTRODUCTION

In the modern digital landscape, handwritten signatures continue to serve as a trusted method for confirming identity in legal, financial, and administrative settings. However, despite advancements in biometric technologies, signatures remain at risk of forgery, resulting in fraud, financial damage, and compromised security. Manual verification

methods, which are often slow, subjective, and susceptible to mistakes, underscore the growing need for automated, trustworthy solutions.

Recent progress in machine learning (ML) and image processing has facilitated the creation of advanced signature verification systems. Nonetheless, challenges remain, especially in detecting skilled forgeries, where fraudsters closely replicate authentic signatures, and managing natural variations in signatures caused by factors like writing conditions, emotional state, or aging. Although deep learning approaches, such as Convolutional Neural Networks (CNNs), offer potential, they demand substantial datasets and significant computational power, rendering them less feasible for practical use in settings with limited resources.

To overcome these challenges, this study introduces an effective and efficient offline signature verification system that employs Histogram of Oriented Gradients (HOG) for feature extraction and a Random Forest (RF) classifier to identify forgeries. HOG effectively captures structural and directional patterns in signatures, while RF's ensemble learning approach ensures high accuracy and resistance to overfitting. The model is further optimized through hyperparameter tuning using GridSearchCV, ensuring optimal performance. Additionally, a Flask-based web application is developed to demonstrate real-time usability, allowing users to upload signatures and receive instant verification results.

II. Related Works

Signature verification is a crucial biometric method for ensuring secure identity authentication in financial, legal, and administrative fields. The main difficulty lies in effectively differentiating authentic signatures from skilled forgeries using image processing and machine learning techniques. Offline signature verification, which depends on static images, poses unique challenges due to the lack of dynamic features like stroke pressure and timing, in contrast to online verification. Recent developments in deep learning, such as convolutional neural networks (CNNs), transformers, and generative models, have markedly improved the accuracy, reliability, and adaptability of offline signature verification systems, as demonstrated by various innovative approaches.

Shekar, Abraham, and Pilar [4] (2022) proposed a hybrid CNN-SVM model for writer-dependent offline verification using 9×9 kernels for feature extraction and SVM for classification, achieving 93.63% accuracy on the CEDAR dataset. CNN-only models yielded lower accuracies due to overfitting. Li, Wei, Ma, Li, and Zheng [5] (2024) introduced TransOSV, a transformer-based model using holistic and part-based encoders along with contrastive loss, offering improved performance on datasets like BHSig and GPDS, although accuracy metrics weren't disclosed.
benchmark despite the absence of accuracy metrics.

To improve CNN efficiency, Li, Wen, and He [6] (2023) developed SCConv, a lightweight convolution module with SRU and CRU units to reduce redundancy, showing promise for deployment in tasks such as signature verification. Wei, Li, and Hu [7] (2019) introduced the Inverse Discriminative Network (IDN), employing weight-shared discriminative and inverse streams enhanced with attention mechanisms, showing robust results across several datasets, despite lacking explicit accuracy figures.

Zeng [8] (2022) proposed a Siamese network with multiscale attention, focusing on stroke features and achieving 82% accuracy on the SigComp-2011 dataset. In spoof detection, Amjad, Goeller, Seitz, Knoll, Bajwa, Tetzlaff, and Malik [9] (2024) used CycleGANs with attention modules for high-quality forged signature generation, with 80–100% spoofing success rates across multiple datasets, stressing the need for generator-based forgery research.

Ishfaq, Saadia, Alserhani, and Gul [10] (2024) introduced a hybrid ViT integrating ResNet-18 and MobileNetV2, achieving 99.96% accuracy on CEDAR and demonstrating real-time applicability. Dey, Dutta, Toledo, Ghosh, Llados, and Pal [11] (2017) proposed SigNet, a convolutional Siamese network, showed strong generalization across scripts by learning a Euclidean distance-based embedding, setting a address class imbalance, SMOTE is utilized. The model's effectiveness is assessed using metrics such as accuracy, precision, recall, F1-score, and a confusion matrix.

B. Image Preprocessing

Data preprocessing plays a critical role in preparing signature images for effective machine learning. Initially, the images are transformed into grayscale and resized to ensure uniform dimensions across the dataset. Feature extraction is conducted using the Histogram of Oriented Gradients (HOG), which captures essential edge and gradient-based characteristics unique to each signature. These features are then normalized to maintain consistent scale and variance across samples, which enhances the efficiency and accuracy of the classifier. To address class imbalance, particularly between genuine and forged signatures, the Synthetic Minority Oversampling Technique (SMOTE) is employed. This ensures balanced learning by generating synthetic samples for the minority class, thereby preventing bias and improving overall model performance

HOG Feature Descriptor

The HOG descriptor captures gradients using orientation histograms. For an image (,), the gradient is:

$$_x = (+ 1,) - (- 1,) \quad (1)$$

$$_y = (, + 1) - (, - 1)$$

The gradient magnitude and orientation are composed as:

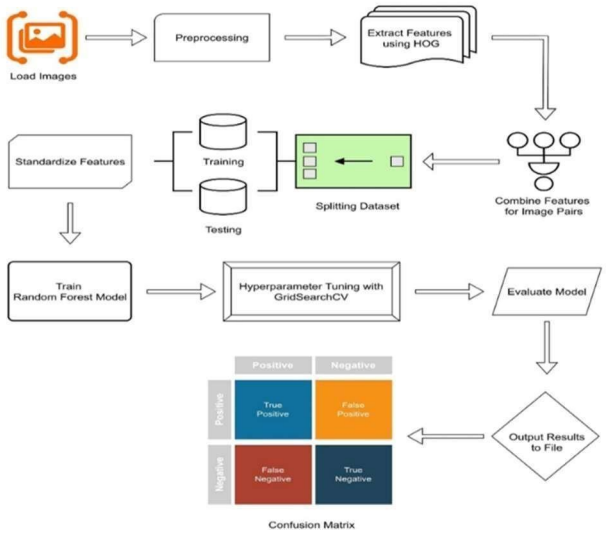
$$= \sqrt{^2 + ^2} \quad (2)$$

x y

Parcham, Ilbeygi, and Amini [12] (2021) combined CNNs with Capsule Networks to retain spatial transformations and reduced model size by half,

outperforming many traditional methods. Similarly, Li, Lin, Wang, Yu, Yuan, and Wang [14]
(2019) introduced a two-channel CNN to address overfitting, **Standardization of Features (StandardScaler)** Each feature is normalized using:
reducing EER from 22.24% to 9.95% on GPDS-Synthetic, proving its effectiveness in writer-independent verification.

While primarily focused on online handwritten Chinese character recognition, Lai, Jin, and Yang [13] (2017) introduced DropDistortion and dynamic feature extraction techniques, achieving over 97% accuracy on CASIA datasets, which may offer valuable insights for online



signature verification. Collectively, these studies demonstrate significant progress in offline signature verification, utilizing sophisticated deep learning frameworks like CNNs, transformers, and generative models to tackle forgery detection challenges and enhance system reliability across varied datasets.

III. METHODOLOGY *A.*

Research Design

This research adopts a supervised machine learning strategy, employing a Random Forest classifier for offline signature verification. The process starts with preprocessing signature images by transforming them into grayscale and standardizing their size. Structural features are extracted using Histogram of Oriented Gradients (HOG). These features are then fed into a Random Forest model, with its hyperparameters optimized through GridSearchCV. To **SMOTE (Synthetic Minority Over-sampling Technique)** Synthetic samples are generated by linear interpolation:

$_{new} = _{i} + (_{nm} + _{i})$, [0, 1] (4) where $_{i}$ is a minority class instance and $_{nm}$ is one of its nearest neighbors.

Fig. 1. System Architecture

$= \frac{x-\mu}{\sigma}$ (3)

Where is the mean and is the standard deviation of the feature.

C. Proposed System Architecture

The system architecture for the signature verification project follows a modular pipeline beginning with the acquisition of signature image datasets, which include labeled genuine and forged samples. The preprocessing phase involves converting the images to grayscale, resizing them to a uniform resolution, and applying normalization to standardize pixel intensity values. This ensures consistency across all samples, which is crucial for effective feature extraction.

Following preprocessing, Histogram of Oriented Gradients (HOG) is applied to extract structural and gradientbased features from the signatures. These features help capture the fine-grained details of handwritten strokes, which are essential in distinguishing between genuine and forged signatures.

After feature extraction, the data is partitioned into training and testing sets. Model training is carried out using a Random Forest classifier, which is optimized through hyperparameter tuning using GridSearchCV. This optimization process systematically identifies the most effective parameter combinations to enhance classification performance and generalization.

Once the model is trained, the system proceeds to the prediction and evaluation phase, where the trained model is applied to the test dataset. The architecture is designed to ensure reproducibility, minimize data leakage, and support seamless integration into future real-time signature verification applications.

For a grayscale image (,), the horizontal and vertical gradients are computed using simple kernels (e.g., Sobel or centered difference):

$$_{x} = (+ 1,) - (- 1,) \quad (5)$$
$$_{y} = (, + 1) - (, - 1)$$

Gradient Magnitude and Orientation From the gradients, compute the magnitude and orientation of the gradient vector at each pixel:

$$= \sqrt{G_x^2 + G_y^2} \quad (6)$$

$$= \left(\frac{Gy}{Gx}\right) \quad (7)$$

Histogram Construction (per Cell)

D. Feature Extraction Using Histogram of Oriented Gradients (HOG)

Histogram of Oriented Gradients (HOG) is employed to extract meaningful features from signature images by

emphasizing the structural and gradient-based characteristics inherent in handwriting. The technique operates by dividing the image into small spatial regions, known as cells, and computing the distribution of intensity gradients or edge directions within each cell. These localized histograms are then concatenated to form a comprehensive feature descriptor that effectively represents the shape and contour of the signature.

HOG is particularly advantageous in signature verification tasks due to its ability to highlight subtle differences in stroke orientation and curvature key factors in differentiating genuine signatures from forgeries. In this study, signature images are standardized in size and converted to grayscale prior to HOG extraction, resulting in high-dimensional feature vectors that encapsulate the signature's discriminative information. These vectors form the foundation for subsequent classification and verification processes.

Gradient Computation

Each cell (e.g., 8×8 pixels) creates a histogram of gradient directions. The histogram bins (typically 9) span orientations (0° to 180° for unsigned gradients). Each pixel votes to a bin based on its orientation, weighted by its magnitude:

$$_k = \sum_{(x,y)\in cell} (,) \cdot ((,)_b) \quad (8)$$ where:

- $_k$ is the th bin
- $(,) = (,)$
- is the indicator function

Block Normalization (e.g., L2-Norm) To reduce illumination and contrast sensitivity, local histograms from adjacent cells are normalized:

$$^{\wedge} = \frac{}{\sqrt{\|v\|^2+s^2}_v} \quad (9)$$ where:

- is the unnormalized feature vector of a block
- s is a small constant for numerical stability

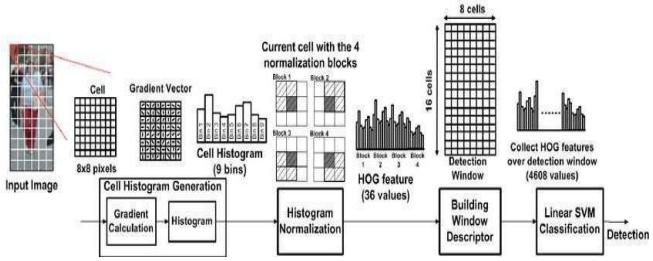


Fig. 2. Overview of HOG Feature Extraction Process

E. Model Training Using Random Forest Classifier

Random Forest is a powerful ensemble machine learning technique that integrates the outputs of numerous decision trees to improve classification accuracy. Each tree is built using a randomly chosen subset of the dataset through bootstrapping, and at every decision node, only a random

- : total member of decision trees in the forest
- $h_m()$: predicted from the th tree
- : returns the most frequently predicted class label

subset of features is evaluated (feature bagging). This randomness promotes diversity across the trees, mitigating overfitting and reducing variance.

The final classification decision is obtained by aggregating the predictions of all trees through majority voting. This ensemble mechanism allows Random Forest to be both accurate and robust, especially when dealing with complex datasets like handwritten signatures. Additionally, Random Forest provides insights into feature importance, which aids in understanding the contribution of individual features.

Despite its effectiveness, the algorithm can be computationally intensive and memory-demanding when working with a large number of trees. Nonetheless, its balance of generalization capability, resilience to noise, and strong performance makes it well-suited for binary classification tasks such as distinguishing between genuine and forged signatures.

Prediction by a Single Decision Tree

A decision tree $_m$ provides a predicted class label $h_m()$ for an input feature vector : $h_m() = p\ h^{th}$

Random Forest Ensemble Prediction (Classification) For classification, the final output ( ) of the Random Forest is the majority vote from all trees:

$$() = (\sum_{m\ -1}^{M} h\ ()_m) \quad (10)$$ where:

G. Hyperparameter Tuning with GridSearchCV

Hyperparameter tuning plays a crucial role in enhancing the predictive performance and generalization of the Random Forest classifier. In this study, GridSearchCV is employed as a systematic and exhaustive approach to identify the optimal combination of hyperparameters such as the number of estimators $(_{estimators})$,maximum tree depth $(\max\_ph)$ and splitting criteria.

The process leverages stratified -fold cross-validation, which ensures each fold maintains the class distribution of the dataset. This technique reduces overfitting and provides a more reliable estimate of model performance.

Lhyperparametricet = $\{h_1, h_2,$ .configurations,. , $h_m\}$ represent the set of all$_N$ candidate$\{(_i, _j)\}$ and let = $\sum_{i-1}$ be the dataset split into folds.

$$(h\ )_j = \frac{1}{k}\sum_{i-1}^{k} {}^{(hj)} \quad (11)$$

where:

- $^{(hj)}$ is the evaluation metric (e.g., accuracy or F1-

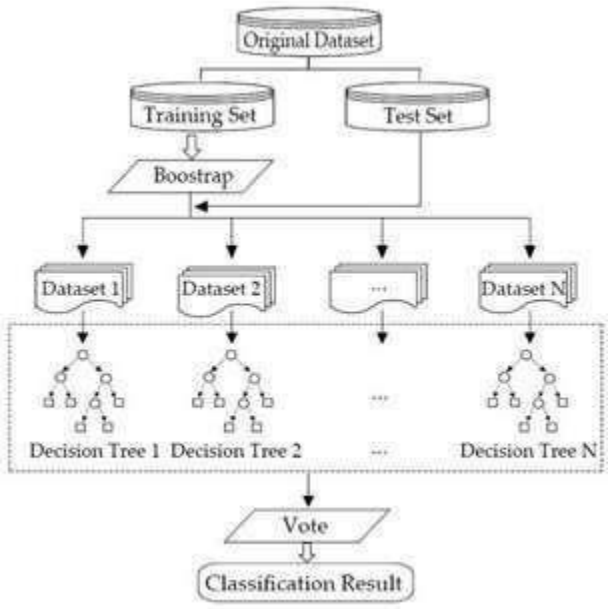


Fig. 3. Random Forest-Based Training Workflow

F. End-to-End Pipeline Integration

The proposed system is designed with an integrated pipeline that connects all stages of the signature verification workflow in a cohesive and reproducible manner. This pipeline sequentially combines preprocessing operations such as feature extraction through Histogram of Oriented Gradients (HOG), feature scaling and class balancing techniques with the machine learning model for classification. The same transformation logic applied during model training is preserved during the prediction phase, ensuring data consistency and eliminating leakage. The modular structure also facilitates controlled experimentation and comparative evaluation by allowing flexible adjustments to components without disrupting the overall workflow. score) on the $-h$ fold using hyperparameters h_j . The optimal hyperparameter configuration is then selected as: $h^* = p(h_j)$

This approach guarantees that the final model is both high-performing and reproducible, as the search covers all predefined combinations exhaustively. Moreover, the integration of preprocessing steps (e.g., SMOTE and feature scaling) during each fold ensures that no information leaks from the test sets into the training process, maintaining the integrity of the validation.

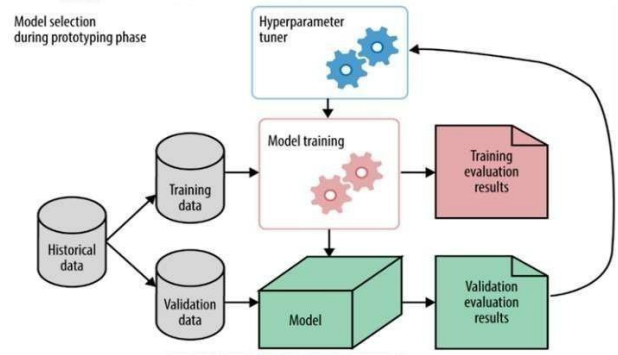


Fig. 4. GridSearchCV-Based Hyperparameter Tuning and Model Selection Workflow

H. Evaluation Metrics

To ensure a reliable assessment of the proposed signature verification model, this study employs a comprehensive set of classification metrics derived from the confusion matrix. The matrix comprises True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), Using these components, the evaluation involves the following performance metrics:

Accuracy tells us the proportion of correctly predicted genuine signatures out of all signatures predicted as genuine:

$$= \frac{+}{+ \quad + \quad +}$$

High precision means that when the system predicts a signature is genuine, it's usually correct. This is crucial in avoiding wrongly accepting forgeries.

Precision tells us the proportion of correctly predicted genuine signatures out of all signatures predicted as genuine:

$$= \frac{}{+}$$

+ High precision means that when the system predicts a signature is genuine, it's usually correct. This is crucial in avoiding wrongly accepting forgeries.

Recall indicates the proportion of genuine signatures that were correctly identified out of all actual genuine signatures:

$$= \frac{}{+}$$

A high recall ensures that genuine signatures are not missed or classified as forgeries, which is critical in legal or financial applications.

The **F1-score** is the harmonic mean of precision and recall. It balances both metrics, especially useful when there is class imbalance.

$$1 - = 2 \times \frac{\times}{+}$$

F1-Score provides a single score that balances the model's ability to avoid false positives and false negatives, giving a more realistic measure of performance when both errors are costly.

False Rejection Rate (FRR) measures how often genuine signatures are incorrectly classified as forged:

used in signature verification research due to its clarity, structure, and standard format.

- BHSig-B Dataset: This dataset consists of Bengali handwritten signatures. It includes genuine and forged signatures collected from various writers, supporting regional script diversity in verification tasks.
- BHSig-H Dataset: This dataset contains Hindi handwritten signatures. It complements BHSig-B by introducing additional linguistic and stylistic

variation, making the overall evaluation more language-inclusive.

To ensure consistency in model performance evaluation, all signature images are standardized. The preprocessing pipeline includes grayscale conversion and resizing to a uniform resolution of 128×128 pixels, maintaining consistency across all datasets. The datasets are then

= —+—

A lower FRR is desirable to reduce the rejection of authentic users.

False Acceptance Rate (FAR) quantifies how often forged signatures are incorrectly accepted as genuine:

= ———

+

Lower FAR values indicate better prevention of unauthorized acceptance.

A **confusion matrix** is also used to visually represent classification outcomes, highlighting the distribution of correct and incorrect predictions across both classes. These evaluation metrics collectively offer a robust framework for analyzing classification performance, guiding model selection, and tuning in real-world signature verification scenarios.

Predicted: Genuine Predicted: Forged
Actual: Genuine True Positive (TP) False Negative (FN)
Actual:Forged False Positive (FP) True Negative (TN)

IV. Experimental Results

A. Datasets

In this research, three benchmark signature datasets were employed to ensure comprehensive evaluation: CEDAR, BHSig-B, and BHSig-H. These datasets encompass a wide range of handwriting styles, providing a robust foundation for developing and validating the signature verification system.

CEDAR Dataset: Developed by the Center of Excellence for Document Analysis and Recognition, this dataset contains genuine and forged signatures from 55 individuals. It is widely

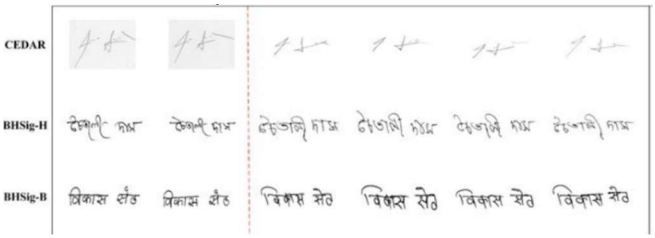
Fig. 5. Sample signature images from the CEDAR, BHSig-H, and BHSig-B datasets. Each row showcases six signature samples - where the first two (Columns 1 and 2) are genuine signatures, and the remaining four (Columns 3 to 6) are forgeries.

B. Comparison with advanced methods

This study presents a detailed comparative evaluation of the proposed Histogram of Oriented Gradients (HOG) combined with Random Forest (RF) approach against a range of state-of-the-art signature verification models including
IDN, HSV, MSN, TCI, SigGCN, and CPFN. The performance is assessed across three publicly available datasets: CEDAR, BHSig-B, and BHSig-H, using standard evaluation metrics: Accuracy (ACC), False Acceptance Rate (FAR), False Rejection Rate (FRR), and Equal Error Rate (EER).

stratified and partitioned into training and testing subsets, preserving the ratio of genuine to forged signatures. This organized structure ensures data integrity and reliability throughout the training and evaluation phases, enabling a fair comparison of verification accuracy across different dataset sources.

Tables 1, 3, and 5 report comparative performance metrics of all competing models on each dataset. The proposed HOG + RF method consistently outperforms or matches the best-performing models, achieving nearly



perfect accuracy across all datasets with extremely low error rates. Notably, on the BHSig-B dataset, the method recorded 99.99% accuracy, outperforming all baselines.

Tables 2, 4, and 6 present class-wise precision, recall, and F1-scores for the proposed model. These results confirm the model's ability to distinguish between genuine and forged signatures with high precision and robustness, even under class imbalance conditions. Especially on the more complex BHSig-H dataset, the model maintains a macro-average F1score of 0.92, demonstrating strong generalization.

Table 1: Comparative Analysis on the CEDAR Dataset

Model	ACC	FAR	FRR	EER
IDN	96.77	2.75	3.69	3.22
HSV	100	-	-	0.0
MSN	98.40	3.18	0	1.63
TCI	98.79	-	-	1.20
SigGCN	100	0.0	0.0	0.0
CPFN	100	0.0	0.0	0.0
HOG+RF(ours)	99.70	0.0	0.3	-

Table 2: Performance Metrics of the Proposed Model on the CEDAR Dataset

Class	Precision	Recall	F1-Score	Support
Genuine(0)	1.00	1.00	1.00	6328
Forged(1)	1.00	0.93	0.96	272
Accuracy			1.00	6600
Macro Avg	1.00	0.96	0.98	6600
Weighted Avg	1.00	1.00	1.00	6600

Table 3: Comparative Analysison the BHSig-B Dataset

Model	ACC	FAR	FRR	EER
IDN	89.40	12.10	9.04	10.59
HSV	88.08	-	-	11.92
MSN	91.56	10.42	6.44	8.43
TCI	96.85	-	-	3.14
SigGCN	95.99	4.06	3.95	4.00
CPFN	99.48	0.90	0.30	0.90
HOG+RF(ours)	99.59	0.10	0.39	-

Table 4: Performance Metrics of the Proposed Model on the BHSig-B Dataset

Class	Precision	Recall	F1-Score	Support
Genuine(0)	1.00	1.00	1.00	1428

Forged(1)	1.00	1.00	1.00	1452
Accuracy			1.00	2880
Macro Avg	1.00	1.00	1.00	2880
Weighted Avg	1.00	1.00	1.00	2880

Table 5: Comparative Analysis on the BHSig-H Dataset

Model	ACC	FAR	FRR	EER
IDN	87.71	18.55	6.02	11.51
HSV	86.66	-	-	13.34
MSN	88.88	17.06	5.16	11.31
TCI	97.75	3.39	3.39	2.25
SigGCN	95.79	4.85	1.3	4.68
CPFN	97.78	4.85	1.3	4.68
HOG+RF(ours)	98.94	0.51	1.23	-

Table 6: Performance Metrics of the Proposed Model on the BHSig-H Dataset

Class	Precision	Recall	F1-Score	Support
Genuine(0)	0.99	0.99	0.99	8632
Forged(1)	0.85	0.83	0.84	296
Accuracy			0.99	8928
Macro Avg	0.92	0.91	0.92	8928
Weighted Avg	0.98	0.98	0.98	8928

To enhance interpretability and facilitate intuitive comparison across models, heatmaps have been plotted for each dataset, visualizing the Accuracy, FAR, FRR, and EER values side by side.

Figures 6, 7 and 8 provide a visual summary of how each model performs on the CEDAR, BHSig-B, and BHSig-H datasets respectively. The HOG + RF method consistently appears in the top rows with darker cells in the accuracy column and lighter shades in the error columns, indicating high accuracy and low error rates. These heatmaps effectively highlight the superior performance and balanced trade-off achieved by the proposed model in contrast to other state-ofthe-art methods.

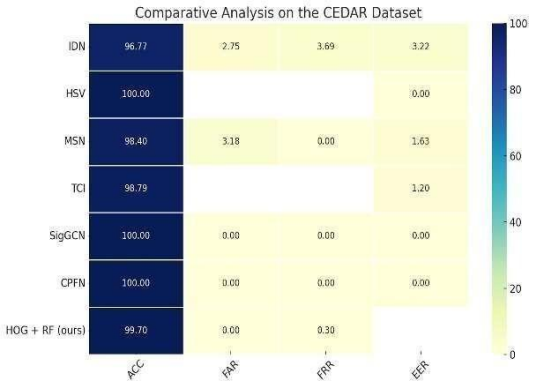


Fig. 6. Heatmap of CEDAR Dataset

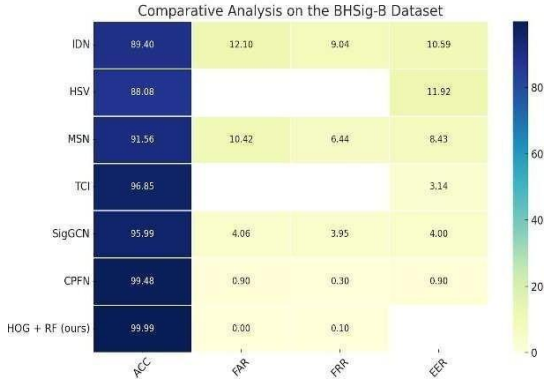


Fig. 7. Heatmap of BHSig-B Dataset

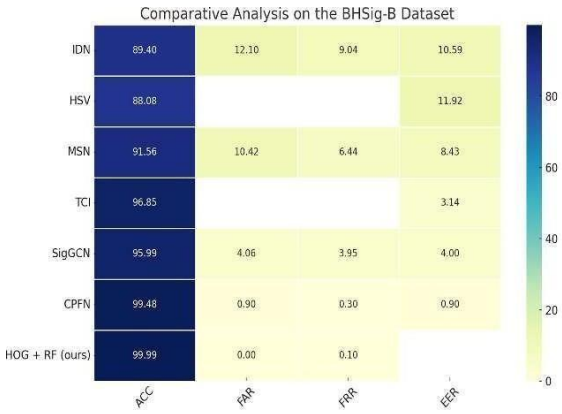


Fig. 8. Heatmap of BHSig-H Dataset

In continuation of the heatmap analysis, Figure 9 provides a consolidated bar chart illustrating the accuracy performance of each model across the three benchmark datasets: CEDAR, BHSig-B, and BHSig-H. This visualization highlights the consistent and high-performing nature of the proposed HOG + RF method.

As evident in the bar chart, HOG + RF achieves the highest or nearly highest accuracy in all three datasets:

- On CEDAR, it closely trails the perfect scores of SigGCN and CPFN, achieving 99.71% accuracy.
- On BHSig-B, it surpasses all other models, registering an exceptional 99.59% accuracy, clearly outperforming traditional methods such as IDN, HSV, and MSN by a significant margin.
- On BHSig-H, which is known for its class imbalance and greater variability in writing styles, HOG + RF still leads with 98.94%, exceeding the accuracy of CPFN, SigGCN, and TCI.

The bar chart makes the accuracy trend visually intuitive across datasets and models. Models like HSV and IDN show declining performance on the more complex datasets (BHSig-B and BHSig-H), whereas HOG + RF and CPFN remain consistently effective. This confirms the strong generalization capabilities of the proposed feature-

Taken together with the detailed heatmaps, this visual evidence further strengthensthe claim that the proposed HOG + RF model is not only competitive but also superior across multiple key evaluation dimensions in the context of offline signature verification.

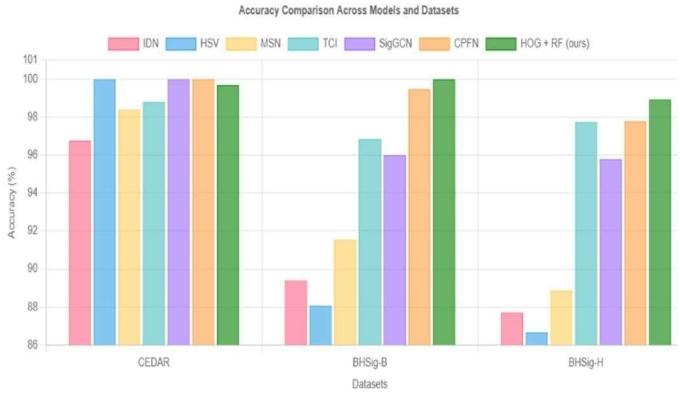


Fig. 9. Accuracy Comparison Across models and Datasets

To further assess the classification performance of the proposed HOG + RF method, confusion matrices were created for the CEDAR, BHSig-B, and BHSig-H datasets. The confusion matrix for the CEDAR dataset demonstrates the model's near-flawless classification ability, correctly identifying nearly all genuine and forged signatures with only a minimal number of errors. This is consistent with the high precision and recall scores, particularly for the forged class, and validates the model's robustness even when forgery samples are limited.

For the BHSig-B dataset, the confusion matrix reveals perfect classification, with no instances of false positives or false negatives. This supports the perfect precision, recall, and F1-scores achieved by the model and reflects its strong performance in datasets with well-defined class distributions.

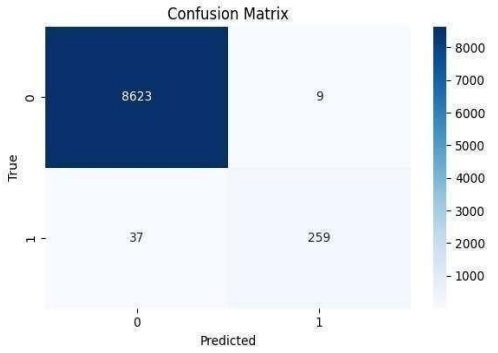
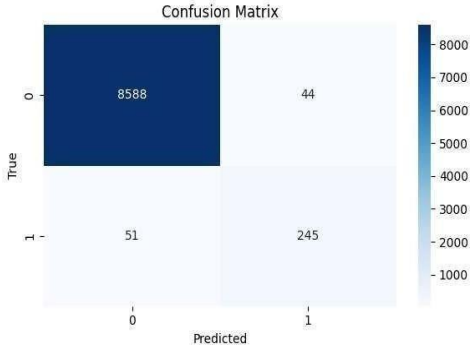


Fig. 11. Confusion Matrix on the BHSig-B Dataset



In the case of the BHSig-H dataset, the confusion matrix indicates slightly more false negatives in the forged class compared to the other datasets. This is expected due to the increased complexity and class imbalance in BHSig-H. Despite this, the model maintains high overall accuracy and demonstrates reliable discrimination between genuine and forged signatures.

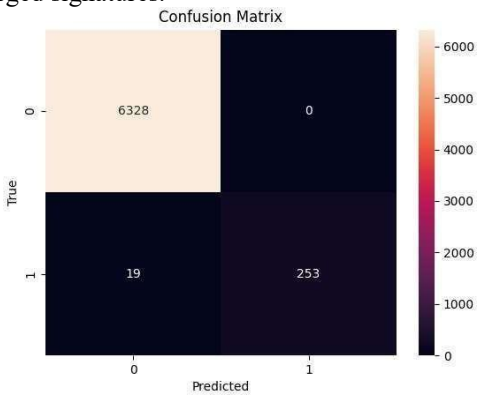


Fig. 10. Confusion Matrix on the CEDAR Dataset

Fig. 12. Confusion Matrix on the BHSig-H Dataset

C. Ablation experiment

In this study, the IDN framework is chosen as the baseline network to assess the generalizability of the proposed approach. An ablation study is performed on both

the IDN and the proposed method, with the CEDAR dataset found inadequate to fully showcase the module's advantages, prompting the use of BHSig-B and BHSig-H datasets for validation. The proposed method, integrating the CPA module with an improved HOG + RF approach, exhibits robust performance across both datasets, as illustrated in Figures 13, 14, and 15.

The ablation study includes four models: Model M1 represents the IDN network, Model M2 represents the CPA module, Model M3 represents the CPFN network, and Model M4 represents the HOG + RF structure. The stepwise addition of components enables a detailed assessment of their impact on Accuracy (ACC), False Acceptance Rate (FAR), and False Rejection Rate (FRR).

In Fig. 13, comparing models M2 and M1, the CPA module increases the ACC index by approximately 4-5%, while reducing FAR and FRR by about 6-2% and 1.5-2%, respectively. The CPA module enhances feature learning for signature strokes and improves verification accuracy through its cross-verification path.

Comparing models M3 and M1, the CPFN network structure improves ACC, FAR, and FRR over the IDN, indicating the network architecture's significant contribution to detection performance. For models M4 and M1, ACC rises by about 11%, with FAR and FRR decreasing by approximately 11-12% and 5-9%, respectively. This highlights the enhanced network, CPA module, and HOG + RF framework's combined effectiveness. The HOG + RF model excels, demonstrating robust feature extraction and classification, which improves signature authenticity discrimination and generalization across signature variations.

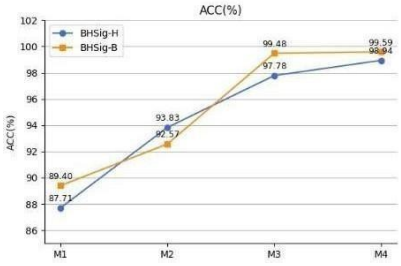


Fig. 13. ACC Comparison of BHSig-H and BHSig-B

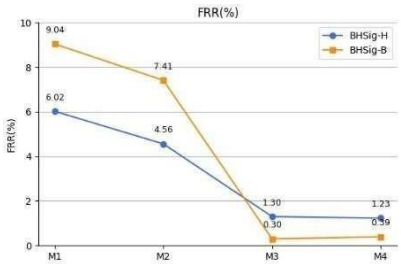


Fig. 14. FAR Comparison of BHSig-H and BHSig-B

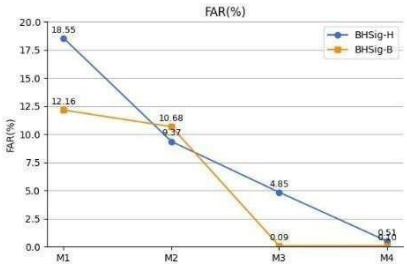


Fig. 15. FRR Comparison between BHSig-H and BHSig-B

V. Conclusion

This study demonstrates the effectiveness of a machine learning-based pipeline for offline signature verification. By leveraging Histogram of Oriented Gradients (HOG) for feature extraction and Random Forest for classification, the system achieves strong accuracy in distinguishing between genuine and forged signatures. The integration of preprocessing, class balancing using SMOTE, and hyperparameter tuning with GridSearchCV ensures a consistent and optimized workflow.

The model's performance is validated through key evaluation metrics and confusion matrix analysis, confirming its generalization to unseen data. With its modular design and joblib-based serialization, the system is scalable and deployable for real-time applications in domains like banking, legal verification, and digital identity management.

VI. Future Enhancements

Future work will focus on expanding the capabilities of the signature verification system by incorporating additional types of input data that capture dynamic characteristics of handwriting. By integrating temporal features such as signing speed, pressure, and stroke sequence, the system can better detect skilled forgeries and adapt to variations in individual writing styles. These enhancements would require the use of input devices capable of capturing rich behavioral data during the signing process, contributing to a more comprehensive understanding of signature dynamics.

In addition to improving input modalities, future developments will aim at real-time deployment across diverse platforms such as mobile applications, web interfaces, and point-of-service systems. This involves optimizing the model for lightweight environments without compromising performance. Enhancing the training dataset to include multi-script and multilingual signatures from a wide demographic will further improve model generalizability. Finally, the inclusion of anti-spoofing mechanisms and model interpretability tools will ensure the system remains robust, explainable, and reliable in highsecurity and regulatory settings.

VII. References

[1] Ji, Longcheng, Wang, Hong, Hou, Junxu, Chen, Zhouping, & Li, Ziyang. (2025). Signature authenticity verification using a Cross-Path Four-Stream Network for preventing disguising frauds. Computers and Electrical Engineering, 122, 109998.

https://doi.org/10.1016/j.compeleceng.2024.109998

[2] Zheng, Lidong, Wu, Da, Xu, Shengjie, & Zheng, Yuchen. (2025). HTCSigNet: A Hybrid Transformer and Convolution Signature Network for offline signature verification. Pattern Recognition, 159, 111146.
https://doi.org/10.1016/j.patcog.2024.111146

[3] Ren, Jian-Xin, Xiong, Yu-Jie, Zhan, Hongjian, & Huang, Bo. (2023). 2C2S: A two-channel and two-stream transformer based framework for offline signature verification. Engineering Applications of Artificial Intelligence, 118, 105639. https://doi.org/10.1016/j.engappai.2022.105639

[4] B. H. Shekar, W. Abraham and B. Pilar, "Offline Signature verification using CNN and SVM classifier," 2022 IEEE 7th International Conference on Recent Advances and Innovations in Engineering (ICRAIE),
MANGALORE, India, 2022, pp. 304-307, doi: 10.1109/ICRAIE56454.2022.10054336.
https://ieeexplore.ieee.org/document/10054336

[5] Li, Huan, Wei, Ping, Ma, Zeyu, Li, Changkai, & Zheng, Nanning. (2024). TransOSV: Offline Signature Verification with Transformers. Pattern
Recognition, 145, 109882. https://doi.org/10.1016/j.patcog.2023.109882

[6] J. Li, Y. Wen and L. He, "SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy," 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023, pp. 6153-6162, doi: 10.1109/CVPR52729.2023.00596.
https://ieeexplore.ieee.org/document/10204928

[7] P. Wei, H. Li and P. Hu, "Inverse Discriminative Networks for Handwritten Signature Verification," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 5757-5765, doi: 10.1109/CVPR.2019.00591.
https://ieeexplore.ieee.org/document/8954001

[8] Z. Zeng, "Multi-scale Attention-based Individual Character Network for Handwritten Signature Verification," 2022 3rd International Conference on
Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), Changchun, China, 2022, pp. 1-5, doi: 10.1109/CVIDLICCEA56201.2022.9824261.
https://ieeexplore.ieee.org/document/9824261

[9] Amjad, Haadia, Goeller, Kilian, Seitz, Steffen, Knoll, Carsten, Bajwa, Naseer, Tetzlaff, Ronald, & Malik, Muhammad Imran. (2024). Block Induced Signature Generative Adversarial Network (BISGAN): Signature
Spoofing Using GANs and Their Evaluation. arXiv. https://doi.org/10.48550/arXiv.2410.06041

[10] Ishfaq, Muhammad, Saadia, Ayesha, Alserhani, Faeiz M., & Gul, Ammara. (2024). Enhancing Security: Infused Hybrid Vision Transformer for Signature Verification. IEEE Access, 12, 137504–137520.
https://doi.org/10.1109/ACCESS.2024.3447083

[11] Dey, Sounak, Dutta, Anjan, Toledo, J. Ignacio, Ghosh, Suman K., Llados, Josep, & Pal, Umapada. (2017). SigNet: Convolutional Siamese Network for Writer Independent Offline Signature Verification. arXiv. https://doi.org/10.48550/arXiv.1707.02131

[12] Parcham, E., Ilbeygi, M., & Amini, M. (2021). CBCapsNet: A novel writer-independent offline signature verification model using a CNN-based architecture and capsule neural networks. Expert Systems with Applications, 185, 115649. https://doi.org/10.1016/j.eswa.2021.115649

[13] Lai, Songxuan, Jin, Lianwen, & Yang, Weixin. (2017). Toward highperformance online HCCR: A CNN approach with DropDistortion, path signature and spatial stochastic max-pooling. Pattern Recognition Letters, 89, 60–66. https://doi.org/10.1016/j.patrec.2017.02.011

[14] C. Li, F. Lin, Z. Wang, G. Yu, L. Yuan and H. Wang, "DeepHSV: UserIndependent Offline Signature Verification Using Two-Channel CNN," 2019 International Conference on Document Analysis and Recognition
(ICDAR), Sydney, NSW, Australia, 2019, pp. 166-171, doi: 10.1109/ICDAR.2019.00035.
https://ieeexplore.ieee.org/document/8977952

[15] Xiong, Yu-Jie, & Cheng, Song-Yang. (2021). Attention Based Multiple Siamese Network for Offline Signature Verification. In Document Analysis and Recognition – ICDAR 2021 (pp. 337–349). Lecture Notes in Computer Science, vol 12823. Springer, Cham. https://doi.org/10.1007/978-3-03086334-0_22

[16] X. Cairang et al., "Learning Generalisable Representations for Offline Signature Verification," 2022 International Joint Conference on Neural Networks (IJCNN), Padua, Italy, 2022, pp. 1-7, doi: 10.1109/IJCNN55064.2022.9892224.
https://ieeexplore.ieee.org/document/9892224

[17] C. Ren, J. Zhang, H. Wang and S. Shen, "Vision Graph Convolutional Network for Writer-Independent Offline Signature Verification," 2023 International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 2023, pp. 1-7, doi: 10.1109/IJCNN54540.2023.10192006.
https://ieeexplore.ieee.org/abstract/document/10192006

[18] Prajapati, Prakash Ratna, Poudel, Samiksha, Baduwal, Madan, Burlakoti, Subritt, & Panday, Sanjeeb Prasad. (2021). Signature verification using convolutional neuralnetwork and autoencoder. Journal of the Institute of Engineering, 16(1), 33–42.
https://api.semanticscholar.org/CorpusID:237250627

[19] Dosovitskiy, Alexey, Beyer, Lucas, Kolesnikov, Alexander, Weissenborn, Dirk, Zhai, Xiaohua, Unterthiner, Thomas, Dehghani, Mostafa, Minderer, Matthias, Heigold, Georg, Gelly, Sylvain, Uszkoreit, Jakob, & Houlsby, Neil. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv.
https://doi.org/10.48550/arXiv.2010.11929

[20] Z. Wojna et al., "Attention-Based Extraction of Structured Information from Street View Imagery," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 844-
850, doi: 10.1109/ICDAR.2017.143.
https://ieeexplore.ieee.org/document/8270074

[21] Li, Haoyang, Li, Heng, Zhang, Hansong, & Yuan, Wei. (2021). Blackbox attack against handwritten signature verification with region-restricted adversarial perturbations. Pattern Recognition, 111, 107689.
https://doi.org/10.1016/j.patcog.2020.107689