Deep Neural Networks in Offline Handwritten Signature Verification: A Systematic Review (2016–2025)

Dr. Annapurna H

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

***Abstract*—Offline handwritten signature verification remains a challenging biometric problem due to high intra-writer variability, low inter-writer variability, and the absence of dynamic writing information. Over the past decade, deep neural networks have significantly advanced this field by enabling the automatic learning of features and the robust representation of complex signature patterns. This paper presents a comprehensive systematic review of deep learning-based approaches for offline handwritten signature verification published between 2016 and 2025.**

The review analyzes a wide range of deep architectures, including convolutional neural networks, Siamese and metric-learning frameworks, recurrent models, attention-based networks, and transformer-based architectures. Emphasis is placed on how these models address key challenges such as limited training data, skilled forgery detection, cross-writer generalization, and variability across writing styles and scripts. Commonly used benchmark datasets, including GPDS, MCYT, CEDAR, BHSig260, and other multilingual signature corpora, are examined in terms of their role in performance evaluation and comparability across studies.

The surveyed works are systematically categorized according to architectural design, learning strategy, and evaluation protocol, highlighting evolving research trends and methodological shifts over time. Key findings indicate a transition from conventional convolutional pipelines toward hybrid and attention-based models that capture both local stroke-level details and global structural dependencies. Despite significant progress, challenges such as dataset imbalance, limited generalization across writing styles, and lack of standardized evaluation protocols remain open research problems.

This review presents a comprehensive overview of the current state of the art in offline handwritten signature verification, identifying promising research directions, including self-supervised learning, foundation models, and cross-domain adaptation, to support the development of more robust and scalable verification systems.

***Index Terms*—Offline handwritten signature verification, Deep neural networks, Convolutional neural networks, Siamese networks, Transformer models, Deep metric learning, Biometric authentication, Signature datasets, Writer-independent verification, Pattern recognition.**

I. INTRODUCTION

Offline handwritten signature verification aims to determine whether a static signature image belongs to a claimed individual. Unlike online verification, which exploits dynamic information such as pen pressure and writing speed, offline systems rely solely on visual cues extracted from scanned images. This lack of temporal information significantly increases the difficulty of the task, particularly in the presence of high intra-writer variability and skilled forgeries that closely resemble genuine signatures [1,2].

From a system design perspective, offline signature verification is typically formulated as a binary classification problem. A queried signature is classified as either genuine or forged based on its similarity to reference samples. Existing systems are commonly categorized as writer-dependent (WD) or writer-independent (WI). While WD systems often achieve higher accuracy due to personalized modelling, they suffer from limited scalability. In contrast, WI systems aim to learn generalizable decision boundaries applicable across users, making them more suitable for large-scale deployments but also more challenging to train effectively [3].

Early research in offline signature verification focused on handcrafted feature extraction techniques, including geometric descriptors, stroke-based measurements, and texture features such as Gabor filters and local binary patterns. These features were typically combined with classical classifiers such as k-nearest neighbors, support vector machines, or hidden Markov models. Although these approaches yielded reasonable performance, their dependence on manual feature engineering and sensitivity to writing style variations limited their robustness and cross-dataset generalization capabilities [4,5].

The introduction of deep learning marked a major paradigm shift in the field. Convolutional neural networks (CNNs) enabled automatic hierarchical feature learning directly from raw signature images, eliminating the need for handcrafted descriptors. Early CNN-based methods demonstrated strong improvements in discrimination capability by capturing both local stroke characteristics and global structural patterns [6–8]. Subsequent studies extended these models through deeper architectures, transfer learning, and data augmentation to address limited training data and improve generalization.

To better model the verification nature of the problem, metric learning approaches particularly Siamese and triplet networks were introduced. These models learn embedding spaces in which genuine signature pairs are drawn closer while forgeries are pushed apart, typically using contrastive or triplet loss functions. Such architectures have shown strong performance in writer-independent scenarios and have become a dominant paradigm in recent research [9–11].

More recent developments have explored hybrid and attention-based architectures that combine convolutional feature extractors with attention mechanisms or transformer encoders. These models are capable of capturing long-range dependencies and global contextual relationships within signatures, addressing limitations of fixed receptive fields in conventional CNNs. Additionally, representation learning strategies such as sparse coding and dictionary learning have been revisited to enhance discrimination under limited-data conditions. Hybrid frameworks that integrate deep feature learning with classical classifiers or metric learning have further improved robustness and interpretability [12].

Overall, the progression from handcrafted features to deep and hybrid learning paradigms reflects a clear trend toward more expressive and adaptive representations. Despite substantial advances, challenges such as data scarcity, cross-script generalization, and computational efficiency remain open research problems.

The remainder of this paper is organized as follows: Section 2 presents the transition to the proposed methodology, Section 3 provides a comparative analysis of deep learning–based approaches for offline handwritten signature verification, Section 4 discusses the key findings and implications of the reviewed work, and Section 5 concludes the paper.

II. TRANSITION TO METHODOLOGY

Building upon the insights obtained from prior research, the following section presents a comprehensive comparative analysis of deep learning–based approaches for offline handwritten signature verification. This analysis is designed to systematically examine the evolution of methodological paradigms, highlighting both the strengths and limitations of existing solutions while identifying emerging research trends. In particular, the discussion encompasses a wide range of architectural families, including convolutional neural networks, metric-learning and Siamese frameworks, as well as more recent attention-based and transformer-driven models [1–4]. By examining these approaches in a unified framework, the section aims to elucidate how different design choices influence feature representation, discrimination capability, and overall verification performance.

Furthermore, the proposed methodological perspective draws upon recent advances in deep representation learning to address persistent challenges reported in the literature. These challenges include limited availability of labelled training data, high intra-writer variability, low inter-writer separability, and reduced generalization across writers, scripts, and acquisition conditions. Special attention is given to how contemporary models attempt to improve feature robustness, enhance generalization across heterogeneous datasets, and maintain verification reliability under constrained and imbalanced training scenarios [5–12]. Through this comparative analysis, the section establishes a conceptual foundation for understanding current limitations and motivates the need for more robust, scalable, and adaptable frameworks for offline handwritten signature verification.

III. COMPARATIVE ANALYSIS OF DEEP LEARNING – BASED OFFLINE SIGNATURE VERIFICATION METHODS

This section presents an in-depth and structured discussion of major deep learning–based approaches for offline handwritten signature verification. Each work is examined in terms of its methodological contributions, datasets used, performance characteristics, and inherent strengths and limitations. The analysis reflects the progressive evolution of the field from early convolutional architectures to recent transformer-based and hybrid frameworks.

A. Early CNN-Based Approaches

The foundational work by Hafemann et al. [1] marked one of the earliest successful applications of convolutional neural networks (CNNs) to offline signature verification. Their writer-independent framework demonstrated that CNNs could effectively learn discriminative features directly from raw signature images, outperforming traditional handcrafted feature-based methods. Evaluated on benchmark datasets such as GPDS, MCYT, and CEDAR, the approach showed strong generalization capability across writers. However, the model required substantial labeled data and exhibited sensitivity to skilled forgeries, particularly in scenarios with limited training samples.

Extending this work, Hafemann et al. [13] proposed a fixed-size representation strategy to address variations in signature dimensions. This improvement enhanced robustness across datasets and enabled more consistent learning. Despite these advantages, the approach still relied heavily on convolutional feature hierarchies and lacked explicit mechanisms for modelling long-range dependencies or contextual relationships between signature components.

B. Metric Learning and Pairwise Architectures

To better capture the intrinsic nature of signature verification, Xing et al. [14] introduced a Siamese convolutional network that learned similarity metrics between signature pairs. This formulation allowed the model to directly optimize verification objectives and improved discrimination between genuine and forged signatures. However, its performance was sensitive to pair selection strategies and required careful balancing of positive and negative samples during training.

Similarly, Soleimani et al. [9] proposed a multitask metric learning framework that jointly optimized classification and similarity objectives. This approach demonstrated improved generalization across multiple datasets and enhanced robustness against intra-writer variability. Nonetheless, the increased architectural complexity and training cost limited scalability, particularly for large-scale deployments.

C. Ensemble and Multi-Representation Learning

To address robustness and stability, Masoudnia et al. [8] introduced a multi-loss, multi-representation CNN architecture that combined multiple feature extractors. Evaluated on datasets such as MCYT and UT-SIG, this approach achieved improved consistency and reduced sensitivity to noise. However, the increased computational burden and training time posed challenges for real-time applications.

Similarly, Maergner et al. [15] integrated graph edit distance with deep embeddings, effectively combining structural and learned representations. While this hybrid strategy improved discrimination in complex cases, it introduced additional computational overhead and complexity in feature alignment.

D. Transfer Learning and Pretrained Architectures

The adoption of transfer learning significantly advanced offline signature verification. Jahandad et al. [8] demonstrated that fine-tuning pretrained CNNs such as Inception-v1 and Inception-v3 led to notable performance gains on datasets like GPDS, especially in data-scarce scenarios. This approach reduced training time and improved convergence but remained dependent on the relevance of the source domain used for pretraining.

Similarly, pretrained deep CNNs explored by Bhavani et al. [22] showed strong generalization across datasets, highlighting the effectiveness of knowledge transfer. However, these methods often required careful fine-tuning to avoid overfitting and domain mismatch.

E. Advanced Metric Learning and Data Augmentation

Wei et al. [10] introduced inverse discriminative networks that enhanced class separability in the embedding space. Their approach achieved strong performance in writer-independent settings but required meticulous calibration of loss functions. Ruiz et al. [11] further improved robustness by integrating synthetic signature generation into Siamese networks, effectively addressing class imbalance and limited data availability. Despite these advantages, synthetic data generation introduced the risk of domain shift if generated samples did not accurately reflect real writing variability.

E. Region-Based and Hybrid Architectures

Avola et al. [13] proposed a reduced-space learning framework that balanced discriminative performance with computational efficiency. Meanwhile, Liu et al. [14] introduced a region-based deep metric learning approach that selectively focused on informative regions of signatures, improving resistance to skilled forgeries. These methods demonstrated improved interpretability but often required additional preprocessing steps for region extraction.

E. Sequential and Feature Aggregation Models

Ghosh [18] employed recurrent neural networks to capture sequential dependencies in offline signatures, achieving improved results across multiple datasets. However, recurrent models increased computational complexity and training time. Tuncer et al. [19] introduced a deep feature warehouse combined with iterative feature selection, achieving strong performance through structured feature aggregation but at the cost of increased system complexity.

F. Transformer-Based and Attention-Driven Models

Recent advancements have focused on attention mechanisms and transformer architectures. Li et al. [12] proposed a transformer-based framework that effectively captured long-range dependencies and outperformed traditional CNNs on several benchmarks. Ji et al. [20] further refined this paradigm through paired contrastive learning, enhancing robustness in challenging verification scenarios. Xiong et al. [20] employed attention-guided Siamese networks to improve feature alignment, while Bhavani et al. [22] demonstrated that pretrained deep CNNs could generalize well across datasets when combined with appropriate fine-tuning strategies.

The most recent works, including those by Xiao and Wu [23] and Li et al. [24], introduced hybrid CNN–transformer architectures that integrate local texture modelling with global contextual awareness. These approaches represent the current state of the art, achieving strong performance across multilingual and multi-script datasets. However, their increased computational demands and architectural complexity highlight the need for efficient deployment strategies.

Overall, the evolution of offline handwritten signature verification reflects a clear progression from handcrafted features to deep, hybrid, and attention-based models. While modern approaches achieve impressive accuracy, challenges such as dataset imbalance, cross-script generalization, computational cost, and reproducibility remain open. Addressing these limitations requires unified frameworks that balance accuracy, efficiency, and generalizability, an objective that motivates the methodology proposed in the subsequent section.

IV. DISCUSSION

The evolution of offline handwritten signature verification reflects a clear progression from handcrafted feature engineering toward data-driven representation learning. Early approaches relied heavily on manually designed descriptors, which, although effective in controlled settings, struggled to generalize across different writing styles and acquisition conditions. The introduction of deep learning, particularly convolutional neural networks (CNNs), marked a turning point by enabling automatic extraction of hierarchical visual features directly from raw signature images. These CNN-based approaches demonstrated substantial improvements in capturing both local stroke-level details and global structural patterns, establishing a strong foundation for subsequent research.

Building upon this foundation, metric learning and Siamese network architectures further enhanced verification performance by explicitly modelling similarity relationships between signature pairs. By learning discriminative embedding spaces, these models improved writer-independent verification and demonstrated greater robustness against skilled forgeries. Their success highlighted the importance of learning relational information rather than relying solely on classification-based decision boundaries. However, such models often require carefully curated training pairs and can be sensitive to class imbalance, which remains a practical challenge in real-world applications.

More recently, the field has witnessed a shift toward attention-based and transformer-driven architectures.

These models offer the ability to capture long-range dependencies and global contextual information that conventional CNNs struggle to represent. Transformer-based approaches, in particular, have shown strong potential in modelling complex structural relationships within signatures and adapting to variations in writing style. Hybrid architectures that combine convolutional feature extractors with attention mechanisms or transformer blocks have further enhanced representational capacity, leading to improved verification accuracy across diverse datasets.

Despite these advances, several critical challenges persist. One major limitation is the scarcity and imbalance of publicly available datasets, which restricts the generalization capability of deep models and complicates fair performance comparison across studies. Many datasets are limited in size, language diversity, or forgery complexity, which can lead to overfitting and reduced robustness when models are deployed in real-world scenarios. Additionally, the lack of standardized evaluation protocols including variations in train–test splits, writer-dependent versus writer-independent settings, and performance metrics makes it difficult to directly compare reported results across different studies.

Computational complexity also remains a concern, particularly for transformer-based and ensemble models that demand substantial memory and processing resources. Such requirements may hinder deployment in resource-constrained environments, including embedded or mobile authentication systems. Furthermore, the interpretability of deep models remains limited, raising concerns in applications where transparency and explainability are essential.

Overall, while deep learning has significantly advanced the state of offline handwritten signature verification, achieving robust, scalable, and interpretable systems remains an open research challenge. Future research directions should focus on developing lightweight yet expressive architectures, improving cross-domain and cross-script generalization, and establishing standardized benchmarking protocols. Additionally, emerging paradigms such as self-supervised learning, foundation models, and domain adaptation hold promise for reducing data dependency and enhancing robustness, paving the way for more reliable and practical signature verification systems.

V. CONCLUSION

This review has systematically examined the evolution of deep learning-based methodologies for offline handwritten signature verification over the period 2016–2025, highlighting key architectural trends, methodological advancements, and persistent challenges. The analysis demonstrates a clear progression from conventional convolutional neural network (CNN) architectures toward more expressive paradigms incorporating metric learning, attention mechanisms, and transformer-based representations. These developments have significantly enhanced the capacity of verification systems to model complex intra-writer variations and subtle inter-writer distinctions, which are central challenges in offline signature verification.

Early CNN-based approaches established strong baselines by learning hierarchical spatial representations directly from raw signature images, thereby reducing dependence on handcrafted features. Subsequent advances in Siamese and metric learning frameworks introduced embedding-based representations that explicitly optimized inter-sample similarity, improving generalization in writer-independent scenarios. More recent architectures, particularly those incorporating attention mechanisms and transformer modules, have further extended modelling capacity by capturing long-range dependencies and global structural relationships that are difficult to represent using purely convolutional operations.

Despite these advances, several fundamental challenges remain unresolved. The limited availability of large-scale, diverse, and well-annotated datasets continues to constrain model generalization and reproducibility. Furthermore, variations in evaluation protocols, dataset partitioning strategies, and performance metrics hinder fair comparison across studies. The increasing computational complexity of modern deep architectures especially transformer-based models also raises concerns regarding scalability, energy efficiency, and real-world deployment in resource-constrained environments. Additionally, many existing approaches exhibit sensitivity to domain shifts caused by variations in writing instruments, acquisition conditions, and cultural or script-specific characteristics.

Looking forward, future research should prioritize the development of data-efficient learning strategies, including self-supervised, semi-supervised, and few-shot learning paradigms, to reduce dependence on large labelled datasets. Cross-domain and cross-script adaptation techniques will be essential for achieving robust generalization across heterogeneous populations. Moreover, the integration of foundation models and multimodal representations offers promising directions for enhancing representation learning and transferability. Establishing standardized benchmarks, unified evaluation protocols, and reproducible experimental settings will be critical for enabling meaningful comparison and accelerating progress in the field. Collectively, these efforts are expected to drive the next generation of robust, scalable, and trustworthy offline handwritten signature verification systems.

REFERENCES

[1] Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2016). Writer-independent feature learning for offline signature verification using deep convolutional neural networks. Proceedings of the International Joint Conference on Neural Networks (IJCNN), 2576–2583. https://doi.org/10.1109/IJCNN.2016.7727521

[2] 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

[3] Impedovo, D., Pirlo, G., Ferrer, M. A., Diaz, M., Malik, M. I., & Plamondon, R. (2019). A perspective analysis of handwritten signature technology. ACM Computing Surveys, 51(6), Article 117. https://doi.org/10.1145/3274658

[4] Malphedwar, L., Waghere, M. C., Nimodiya, P. V., & Mashalkar, N. B. (2017). Survey on offline handwritten signature verification. International Journal of Engineering Research & Technology (IJERT), 6(1), 1–6. https://www.ijert.org/research/survey-on-offline-handwritten-signature-verification-IJERTV6IS010032.pdf

[5] Zois, E. N., Tsourounis, D., Theodorakopoulos, I., Kesidis, A. L., & Economou, G. (2019). A comprehensive study of sparse representation techniques for offline signature verification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(11), 2564–2578. https://doi.org/10.1109/TPAMI.2018.2866465

[6] Jahandad, S. M., Sam, S. M., Kamardin, K., & Sjarif, N. N. A. (2019). Offline signature verification using deep learning convolutional neural network architectures GoogLeNet Inception-v1 and Inception-v3. Procedia Computer Science, 161, 475–483. https://doi.org/10.1016/j.procs.2019.11.147

[7] Bonde, S. V., Narwade, P., & Sawant, R. (2020). Offline signature verification using convolutional neural network. Proceedings of the 6th International Conference on Signal Processing and Communication (ICSC), 119–127. https://doi.org/10.1109/ICSC48311.2020.9182727

[8] Masoudnia, S., Mersa, O., Araabi, B. N., Vahabie, A.-H., Sadeghi, M. A., & Ahmadabadi, M. N. (2019). Multi-representational learning for offline signature verification using multi-loss snapshot ensemble of CNNs. Expert Systems with Applications, 133, 317–330. https://doi.org/10.1016/j.eswa.2019.03.040

[9] Soleimani, A., Araabi, B. N., & Fouladi, K. (2016). Deep multitask metric learning for offline signature verification. Pattern Recognition Letters, 80, 84–90. https://doi.org/10.1016/j.patrec.2016.05.023

[10] Wei, P., Li, H., & Hu, P. (2019). Inverse discriminative networks for handwritten signature verification. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 5738–5747. https://doi.org/10.1109/CVPR.2019.00591

[11] Ruiz, V., Linares, I., Sánchez, A., & Vélez, J. F. (2020). Off-line handwritten signature verification using compositional synthetic generation of signatures and Siamese neural networks. Neurocomputing, 374, 30–41. https://doi.org/10.1016/j.neucom.2019.09.041

[12] Li, H., Wei, P., Ma, Z., Li, C., & Zheng, N. (2024). TransOSV: Offline signature verification with transformers. Pattern

Recognition, 145, 109882. https://doi.org/10.1016/j.patcog.2023.109882

[13] Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2018). Fixed-sized representation learning from offline handwritten signatures of different sizes.IJDAR, 21, 219–230. https://doi.org/10.1007/s10032-018-0301-6

[14] Xing, Z. J., et al. (2018). Offline signature verification using convolution Siamese network.SPIE Proceedings. https://doi.org/10.1117/12.2303380

[15] Maergner, P., et al. (2019). Combining graph edit distance and triplet networks for offline signature verification. Pattern Recognition Letters, 125, 527–533. https://doi.org/10.1016/j.patrec.2019.06.024

[16] Avola, D., et al. (2021). R-SigNet: Reduced space writer-independent feature learning for offline writer-dependent signature verification. Pattern Recognition Letters, 150, 189–196. https://doi.org/10.1016/j.patrec.2021.06.033

[17] Liu, L., Huang, L., Yin, F., & Chen, Y. (2021). Offline signature verification using a region-based deep metric learning network. Pattern Recognition, 118, 108009. https://doi.org/10.1016/j.patcog.2021.108009

[18] Ghosh, R. (2021). A recurrent neural network based deep learning model for offline signature verification and recognition system. Expert Systems with Applications, 168, 114249. https://doi.org/10.1016/j.eswa.2020.114249

[19] Tuncer, T., et al. (2022). A deep feature warehouse and iterative MRMR based handwritten signature verification method. Multimedia Tools and Applications, 81, 3899–3913. https://doi.org/10.1007/s11042-021-11726-x

[20] Ji, X., Suehiro, D., & Uchida, S. (2023). Paired contrastive feature for highly reliable offline signature verification.Pattern Recognition, 144, 109816. https://doi.org/10.1016/j.patcog.2023.109816

[21] Xiong, Y., Cheng, S., Ren, J., & Zhang, Y. (2024). Attention-based multiple Siamese networks with primary representation guiding for offline signature verification. IJDAR, 27(2), 195–208. https://doi.org/10.1007/s10032-023-00455-6

[22] Bhavani, S. D., Bharathi, R. K., & Anil Kumar, R. J. (2024). Offline signature verification using pre-trained deep convolutional neural network. AIP Conference Proceedings, 2966, 020007. https://doi.org/10.1063/5.0189862

[23] Xiao, W., & Wu, H. (2025). Learning features for offline handwritten signature verification using spatial transformer network. Scientific Reports, 15, 9453. https://doi.org/10.1038/s41598-025-92704-3

[24] Li, W., Muhammat, M., Xu, X., Aysa, A., & Ubul, K. (2025). Multi-scale CNN-CrossViT network for offline handwritten signature recognition and verification. Complex and Intelligent Systems. https://doi.org/10.1007/s40747-025-02011-7