

A Hybrid Deep Visual and Structural Graph-Based Approach for Writer-Dependent Offline Signature Verification

Dr. Annapurna H

Associate Professor and HoD: Dept. of Computer Science Yuvaraja's College, University of Mysore, Mysuru
Karnataka, 570005, India

Abstract - This paper proposes a hybrid writer-dependent offline signature verification framework that integrates deep visual features with structural graph-based representations. A convolutional neural network inspired by SigNet is employed to extract global visual descriptors, while a graph neural network captures stroke-level structural characteristics from skeletonized signatures. The two feature sets are fused and refined using ReliefF-based feature selection. Writer-specific models are then constructed using centroid-based distance measures and adaptive thresholds estimated through Equal Error Rate (EER) optimization. Extensive experiments conducted on the CEDAR offline signature dataset using multiple train and test splits demonstrate that the proposed approach consistently achieves low error rates and high verification accuracy of 93.40%. The results confirm that combining visual and structural information significantly improves robustness against skilled forgeries.

Keywords - Offline signature verification; writer-dependent threshold; deep learning; graph neural networks; feature fusion;

I. INTRODUCTION

Offline handwritten signature verification remains a challenging biometric problem due to significant intra-writer variability and the presence of skilled forgeries. Despite the increasing adoption of biometric technologies, handwritten signatures continue to be widely used in financial, administrative, and legal applications because of their social acceptance and ease of acquisition. Unlike physiological biometrics, signatures are behavioral in nature and are therefore subject to variations in writing style, mood, and writing conditions, making reliable verification particularly difficult. This challenge is further amplified in offline scenarios, where dynamic information such as stroke order, writing speed, and pen pressure is unavailable.

Early approaches to offline signature verification primarily relied on handcrafted features, including geometric descriptors, directional histograms, and texture-based representations. While these methods demonstrated reasonable performance, their effectiveness often deteriorated in the presence of skilled forgeries. The advent of deep learning has significantly advanced the field by enabling automatic and hierarchical feature learning directly from raw signature images. In particular, convolutional neural networks (CNNs), such as SigNet, have shown strong capability in learning discriminative global representations for signature verification tasks [1].

However, CNN-based approaches predominantly focus on visual appearance and may not fully capture the structural relationships among strokes that characterize individual writing styles. To address this limitation, graph-based representations have been explored as a complementary modelling strategy, as they explicitly encode spatial and topological relationships between stroke components. Recent studies demonstrate that graph neural networks (GNNs) are effective in capturing such structural dependencies and can substantially enhance verification performance when combined with visual features [2].

Motivated by these observations, this work proposes a hybrid writer-dependent verification framework that integrates deep visual features with structural graph-based representations. The proposed approach leverages the complementary strengths of convolutional and graph-based learning to improve robustness against skilled forgeries. The remainder of this paper is organized as follows: Section 2 reviews related work; Section 3 describes the proposed methodology; Section 4 presents the experimental setup and results; Section 5 provides a comparative analysis with existing methods; and Section 6 concludes the paper.

II. RELATED WORKS

Offline handwritten signature verification has been extensively investigated due to its importance in biometric authentication systems. Early studies primarily relied on handcrafted features combined with traditional classifiers. Kalera et al. [3] proposed a distance-based verification framework using geometric and statistical descriptors, while Chen and Srihari [4] explored graph-based matching techniques to model structural relationships between signature strokes. Although these methods achieved moderate success, their performance was often limited by sensitivity to intra-writer variability and the inability to generalize well to skilled forgeries.

With the advent of deep learning, convolutional neural networks (CNNs) significantly advanced the field by enabling automatic feature learning directly from raw signature images. Hafemann et al. [5] demonstrated that CNNs can learn highly discriminative representations for offline signature verification, outperforming traditional handcrafted approaches. Subsequent works adopted Siamese and triplet network architectures to learn similarity metrics between signature pairs, further improving

verification performance under writer-dependent and writer-independent scenarios [6].

Despite these advances, most CNN-based approaches primarily focus on global appearance features and often neglect the underlying structural information present in handwriting. To address this limitation, several studies have explored graph-based representations that explicitly model stroke connectivity and spatial relationships. Graph-based techniques have been shown to capture discriminative structural cues such as stroke intersections and continuity, which are difficult to model using convolutional filters alone. In this context, graph neural networks (GNNs) have emerged as a powerful tool for learning relational patterns in handwritten data [7].

Recent works have investigated hybrid approaches that combine deep visual features with structural or topological representations. Maergner et al. [8] demonstrated that integrating structural descriptors with deep features improves robustness against skilled forgeries. Similarly, hybrid deep-graph frameworks have been shown to outperform single-modality approaches by leveraging complementary information from both appearance and structure [9]. However, many of these methods employ complex architectures or lack a unified verification strategy, limiting their practical applicability.

With these observations, the present work proposes a unified writer-dependent verification framework that integrates CNN-based visual representations with graph-based structural features. Unlike existing methods, the proposed approach employs an adaptive writer-specific thresholding mechanism based on Equal Error Rate (EER) analysis and evaluates performance across multiple training–testing splits. This design enables robust verification while maintaining computational efficiency and practical applicability.

III. PROPOSED METHODOLOGY

The proposed framework consists of six major stages:

- Preprocessing
- Feature Extraction
- Feature Fusion
- Feature Selection
- Writer-dependent Threshold estimation
- Writer-dependent verification

A schematic overview of the proposed system is illustrated in Fig. 1.

A. Preprocessing

Each signature image is first converted to grayscale and denoised using median filtering to suppress impulsive noise while preserving edge information [10]. Subsequently, Otsu's thresholding method is applied to obtain a binary representation by automatically determining an optimal threshold that maximizes inter-class variance between foreground and background pixels [11]. Morphological operations are then employed to remove small artifacts and enhance stroke continuity. All images are resized to a fixed resolution of 224×224 pixels to ensure uniformity across samples. Finally, skeletonization is performed to reduce the binary signature to a one-pixel-wide representation while preserving its essential structural and topological characteristics, which are crucial for subsequent graph construction and analysis.

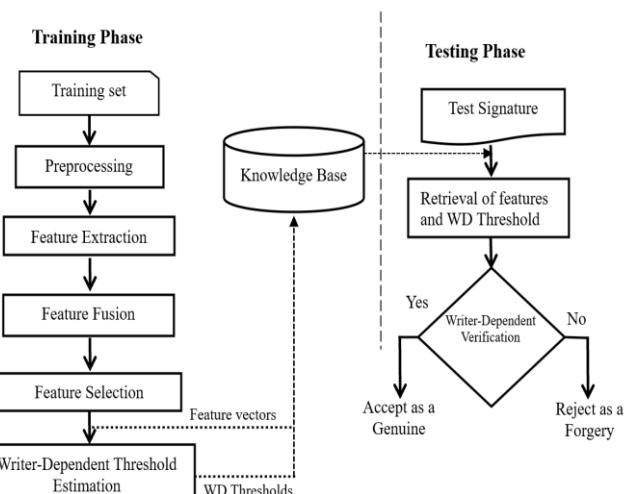


Fig. 1. Block diagram of the proposed methodology

B. Feature Extraction

After preprocessing, the cleaned and normalized signature images are used for feature extraction, which aims to transform the visual information into meaningful numerical representations suitable for verification.

In the proposed work, feature extraction is performed in two complementary stages:

- 1) Deep visual feature extraction
- 2) Structural feature extraction

1) Deep visual feature extraction :

To effectively capture the global visual characteristics of handwritten signatures, a deep convolutional neural network (CNN) inspired by the SigNet architecture is employed in this work. SigNet, originally proposed for offline signature verification, has demonstrated strong capability in learning discriminative representations directly from raw signature images by exploiting hierarchical feature learning mechanisms [5]. The motivation for adopting a CNN-based approach lies in its ability to automatically extract meaningful patterns without relying on handcrafted descriptors, which often fail to generalize across diverse writing styles.

The network architecture consists of multiple convolutional and pooling layers arranged in a hierarchical manner. In the initial layers, convolutional filters learn low-level features such as edges, stroke boundaries, and local orientation patterns. As the depth of the network increases, higher-level layers capture more abstract and semantically meaningful information, including stroke arrangements, global shape characteristics, and writing style variations that are distinctive to individual writers.

Max-pooling layers are interleaved between convolutional layers to progressively reduce spatial resolution while retaining salient features. This design not only improves computational efficiency but also introduces a degree of translation invariance, making the learned representations more robust to minor variations in writing position and scale. The extracted feature maps are subsequently flattened and passed through fully

connected layers, which integrate spatial and contextual information into a compact representation.

The output of the final fully connected layer is a fixed-length feature vector of 2048 dimensions, serving as a high-level descriptor of the signature's visual appearance. This representation effectively encodes global attributes such as stroke distribution, shape consistency, and texture patterns, which are crucial for distinguishing genuine signatures from skilled forgeries. Compared to handcrafted features, the deep representations learned by the CNN exhibit stronger discriminative power and improved generalization across different writers.

By leveraging the representational strength of deep convolutional networks, the proposed framework captures the global visual structure of signatures and provides a robust foundation for subsequent feature fusion with structural graph-based representations.

2) Structural feature extraction:

To capture the structural characteristics of handwritten signatures, skeletonized images are transformed into graph-based representations. In this formulation, each skeleton pixel is modelled as a node, while edges are established between spatially adjacent pixels using an 8-neighborhood connectivity rule. This representation preserves essential structural properties such as stroke continuity, curvature, branching patterns, and junction points, which play a critical role in distinguishing genuine signatures from skilled forgeries.

Each node is initially represented by a low-dimensional feature vector corresponding to its normalized spatial coordinates (x, y). These coordinates encode the relative position of stroke elements within the signature and provide a spatial reference for subsequent learning. To model the relationships among neighboring nodes, a Graph Convolutional Network (GCN) is employed. GCNs enable effective learning from non-Euclidean data by propagating and aggregating information across connected nodes, thereby capturing both local and global structural dependencies [7].

Through successive graph convolution layers, each node aggregates information from its local neighborhood, allowing the network to learn higher-level structural patterns such as stroke continuity, curvature flow, and junction complexity. This hierarchical message-passing mechanism enables the network to model the topological organization of handwriting more effectively than conventional grid-based approaches.

In the proposed framework, the GCN produces a fixed-length structural embedding of 128 dimensions. This embedding is obtained through a global pooling operation that aggregates node-level features into a single vector, effectively summarizing the overall structural characteristics of the signature. The resulting representation encodes several discriminative attributes, including stroke connectivity patterns, junction behavior, global stroke distribution, and structural complexity.

These graph-based structural features complement the appearance-based representations extracted by the convolutional neural network. While CNNs focus on texture, shape, and visual composition, the GNN emphasizes relational and topological information that is less sensitive to writing style variations. The integration of these complementary feature modalities enhances

robustness against skilled forgeries and improves overall verification performance.

IV. FEATURE FUSION

To combine complementary visual and structural information, deep visual and graph-based features are fused at the feature level. The deep visual representation consists of a 2048-dimensional vector extracted from a convolutional neural network, capturing global appearance attributes such as stroke shape and texture. In parallel, a 128-dimensional structural feature vector is obtained from a graph neural network, encoding stroke connectivity, spatial arrangement, and junction patterns.

The two feature sets are concatenated to form a unified 2176-dimensional representation that jointly models appearance and structural characteristics. This fusion enables the system to leverage both global visual cues and fine-grained topological information, resulting in a more discriminative representation for distinguishing genuine signatures from skilled forgeries. The fused feature vector serves as the input to the subsequent feature selection and verification stages.

V. FEATURE SELECTION

After feature fusion, the resulting high-dimensional representation is further refined using the ReliefF feature selection algorithm. The input to the ReliefF module is the fused feature vector obtained by concatenating the deep visual features (2048 dimensions) and the structural graph-based features (128 dimensions), resulting in a combined feature space of 2176 dimensions for each signature sample.

ReliefF is a supervised, instance-based feature selection method designed to estimate the relevance of individual features based on their ability to discriminate between classes while preserving intra-class consistency [12]. For each training sample, the algorithm identifies a set of nearest neighbors belonging to the same class (near-hits) and to different classes (near-misses). Feature weights are then updated by analyzing how feature values vary between these neighboring samples. Features that exhibit large inter-class differences and small intra-class variations are assigned higher importance scores, whereas features that contribute little to class discrimination are penalized.

By ranking features according to their discriminative strength, ReliefF enables the selection of the most informative subset while suppressing redundant or noisy dimensions. This process significantly reduces the dimensionality of the fused feature space and enhances generalization performance. In this work, the top-ranked features are retained to form a compact and discriminative representation, which serves as the input for the subsequent writer-dependent verification stage.

VI. WRITER-DEPENDENT THRESHOLD ESTIMATION

In the proposed framework, verification is performed in a writer-dependent manner, where an independent decision model is constructed for each enrolled writer. During the training phase, only genuine signature samples belonging to a specific writer are used. This strategy allows the system to learn writer-specific characteristics while avoiding interference from other writers' writing styles.

Let $f_i \in \mathbb{R}^d, i = 1, 2, \dots, N$ denote the selected feature vectors obtained after feature fusion and ReliefF-based feature selection, where d represents the dimensionality of the selected feature space and N is the number of genuine training samples for a given writer.

A representative model for each writer is constructed by computing the centroid of the corresponding feature vectors:

$$\mu_w = \frac{1}{N} \sum_{i=1}^N f_i \quad (1)$$

This centroid serves as a compact prototype that characterizes the typical writing behavior of the writer in the learned feature space.

To establish an appropriate decision boundary, an adaptive threshold is estimated for each writer using the Equal Error Rate (EER) criterion. Distances between the centroid and both genuine and forgery samples are computed, and the threshold is selected at the point where the False Acceptance Rate (FAR) equals the False Rejection Rate (FRR). This approach ensures a balanced trade-off between security and usability while accommodating individual variations in handwriting style.

The resulting writer-specific threshold enables robust and personalized verification, allowing the system to effectively distinguish genuine signatures from skilled forgeries.

VII. WRITER-DEPENDENT VERIFICATION

In this work, verification is performed using a writer-dependent distance-based strategy, where each enrolled writer is associated with an individual decision model. Unlike global classifiers that attempt to learn a universal boundary across all writers, the proposed approach constructs writer-specific reference models and thresholds. This allows the system to better capture individual writing characteristics and effectively handle intra-writer variability. The verification decision is made by measuring the similarity between a query signature and the corresponding writer's learned representation.

During the verification stage, a query signature undergoes the same preprocessing, feature extraction, and feature selection steps as those applied during training.

Let $f_q \in \mathbb{R}^d$ denote the resulting feature vector of the query signature, where d represents the dimensionality of the selected feature space.

The similarity between the query signature and the enrolled writer model is measured using the Euclidean distance between the query feature vector and the writer-specific centroid:

$$D(f_q, \mu_w) = \|f_q - \mu_w\|_2 \quad (2)$$

where μ_w denotes the centroid computed from the genuine training samples of the corresponding writer.

The verification decision is then made by comparing the computed distance with a writer-specific threshold T_w , estimated during training using the Equal Error Rate (EER) criterion.

If the condition

$$D(f_q, \mu_w) \leq T_w \quad (3)$$

is satisfied, the signature is accepted as genuine; otherwise, it is classified as a forgery.

This writer-dependent verification strategy provides an efficient and interpretable decision mechanism. By employing individualized centroids and adaptive thresholds, the system effectively captures intra-writer variability while maintaining strong discrimination against skilled forgeries.

VIII. EXPERIMENTAL SETUP AND RESULTS

This section presents the experimental setup and performance evaluation of the proposed offline signature verification framework. All experiments are conducted on the CEDAR offline signature dataset [3], which contains signatures from 55 writers. Each writer contributes 24 genuine signatures and 24 skilled forgeries, resulting in a total of 2,640 samples.

A. Experimental Setup

To evaluate the robustness and generalization capability of the proposed approach, experiments are performed using five different training-testing splits: 30–70, 40–60, 50–50, 60–40, and 70–30. For each split, only genuine signatures are used during training to construct writer-specific models, while both genuine and skilled forgeries are used during testing.

This experimental protocol ensures an independent evaluation for each writer and reflects realistic deployment scenarios where only a limited number of genuine samples are available for user enrollment. The use of multiple training-testing configurations further enables analysis of system behavior under varying levels of training data availability.

B. Evaluation Metrics

System performance is assessed using four widely adopted biometric evaluation metrics: False Acceptance Rate (FAR), False Rejection Rate (FRR), Equal Error Rate (EER), and Overall Accuracy. FAR measures the proportion of forged signatures incorrectly accepted as genuine, while FRR indicates the proportion of genuine signatures incorrectly rejected. EER corresponds to the operating point where FAR and FRR are equal and serves as a reliable indicator of verification performance. Overall accuracy reflects the proportion of correctly classified samples across all test cases.

The verification decision is made using a distance-based strategy with writer-dependent thresholds determined using the EER criterion. This ensures fair and consistent evaluation across all writers.

C. Results and Discussion

The experimental results obtained under different training-testing splits are summarized in Table I, while the corresponding accuracy trends are illustrated in Fig. 2.

From the Table I it is clear that, increasing the proportion of training samples consistently improves verification performance. In particular, lower EER values and higher accuracy are achieved as the training ratio increases from 30% to 70%, indicating improved modelling of writer-specific characteristics.

TABLE I. RESULTS OBTAINED FOR DIFFERENT TRAINING AND TESTING SPLITS ON THE CEDAR DATASET

Testing Set (%)	Testing Set (%)	FAR (%)	FRR (%)	EER (%)	Accuracy (%)
30	70	8.56	8.77	8.66	91.34
40	60	7.52	8.85	8.18	91.82
50	50	6.67	7.27	6.97	93.03
60	40	6.55	7.82	7.18	92.82
70	30	5.91	7.27	6.59	93.41

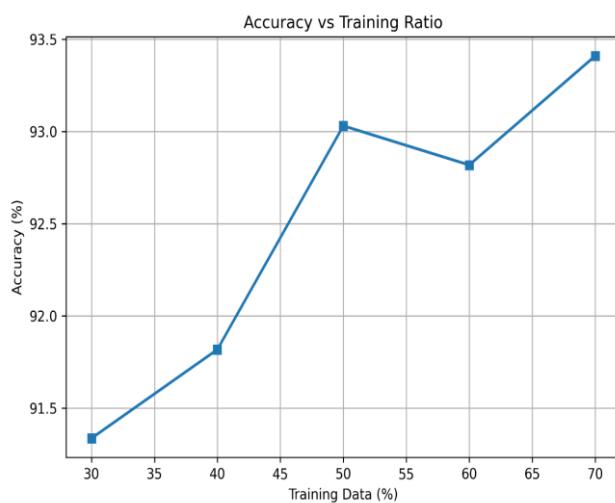


Fig. 2. Accuracy obtained for different Training and Testing splits on the CEDAR dataset

The results demonstrate that the proposed framework effectively benefits from additional training data, enabling more reliable discrimination between genuine signatures and skilled forgeries. The integration of deep visual features and structural graph-based representations plays a crucial role in achieving this performance improvement. By capturing both appearance-based and topological characteristics, the system exhibits enhanced robustness against skilled forgery attempts.

Furthermore, the application of ReliefF-based feature selection contributes to improved stability and discrimination by reducing redundancy and emphasizing informative features. Overall, the proposed approach demonstrates strong verification performance across all experimental settings, confirming its effectiveness and suitability for real-world offline signature verification applications.

IX. CONCLUSION

This paper presented a writer-dependent offline signature verification framework that integrates deep visual features with structural graph-based representations. By combining convolutional neural networks and graph neural networks, the proposed approach effectively captures both appearance and structural characteristics of handwritten signatures. The use of ReliefF-based feature selection and writer-specific thresholding further enhances discriminative capability and robustness.

Experimental results on the CEDAR dataset demonstrate consistent performance improvements across different training–testing splits, with reduced error rates and improved accuracy as more genuine samples are used for training. The results confirm the effectiveness of the proposed method in handling intra-writer variability and detecting skilled forgeries.

Future work will focus on extending the framework with advanced deep learning architectures and evaluating its generalization across larger and more diverse datasets.

REFERENCES

- [1] Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2017). Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognition*, 70, 163–176. <https://doi.org/10.1016/j.patcog.2017.05.012>
- [2] Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *Proceedings of the International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1609.02907>
- [3] Kalera, M. K., Zhang, B., & Srihari, S. N. (2004). Offline signature verification and identification using distance statistics. *International Journal of Pattern Recognition and Artificial Intelligence*, 18(7), 1339–1360. <https://doi.org/10.1142/S0218001404003630>
- [4] Chen, S., & Srihari, S. N. (2006). A new off-line signature verification method based on graph matching. *Proceedings of the 18th International Conference on Pattern Recognition (ICPR)*, 1, 86–89. <https://doi.org/10.1109/ICPR.2006.825>
- [5] Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2017). Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognition*, 70, 163–176. <https://doi.org/10.1016/j.patcog.2017.05.012>
- [6] Bromley, J., Bentz, J. W., Bottou, L., Guyon, I., LeCun, Y., Moore, C., & Shah, R. (1994). Signature verification using a Siamese time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, 7(4), 669–688. <https://doi.org/10.1142/S0218001493000339>
- [7] Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1609.02907>
- [8] Maergner, P., Sabourin, R., & Plamondon, R. (2019). Combining structural and deep features for offline signature verification. *Pattern Recognition Letters*, 125, 527–533. <https://doi.org/10.1016/j.patrec.2019.07.018>
- [9] Zhang, Y., Chen, X., & Huang, K. (2022). Hybrid deep representation learning for offline handwritten signature verification. *Expert Systems with Applications*, 198, 116825. <https://doi.org/10.1016/j.eswa.2022.116825>
- [10] Huang, T. S., Yang, G. J., & Tang, G. Y. (1979). A fast two-dimensional median filtering algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 27(1), 13–18. <https://doi.org/10.1109/TASSP.1979.1163188>
- [11] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
- [12] Kononenko, I. (1994). Estimating attributes: Analysis and extensions of Relief. *European Conference on Machine Learning (ECML)*, 171–182. https://doi.org/10.1007/3-540-57868-4_57