International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

A Deep Feature Reliability-Based Framework for Offline Handwritten Signature Verification

Dr. Annapurna H
Associate Professor & HoD, Department of Computer Science, Yuvaraja's College(Autonomous)
University of Mysore, Karnataka, India

ABSTRACT: Offline handwritten signature verification is widely employed as a biometric authentication technique in financial, legal, and administrative applications. However, accurately distinguishing genuine signatures from skilled forgeries remains a challenging task because of natural intra-writer variations and the inconsistent discriminative capability of deep feature representations. This paper presents a Deep Feature Reliability-Based Framework for Offline Handwritten Signature Verification that enhances verification performance by identifying and utilizing stable writer-specific deep features. Initially, signature images are preprocessed and represented using 2048-dimensional deep features extracted from a pre-trained ResNet50 network. A feature reliability estimation scheme is then introduced to evaluate the statistical consistency of individual feature dimensions across genuine signatures of each writer. Based on the estimated reliability scores, the most reliable features are selected and incorporated into a weighted cosine similarity measure for signature verification. The proposed framework is evaluated on the CEDAR offline handwritten signature dataset using three training–testing splits of 25–75, 50–50, and 75–25. Experimental results demonstrate that the proposed method achieves its best performance with the 75–25 split, obtaining an accuracy of 77.61%, precision of 83.77%, recall of 68.48%, and an F1-score of 75.33%. The findings indicate that incorporating feature reliability into the verification process improves the robustness of deep feature representations and provides an effective framework for offline handwritten signature verification.

KEYWORDS: Offline handwritten signature verification, biometrics, ResNet50, deep feature extraction, feature reliability estimation, reliable feature selection, weighted cosine similarity.

INTRODUCTION

Handwritten signatures continue to be one of the most widely accepted biometric traits for personal authentication because of their legal validity, ease of acquisition, and extensive use in banking, financial transactions, business contracts, and administrative applications (Jain et al., 2004; Hameed et al., 2021). Automatic signature verification systems have therefore attracted considerable research interest as they provide a reliable and efficient alternative to manual verification, which is often subjective, time-consuming, and susceptible to human error (Hameed et al., 2021). Based on the mode of acquisition, signature verification techniques are broadly classified into online and offline approaches. Among these, offline handwritten signature verification is more challenging because only the static image of the signature is available, without dynamic information such as writing speed, pen pressure, stroke order, or pen trajectory (Hafemann et al., 2017).

The performance of offline signature verification is primarily affected by two major challenges: significant intra-writer variation and high similarity between genuine signatures and skilled forgeries (Hafemann et al., 2017). Genuine signatures produced by the same individual may vary naturally due to writing conditions, while skilled forgeries are intentionally designed to imitate the visual appearance of genuine signatures. These challenges make accurate discrimination between genuine and forged signatures a difficult pattern recognition problem.

Early offline signature verification methods relied on handcrafted features, including geometric, statistical, contour, and texture descriptors, combined with conventional machine learning classifiers. Although these methods demonstrated satisfactory performance under controlled conditions, handcrafted representations often failed to capture the complex variations present in handwritten signatures (Hameed et al., 2021). Recent advances in deep learning, particularly Convolutional Neural Networks (CNNs), have significantly improved verification performance by automatically learning hierarchical and discriminative feature

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

representations directly from signature images (Hafemann et al., 2017). Pre-trained deep networks such as ResNet have further enhanced feature extraction through transfer learning, enabling robust representation even with relatively limited signature datasets. Despite the effectiveness of deep feature extraction, most existing approaches utilize the complete feature vector for verification by assigning equal importance to every extracted feature. In practice, however, not all deep features contribute equally to writer discrimination. Some features consistently represent stable writer-specific characteristics across genuine signatures, whereas others are influenced by natural writing variations, image noise, or redundant information. Consequently, treating all features equally may reduce the discriminative capability of the verification system.

To address this limitation, this paper proposes a Deep Feature Reliability-Based Framework for Offline Handwritten Signature Verification. The proposed framework employs a pre-trained ResNet50 network to extract deep feature representations and introduces a Feature Reliability Estimation stage to identify statistically stable and discriminative features. Based on the estimated reliability, only the most informative features are retained for signature verification, thereby reducing the influence of unstable or redundant features.

The remainder of this paper is organized as follows. The second Section reviews the related work on offline handwritten signature verification. Section three presents the proposed Deep Feature Reliability-Based Framework, including image preprocessing, deep feature extraction, feature reliability estimation, reliable feature selection, and signature verification. Section four describes the experimental setup, evaluation metrics, and discusses the obtained results. Finally, the last Section concludes the paper and outlines potential directions for future research.

LITERATURE SURVEY

Offline handwritten signature verification has been extensively studied as a behavioural biometric authentication problem because of its widespread use in financial, legal, and administrative applications. Early research primarily relied on handcrafted feature extraction techniques, including geometric, statistical, contour, texture, and directional descriptors, which were combined with conventional machine learning classifiers such as Support Vector Machines (SVMs), Hidden Markov Models (HMMs), Artificial Neural Networks (ANNs), and k-Nearest Neighbour (k-NN) classifiers. Although these approaches achieved satisfactory performance under controlled conditions, their effectiveness depended heavily on manually designed features and often deteriorated in the presence of significant intra-writer variations and skilled forgeries (Impedovo and Pirlo, 2008; Hafemann et al., 2017).

The emergence of deep learning has significantly advanced offline signature verification by enabling automatic extraction of hierarchical feature representations. Convolutional Neural Networks (CNNs) have demonstrated superior performance over handcrafted methods by learning discriminative visual patterns directly from signature images. Hafemann et al. (2017) showed that deep CNN-based representations substantially improve writer-independent signature verification. Consequently, transfer learning using pre-trained architectures such as VGG, DenseNet, MobileNet, and ResNet has become a widely adopted strategy for extracting robust deep features from relatively small signature datasets (Ozyurt et al., 2024).

Recent studies have focused on improving deep learning models through enhanced feature extraction and representation learning. Zhang et al. (2024) proposed a feature disentangling variational autoencoder to obtain more discriminative signature representations, improving the separation between genuine signatures and skilled forgeries. Similarly, Xiao and Wu (2025) incorporated a Spatial Transformer Network into a deep learning framework to improve feature alignment and robustness against geometric variations in handwritten signatures.

Researchers have also investigated feature selection and optimization techniques to reduce redundancy in deep feature vectors. Ozyurt et al. (2024) combined transfer learning with feature selection methods to identify informative deep features, leading to improved classification performance while reducing computational complexity. In addition, recent surveys have emphasized that current research is increasingly focused on deep learning, hybrid architectures, and feature optimization to enhance offline signature verification accuracy (Nori and Murshid, 2025).

Despite these advances, most existing approaches utilize the complete deep feature vector during verification, implicitly assuming that all extracted features contribute equally to writer discrimination. In practice, however, individual deep features exhibit different levels of stability across genuine signatures. Some features consistently capture writer-specific characteristics, whereas others are

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

affected by natural writing variations, acquisition noise, or redundant information. The reliability of individual deep features has received comparatively little attention in the existing literature.

Motivated by this observation, the present work proposes a **Deep Feature Reliability-Based Framework** for offline handwritten signature verification. Unlike conventional approaches that directly employ all extracted deep features, the proposed framework estimates the statistical reliability of each feature extracted by ResNet50 and selects only the most stable and discriminative features for verification. This reliability-guided strategy aims to reduce the influence of unstable features while improving verification accuracy and robustness against skilled forgeries.

PROPOSED METHODOLOGY

This section presents the proposed deep feature reliability-based framework for offline handwritten signature verification. The major stages of the proposed model, including signature image preprocessing, deep feature extraction, deep feature reliability estimation, reliable feature selection, signature verification, and performance evaluation, are described in detail. The overall workflow of the proposed Deep Feature Reliability-Based Framework for Offline Handwritten Signature Verification is illustrated in Figure 1. The framework consists of separate training and testing phases, where reliable deep features are estimated during training and subsequently utilized to perform robust signature verification.

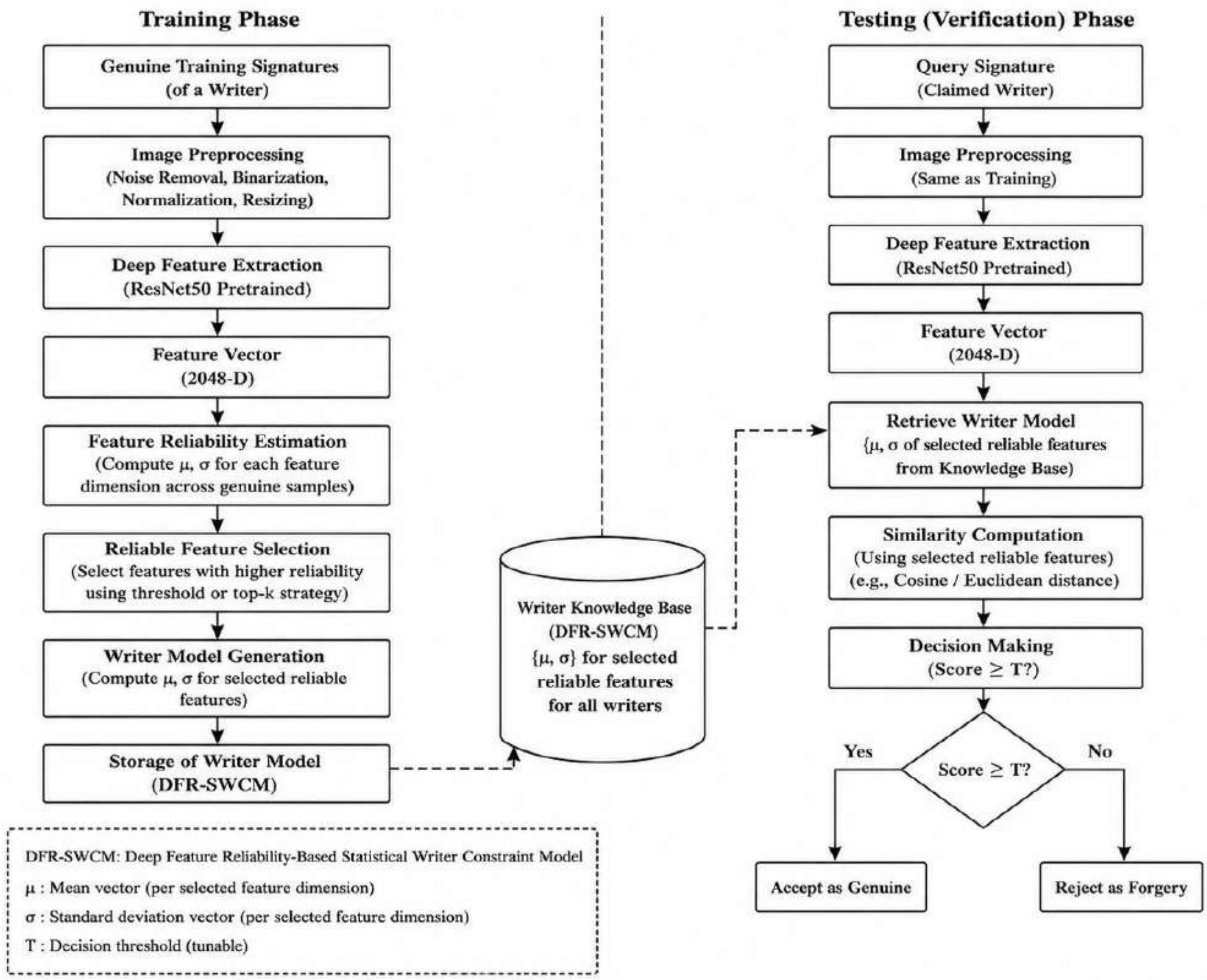


Figure 1. Block diagram of the proposed work.

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

- **Image Preprocessing**

Image preprocessing is performed to improve the quality and consistency of signature images before deep feature extraction. The objective is to remove noise, eliminate irrelevant background information, and standardize the input images while preserving the structural characteristics of the handwritten signature (Hafemann et al., 2017; Hameed et al., 2021).

Initially, each signature image is converted to grayscale since colour information is not required for offline signature verification. A 3×3 median filter is then applied to suppress impulsive noise while preserving the signature strokes (Gonzalez and Woods, 2018). The filtered image is binarized using Otsu's thresholding method, which automatically separates the foreground signature from the background (Otsu, 1979).

Subsequently, connected component analysis is employed to remove small isolated noise components. The signature region is localized using the foreground bounding box and cropped with a small margin to eliminate unnecessary background while retaining the complete signature. The cropped image is then resized without altering its aspect ratio and centred on a 224×224 white canvas to satisfy the input requirements of the pre-trained ResNet50 model (He et al., 2016). Finally, pixel intensities are normalized to the range [0,1], producing a standardized representation suitable for deep feature extraction.

This preprocessing stage generates clean, uniformly sized signature images that facilitate reliable feature extraction and improve the robustness of the subsequent verification process.

- **Deep Feature Extraction**

Deep feature extraction is performed to obtain a robust and discriminative representation of handwritten signatures for subsequent verification. Unlike handcrafted descriptors, deep convolutional neural networks (CNNs) automatically learn hierarchical feature representations that effectively capture both local stroke characteristics and global structural patterns, resulting in improved verification performance (Hafemann et al., 2017; Hameed et al., 2021).

In the proposed framework, a pre-trained ResNet50 model is employed as the feature extraction backbone. ResNet50 is selected because of its strong representational capability, efficient residual learning mechanism, and proven effectiveness in various image recognition and biometric applications (He et al., 2016). The network comprises an initial convolutional layer followed by four residual stages interconnected through identity shortcut connections, which facilitate effective feature propagation and mitigate the degradation problem encountered in deep neural networks.

The preprocessed signature images of size 224×224 are converted into three-channel RGB images and normalized using the ImageNet mean and standard deviation before being provided as input to the network. Instead of utilizing the final fully connected classification layer, the proposed framework removes the classification head and employs the output of the Global Average Pooling (GAP) layer as the signature representation. Consequently, each signature image is represented by a 2048-dimensional deep feature vector, which preserves high-level semantic information while reducing spatial redundancy (He et al., 2016).

Let I_i denote the i^{th} preprocessed signature image and $F(\cdot)$ represent the ResNet50 feature extraction function. The extracted deep feature vector is expressed as

$$\mathbf{x}_i = F(I_i) \qquad (1)$$

Where

$$\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{i2048}] \in \mathbb{R}^{2048} \qquad (2)$$

The extracted deep feature vectors constitute the input to the subsequent Feature Reliability Estimation stage, where the statistical consistency of individual features is analysed to identify the most reliable writer-specific representations.

- **Feature Reliability Estimation**

Although the deep feature vector extracted by ResNet50 provides a rich representation of the handwritten signature, not every feature contributes equally to writer discrimination. Certain features remain highly consistent across genuine signatures of the same writer, whereas others exhibit significant variation due to natural handwriting inconsistencies or image acquisition noise. To address

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

this issue, the proposed framework estimates the reliability of each deep feature based on its statistical stability across the genuine signatures of an individual writer (Everitt and Skrondal, 2010; Montgomery and Runger, 2018).

Let $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{i2048}]$ denote the deep feature vector extracted from the i^{th}genuine signature of a writer. For each feature dimension j, the mean feature value is computed as

$$\mu_j = \frac{1}{N}\sum_{i=1}^{N} x_{ij} \qquad (3)$$

where N represents the number of genuine training signatures for the corresponding writer.
The variability of each feature is then quantified using the sample standard deviation,

$$\sigma_j = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_{ij} - \mu_j\right)^2} \qquad (4)$$

which measures the dispersion of feature values around their mean.
To normalize the variation relative to the feature magnitude, the coefficient of variation (CV) is calculated as

$$CV_j = \frac{\sigma_j}{|\mu_j| + \varepsilon} \qquad (5)$$

where $\varepsilon$is a small positive constant introduced to prevent division by zero.
Finally, the reliability score of the j^{th}feature is computed as

$$R_j = \frac{1}{1 + CV_j} \qquad (6)$$

where $R_j \in [0,1]$. A higher reliability score indicates that the corresponding feature exhibits greater consistency across the genuine signatures of the writer and is therefore more representative of the writer's unique signing characteristics. Conversely, features with lower reliability scores are considered less stable and are more likely to be influenced by natural writing variations or noise.
The reliability scores obtained for all 2048 feature dimensions form a feature reliability vector, which serves as the basis for selecting the most stable and discriminative features in the subsequent Reliable Feature Selection stage. This reliability-driven strategy enables the proposed framework to emphasize writer-consistent deep features while reducing the influence of unstable or redundant information, thereby improving the robustness of offline handwritten signature verification.

- **Reliable Feature Selection**

The feature reliability scores estimated in the previous stage are utilized to identify the most stable and discriminative deep features for signature verification. Although ResNet50 generates a 2048-dimensional feature vector, not every feature contributes equally to writer discrimination. Features exhibiting higher reliability scores remain more consistent across genuine signatures of the same writer and therefore provide greater discriminative information. In contrast, features with lower reliability scores are more susceptible to natural handwriting variations, image noise, and redundant information, which may negatively influence the verification process (Guyon and Elisseeff, 2003; Chandrashekar and Sahin, 2014).
Let

$$\mathbf{R} = [R_1, R_2, \dots, R_{2048}]$$

represent the feature reliability vector obtained from the Feature Reliability Estimation stage. The reliability scores are arranged in descending order such that

$$R_{(1)} \geq R_{(2)} \geq \cdots \geq R_{(2048)} \qquad (7)$$

where $R_{(k)}$ denotes the k^{th} highest reliability score. Ranking the features according to their reliability enables the identification of the most stable writer-specific characteristics while eliminating less informative feature dimensions.
Based on the ranked reliability scores, the top $m$reliable features are selected to construct a reduced feature representation,

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

$$\mathbf{x}' = [x_{j_1}, x_{j_2}, \dots, x_{j_m}] \quad (8)$$

where $j_1, j_2, \dots, j_m$ represent the indices of the selected feature dimensions and $m < 2048$. In the proposed implementation, the **512** highest-ranked features are retained for subsequent verification, while the remaining features are discarded. This dimensionality reduction decreases computational complexity without sacrificing the most informative writer-specific characteristics (Guyon and Elisseeff, 2003).

The reliability values corresponding to the selected features are preserved as feature weights and are subsequently incorporated into the weighted similarity computation during the signature verification stage. By emphasizing highly reliable feature dimensions and suppressing unstable representations, the proposed framework enhances the discriminative capability of the deep feature representation and improves the robustness of offline handwritten signature verification (Chandrashekar and Sahin, 2014; Hameed et al., 2021).

- **Signature Verification**

During the verification stage, the query signature is compared with the reference model of the claimed writer using the reliable deep features selected in the previous stage. Unlike conventional methods that treat all feature dimensions equally, the proposed framework incorporates the estimated feature reliability scores as feature weights during similarity computation. As a result, features that remain highly consistent across genuine signatures contribute more significantly to the verification process, whereas unstable features have a reduced influence on the final decision (Hameed et al., 2021; Guyon and Elisseeff, 2003).

For each claimed writer, a reference feature vector is constructed from the selected reliable features of the genuine training signatures. The query signature is processed using the same preprocessing and feature extraction procedures, and its reduced feature vector is obtained using the selected feature indices. The similarity between the query feature vector q and the reference feature vector p is then computed using a weighted cosine similarity measure. Cosine similarity is widely adopted in biometric verification because it effectively measures the angular similarity between high-dimensional feature vectors while being less sensitive to differences in feature magnitude (Aggarwal, 2015).

The weighted cosine similarity is computed as

$$S(q, p) = \frac{\sum_{j=1}^{m} w_j\, q_j p_j}{\sqrt{\sum_{j=1}^{m} w_j\, q_j^2}\sqrt{\sum_{j=1}^{m} w_j\, p_j^2}} \quad (9)$$

where q_j and p_j denote the j^{th} selected feature of the query and reference signatures, respectively, w_j represents the reliability weight associated with the j^{th} feature, and m is the number of selected reliable features. In the proposed framework, $m = 512$.

The computed similarity score lies within the interval $[-1,1]$, where higher values indicate a greater degree of similarity between the query signature and the claimed writer's reference model. By integrating feature reliability into the similarity computation, the proposed framework emphasizes stable writer-specific characteristics while suppressing the influence of unreliable features, thereby improving the robustness and discriminative capability of the offline handwritten signature verification process (Guyon and Elisseeff, 2003; Aggarwal, 2015).

EXPERIMENTAION AND RESULTS

The proposed Deep Feature Reliability-Based Framework for Offline Handwritten Signature Verification was evaluated using the CEDAR offline handwritten signature dataset, which is one of the most widely used benchmark datasets for signature verification research (Kalera et al., 2004). The dataset comprises signatures collected from 55 writers, where each writer has 24 genuine signatures and 24 skilled forgeries, resulting in a total of 2,640 signature images. The availability of both genuine and skilled forgery samples makes the CEDAR dataset suitable for assessing the robustness of offline signature verification systems.

All signature images were first preprocessed using grayscale conversion, median filtering, Otsu's thresholding, noise removal, signature cropping, resizing, and normalization. The preprocessed images were then resized to 224×224pixels and provided as input to a pre-trained ResNet50 network for extracting 2048-dimensional deep feature vectors. Subsequently, the proposed feature reliability estimation algorithm was applied to compute the reliability of each feature dimension, and the 512 most reliable features were selected for signature verification.

To evaluate the generalization capability of the proposed framework, the genuine signatures of each writer were divided into training and testing subsets using three different training–testing ratios: 25–75, 50–50, and 75–25. Only genuine signatures were used during the training phase for writer model construction and feature reliability estimation, whereas the testing phase included the remaining genuine signatures together with the corresponding skilled forgeries. The experiments for each split were repeated 10 times using different random partitions, and the average performance was reported to minimize the influence of random sampling.

- **Decision Making**

The final verification decision is obtained by comparing the weighted cosine similarity score with a predefined decision threshold. The similarity score reflects the degree of correspondence between the query signature and the reference model of the claimed writer. Higher similarity scores indicate that the query signature closely matches the genuine writing characteristics of the claimed writer, whereas lower scores suggest that the signature is more likely to be a skilled forgery.

In the proposed framework, a writer-specific decision threshold is estimated during the training phase using only the genuine training signatures of the corresponding writer. This threshold captures the natural intra-writer variability while preventing the use of testing information during model construction. During the verification stage, the similarity score obtained for a query signature is compared against the estimated threshold to determine its authenticity. Similar threshold-based decision strategies are widely employed in biometric verification systems because of their simplicity and effectiveness in separating genuine and impostor samples (Jain et al., 2004; Aggarwal, 2015).

The final verification decision is defined as

$$D = \begin{cases} \text{Genuine}, & \text{if } S(q,p) \geq T, \\ \text{Forged}, & \text{if } S(q,p) < T. \end{cases} \qquad (10)$$

Where, D is the verification decision, $S(q,p)$denotes the weighted cosine similarity score and Trepresents the writer-specific decision threshold estimated from the genuine training signatures.

A query signature is accepted as **genuine** when its similarity score is greater than or equal to the threshold, indicating sufficient similarity to the claimed writer's reference model. Otherwise, the signature is classified as **forged**. This decision mechanism effectively combines the discriminative capability of deep features with the proposed feature reliability weighting strategy, enabling robust offline handwritten signature verification.

- **Performance Evaluation**

The performance of the proposed Deep Feature Reliability-Based Framework is evaluated using standard biometric verification metrics that quantify the ability of the system to distinguish genuine signatures from skilled forgeries. The experiments are conducted under different training–testing splits, and the average results obtained over multiple independent trials are reported to ensure a reliable assessment of the proposed method. The evaluation is based on the confusion matrix consisting of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), which are widely adopted in offline signature verification research (Jain et al., 2004; Hameed et al., 2021).

The overall verification accuracy is computed as

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (11)$$

which represents the proportion of correctly classified genuine and forged signatures.

Precision measures the proportion of signatures classified as genuine that are actually genuine and is defined as

$$\text{Precision} = \frac{TP}{\frac{TP}{TP+FP}} \qquad (12)$$

whereas Recall represents the proportion of genuine signatures correctly identified by the system,

$$\text{Recall} = \frac{TP}{TP + FN} \qquad (13)$$

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

The harmonic mean of Precision and Recall is calculated using the F1-score,

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (14)$$

which provides a balanced measure of verification performance.

In addition to these classification metrics, two biometric error measures are considered. The False Acceptance Rate (FAR) indicates the proportion of skilled forgeries incorrectly accepted as genuine,

$$\text{FAR} = \frac{FP}{FP + TN} \qquad (15)$$

whereas the **False Rejection Rate (FRR)** represents the proportion of genuine signatures incorrectly rejected,

$$\text{FRR} = \frac{FN}{TP + FN} \qquad (16)$$

Finally, the Equal Error Rate (EER) is reported as an overall measure of verification performance. The EER corresponds to the operating point at which the False Acceptance Rate and False Rejection Rate are equal. A lower EER indicates better verification capability and improved discrimination between genuine and forged signatures (Jain et al., 2004).

These evaluation metrics collectively provide a comprehensive assessment of the effectiveness, reliability, and robustness of the proposed framework under different training–testing configurations.

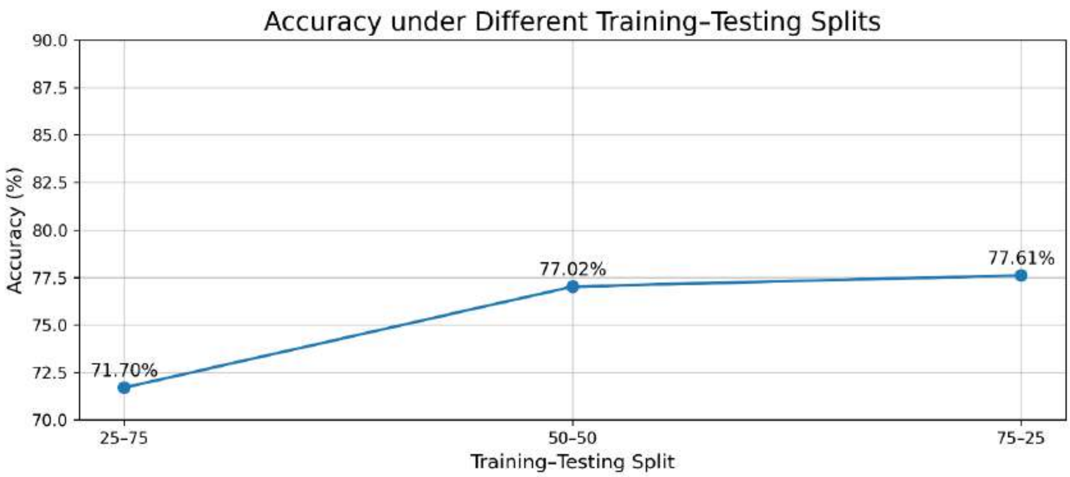
RESUTLS AND DISCUSSION

The proposed framework was evaluated under three different training–testing splits to investigate its verification performance under varying amounts of training data. **Table 1** summarizes the average Accuracy, Precision, Recall, and F1-score obtained for each experimental configuration.

Table 1. Performance Comparison under Different Training–Testing Splits

Training–Testing Split	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
25–75	71.70	83.93	53.88	65.45
50–50	77.02	84.27	66.48	74.29
75–25	77.61	83.77	68.48	75.33

To further illustrate the influence of different training–testing splits on the proposed framework, the verification accuracy obtained for each experimental setting is presented graphically in **Figure 2**.



International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 09 Issue 07 July 2026
DOI: 10.47191/ijcsrr/V9-i7-28, Impact Factor: 8.048
IJCSRR @ 2026
www.ijcsrr.org

Figure 2. Accuracy of the proposed Framework under different training–testing splits.

The experimental results demonstrate that the proposed framework achieves consistent signature verification performance under different training–testing splits. As the proportion of training signatures increases, the verification accuracy improves from 71.70% for the 25–75 split to 77.61% for the 75–25 split. A similar trend is observed for the Recall and F1-score, indicating that a larger number of genuine training signatures enables more reliable estimation of writer-specific feature reliability. These results confirm that the proposed reliability-based feature selection strategy effectively captures stable deep features, thereby enhancing the robustness of offline handwritten signature verification.

CONCLUSION

This paper presented a Deep Feature Reliability-Based Framework for Offline Handwritten Signature Verification that combines ResNet50-based deep feature extraction with a feature reliability estimation strategy to identify stable and discriminative writer-specific features. The selected reliable features were used for weighted signature verification, improving the robustness of the verification process.

The proposed framework was evaluated on the CEDAR dataset under different training–testing splits. The best performance was achieved with the 75–25 split, yielding an accuracy of 77.61%. Future work will focus on developing more discriminative reliability measures and incorporating advanced deep learning models to further improve verification performance.

REFERENCES

1. Jain, A. K., Ross, A., Prabhakar, S. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology. 2004;14(1):4–20. https://doi.org/10.1109/TCSVT.2003.818349
2. Hameed, S., Er-Raoudi, M., Rachik, M., et al. Machine learning-based offline signature verification systems: A systematic review. Signal Processing: Image Communication. 2021;93:116139. https://doi.org/10.1016/j.image.2021.116139
3. Hafemann, L. G., Sabourin, R., Oliveira, L. S. Offline handwritten signature verification—Literature review. Proceedings of the International Joint Conference on Neural Networks (IJCNN). 2017:1–8. https://doi.org/10.1109/IJCNN.2017.7966214
4. Hafemann, L. G., Sabourin, R., Oliveira, L. S. 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
5. Impedovo, D., Pirlo, G. Automatic signature verification: The state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews. 2008;38(5):609–635. https://doi.org/10.1109/TSMCC.2008.923866
6. Hafemann, L. G., Sabourin, R., Oliveira, L. S. 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
7. Özyurt, F., Majidpour, J., Rashid, T. A., Koç, C. Offline handwriting signature verification: A transfer learning and feature selection approach. Traitement du Signal. 2023;40(6):2613–2622. https://doi.org/10.18280/ts.400623
8. Hameed, S., Er-Raoudi, M., Rachik, M., et al. Machine learning-based offline signature verification systems: A systematic review. Signal Processing: Image Communication. 2021;93:116139. https://doi.org/10.1016/j.image.2021.116139
9. Xiao, W., Wu, H. Learning features for offline handwritten signature verification using spatial transformer network. Scientific Reports. 2025;15:9453. https://doi.org/10.1038/s41598-025-92704-3
10. Diaz, M., Ferrer, M. A., Vessio, G. Explainable offline automatic signature verifier to support forensic handwriting examiners.
11. Tsourounis, D., Theodorakopoulos, I., Zois, E. N., Economou, G. Leveraging expert models for training deep neural networks in scarce data domains: Application to offline handwritten signature verification.
12. He, K., Zhang, X., Ren, S., Sun, J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016:770–778. https://doi.org/10.1109/CVPR.2016.90
13. Otsu, N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics. 1979;9(1):62–66. https://doi.org/10.1109/TSMC.1979.4310076
14. Gonzalez, R. C., Woods, R. E. Digital Image Processing. 4th ed. Pearson; 2018.

15. He, K., Zhang, X., Ren, S., Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016:770–778. https://doi.org/10.1109/CVPR.2016.90
16. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L. ImageNet: A large-scale hierarchical image database. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2009:248–255. https://doi.org/10.1109/CVPR.2009.5206848
17. Everitt, B. S., Skrondal, A. The Cambridge Dictionary of Statistics. 4th ed. Cambridge University Press; 2010.
18. Montgomery, D. C., Runger, G. C. Applied Statistics and Probability for Engineers. 7th ed. John Wiley & Sons; 2018.
19. Guyon, I., Elisseeff, A. An introduction to variable and feature selection. Journal of Machine Learning Research. 2003;3:1157–1182.
20. Chandrashekar, G., Sahin, F. A survey on feature selection methods. Computers & Electrical Engineering. 2014;40(1):16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024
21. Kalera, M. K., Srihari, S. N., Xu, A. Offline signature verification and identification using distance statistics. International Journal of Pattern Recognition and Artificial Intelligence. 2004;18(7):1339–1360.
22. Aggarwal, C. C. *Data Mining: The Textbook*. Springer; 2015. https://doi.org/10.1007/978-3-319-14142-8
23. Jain, A. K., Ross, A., Prabhakar, S. An introduction to biometric recognition. IEEE Transactions on Circuits and Systems for Video Technology. 2004;14(1):4–20. https://doi.org/10.1109/TCSVT.2003.818349
24. Sokolova, M., Lapalme, G. A systematic analysis of performance measures for classification tasks. *Information Processing & Management*. 2009;45(4):427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Cite this Article: Annapurna H. (2026). A Deep Feature Reliability-Based Framework for Offline Handwritten Signature Verification. International Journal of Current Science Research and Review, 9(7), pp. 3880-3889. DOI: https://doi.org/10.47191/ijcsrr/V9-i7-28