Identifying Who You Are No Matter What You Write through Abstracting Handwriting Style

Jinyang Huang[1], Yuanhao Feng[1], Feng-Qi Cui[1], Xiang Zhang*, Zhi Liu, Xin Liu, Jianchun Liu, Fusang Zhang, and Meng Li*

***Abstract*—With the increasing use of electronic devices, online handwriting verification has become crucial for biometrics-based identity authentication. Traditional methods, which rely on content-dependent verification of the writer's name, are vulnerable to forgery. This paper introduces a content-independent handwriting authentication system, *Ph-Wri*, designed for commodity smartphones. The core innovation is a multi-path attention feature fusion network that combines both static features (image of the handwritten text) and dynamic features (time-dependent properties during writing), to abstract the handwriting style instead of specific content for recognition, enabling robust user authentication. To extract handwriting style from dynamic writing features, we propose a polarity-aware attention strategy during training. This strategy incorporates Style Channel Attention (SCA) to capture direction-sensitive stylistic features, and Trajectory Spatial Attention (TSA) to highlight key handwriting trajectory regions. In the fine-tuning stage, the Correlation-Aware Attention (CAA) module models inter-channel structural correlations, mitigating the influence of content and enhancing style-consistent representations. By linking content-independent handwriting style to user identity, the system achieves accurate authentication. Extensive experiments on both the self-built CIEHD dataset and the public BiosecurID dataset demonstrate exceptional performance, achieving a 99% Verification Accuracy on CIEHD. Compared to state-of-the-art methods that utilize only static or dynamic data, *Ph-Wri* significantly reduces the Equal Error Rate, showcasing the effectiveness and practicality of the proposed approach.**

***Index Terms*—Handwriting verification, style extraction, content-independent, multi-modal fusion.**

I. Introduction

HANDWRITING represents a fundamental and pivotal biometric feature, primarily employed in the domains of finance, commerce, and forensics [1]. Therefore, it is essential to explore identity verification based on handwriting characteristics. With the widespread use of electronic devices and digital documents, online electronic handwriting has become a prevalent method for identity authentication, significantly enhancing users' information security [2], [3]. The field of handwriting verification presents several formidable challenges, with the most notable ones including high inter-class similarity, significant intra-class variability, and limitations in database size. In particular, intra-class variability stems from inherent differences in handwriting samples collected from the same individual under varying motion states. Furthermore, dataset limitations are rooted in constraints related to data collection conditions [3]–[6].

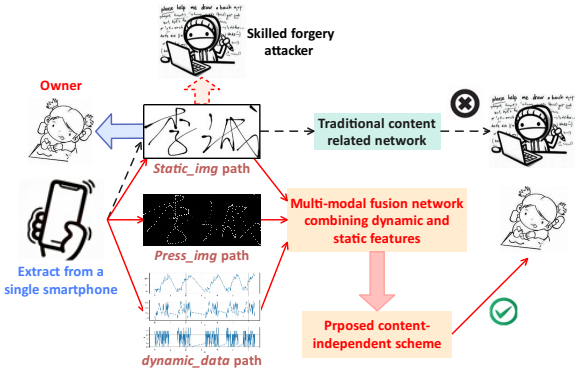


Fig. 1. User authentication process difference between content-dependent methods and the proposed system.

A pressing concern revolves around the issue of high inter-class similarity, primarily fueled by the threat of skilled forgery attacks within handwriting verification systems. Skilled forgers meticulously replicate the handwriting of legitimate users, possessing knowledge of both content and style. These impostors manipulate authentication systems to access secure environments using fraudulent handwritten imitations [7], [8]. Due to the striking resemblance between forged and genuine handwriting, systems with weak security measures are often vulnerable to these sophisticated attacks, jeopardizing the rights and interests of legitimate users. In Fig. 1, we illustrate the distinct authentication processes of content-dependent methods and our proposed system, *Ph-Wri*, for highly similar handwritten signatures from both the signature owner and skilled forgery attackers. The black dotted line depicts the authentication process of content-dependent methods. These methods, which rely on static images and primarily focus on the written content, are prone to errors in signature recognition. In contrast, our proposed system, *Ph-Wri*, leverages dynamic information from the writing process and effectively extracts content-independent writing style features, thus improving handwriting verification performance.

Handwriting verification can be basically classified into two modalities, each based on different input sources:

- **Static Handwriting Verification**: This modality relies on static images of handwritten text, and optical scanners are commonly employed for acquiring such static handwriting samples. Although previous researchers

Jinyang Huang, Xin Liu, and Meng Li are with the School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China.

Feng-Qi Cui, and Jianchun Liu, are with the School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China.

Xiang Zhang is with the School of Cyber Security, Tianjin University, Tianjin 300072, China.

Yuanhao Feng, and Zhi Liu are with the Department of Computer and Network Engineering, The University of Electro-Communications, Tokyo 1828585, Japan.

Fusang Zhang is with Beihang University, Key Laboratory of Space-Air-Ground Integrated Network Security, Ministry of Industry and Information Technology, Beijing 100190, China.

[1]Equal contribution.

Corresponding author*: Xiang Zhang, Meng Li (Email: zhangxiang@ieee.org, mengli@hfut.edu.cn).

have largely focused on improving the effectiveness of image-based descriptors, they inevitably ignore the dynamics of the handwriting process, *i.e.*, the temporal correlation between different handwriting movements in handwriting, which inevitably downgrades the recognition performance.

- **Dynamic Handwriting Verification**: The verification of this modality relies on the time-dependent properties generated during the handwriting act. Dynamic data includes local elements of the handwriting process, *e.g.*, pen position and pressure, from which temporal handwriting characteristics can be extracted. Different from static handwriting predominantly concerns image-based attributes, dynamic handwriting emphasizes temporal and structural information. Apparently, single-modal approaches fail to fully capture the comprehensive spectrum of handwriting behavior.

As a result, a more promising approach is proposed, which involves the integration of these two modalities and leverages the synergies between static and dynamic handwriting characteristics. Prior studies have demonstrated the efficacy of data feature fusion in significantly reducing verification error rates [9]–[11]. However, most of these pioneer methods were confined to extracting basic structural features from static handwriting or conducting simplistic data matching for dynamic handwriting. Furthermore, specific works have predominantly centered on signature synthesis rather than offering a comprehensive exploration of the fusion of dynamic and static handwriting. Hence, there is a need to explore more robust and sophisticated fusion techniques to fully unlock the potential of multi-modal handwriting verification.

To address the limitations of single-modal handwriting verification, this paper introduces a multi-path model that fuses dynamic and static handwriting features. Our model, rooted in a convolutional neural network (CNN), recognizes the temporal and spatial interplay between dynamic and static handwriting. The design's goal is to extract richer and more diverse features from multi-modal handwriting data [12]. On the other hand, handwriting verification, like visual tasks such as face recognition, faces inherent challenges, including high inter-class similarity and substantial intra-class variability. In both domains, it is crucial not only to extract discriminative representations but also to enforce a supervision strategy that enhances feature separability [13]. To address these challenges, we incorporate the *Arcface* loss function, which has proven highly effective in face recognition tasks. The success of *Arcface* in visual domains, where distinguishing subtle differences between individuals is critical, makes it particularly suitable for handwriting verification. By leveraging *Arcface* loss, the model aims to improve the separation of handwriting features, enhancing the overall verification accuracy.

In practical applications of handwriting authentication, like signature verification, the key distinguishing factor between genuine handwriting and skilled forgeries is the writer's distinctive style rather than the content of the handwriting [14]. This highlights the significance of concentrating on the unique features of the writer's handwriting style, which are difficult to replicate, to accurately ascertain the authenticity of the handwriting.

To address this challenge, we propose a content-independent online handwriting verification system *Ph-Wri*. In this system, the specific content of the handwriting, which can be any text written by the individual, is not considered a limiting factor. This system offers strong protection against skilled forgery attacks because the attacker is not required to mimic specific content, making it more challenging for them to replicate the writer's handwriting style. Our goal is to reduce the impact of handwritten content by employing three key techniques: data enhancement, a style attention mechanism, and loss function construction. By diminishing the significance of content in the extracted features, we can highlight the handwriting style, which is extremely difficult for attackers to reproduce.

Specifically, data enhancement serves a dual purpose: it expands the dataset, reduces the risk of model overfitting, and diversifies the handwriting content, providing more resources for extracting content-independent handwriting style features. The style attention mechanism is crafted to enhance the influence of significant channels in the middle network layers, directing the model's attention toward handwriting style features. This method amplifies the importance of these features in the model's analysis. The loss function consists of two components: one quantifies the loss related to the model's ability to accurately recognize the authenticity of handwriting, while the other component aims to minimize the impact of handwriting content on feature extraction, enabling the model to focus on content-independent handwriting style features. This approach reinforces the model's ability to differentiate handwriting based on style while reducing its reliance on content-specific information.

Totally, the main contributions of this paper are as follows:

- As far as we know, we are the first to introduce a content-independent handwriting authentication system *Ph-Wri* based on a commodity smartphone. To better capture handwriting act style features, we design a multi-modal fusion network with full interaction between static and dynamic handwriting at the data level, network layer level, and feature level to compensate for the incompleteness and inaccurate of a single modality.
- To abstract handwriting style from handwriting act features for better user authentication, a polarity-aware attention learning strategy that incorporates Style Channel Attention (SCA) and Trajectory Spatial Attention (TSA) are pertinently designed to extract direction-sensitive stylistic features and highlight trajectory-aware spatial regions that are most indicative of individual writing behavior.
- To further eliminate the handwriting content influence and enhance style-consistent representations, by modeling inter-channel structural dependencies through the handwriting Correlation-Aware Attention (CAA) module in the fine-tuning stage, we significantly strengthen the expression of style irrelevant to content.
- To comprehensively evaluate the performance of *Ph-Wri*, we not only create an online handwriting dataset, CIEHD, consisting of 9000 samples from 50 individuals, but also assess its performance on the widely-used public BiosecurID dataset. Extensive experiments demonstrate that *Ph-Wri* achieves an average Verification Accuracy of 99% on the CIEHD dataset. Furthermore, by leveraging the advantages of both data types, *Ph-Wri* achieves a 7.15% lower average Error Rate compared to state-of-the-art (SOTA) methods on

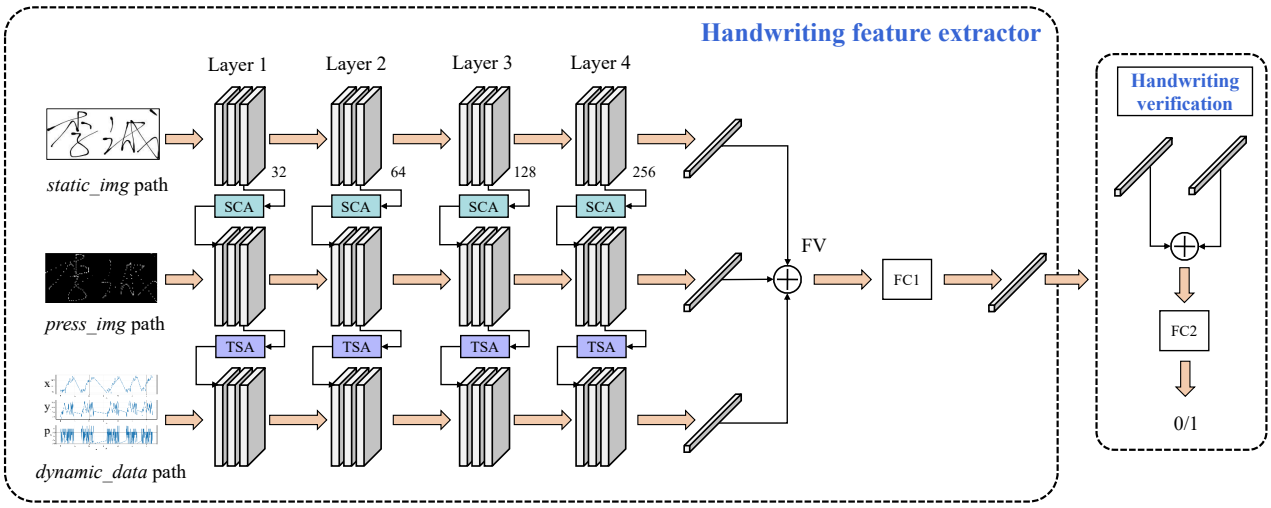


Fig. 2. The architecture of the proposed multi-path fusion network. The *static_img* path takes as input a static image, while the *press_img* path takes a pressure image generated from dynamic pressure data as input. The *dynamic_data* path receives dynamic data as input. The fused feature vector (*FV*) is obtained by combining the results from these three paths and passing them through a fully connected layer (FC1). FC2 incorporates the concept of an *Additive Angular Margin* (*Arcface*).

the BiosecurID dataset.

II. Related Work

In this section, we introduce relevant SOTA works on handwriting verification, including static feature-based verification, dynamic feature-based verification, and multi-modal fusion-based handwriting technologies.

A. Single-modality handwriting verification

Static features for handwriting authentication are typically acquired using optical scanners that capture the written material after the writing process. This approach requires binarization and morphological operations [15], [16] to enhance symbol traces and remove noise from the signature image. A common method for this involves applying a denoising filter, such as a median filter [17], followed by a general binarization technique to establish the optimal threshold for extracting handwritten traces from the background. Subsequently, texture-based methods like LBP and GLCM are used to extract static handwriting features. For example, Fahn *et al.* [18] utilized a refinement algorithm to capture skeleton information of the handwriting trajectory, leveraging branch nodes in the skeleton to extract features. Ghosh *et al.* [19] also focused on graphology-based features to describe handwriting. To capture global features from static handwritten images, CNN-based methods were introduced [20], offering improvements in the feature extraction process. Additionally, the use of Generative Adversarial Networks (GANs) [21] has brought new perspectives to handwriting image processing. While these techniques offer advancements, they still depend heavily on visual attributes of handwriting and do not account for the temporal and dynamic factors that are crucial in distinguishing between genuine handwriting and forgeries.

Despite these efforts, static feature-based authentication suffers from several significant drawbacks. First, the focus on visual attributes, *i.e.*, texture and stroke, direction, makes the system vulnerable to skilled forgery attacks, where attackers can easily replicate the content and appearance of the handwriting. Moreover, this reliance on visual characteristics ignores the dynamic, real-time aspects of the handwriting process, *e.g.*, writing speed and pressure, which could provide additional layers of identity verification.

In dynamic handwriting authentication, data is typically collected using devices like digital handwriting tablets, capturing a temporal sequence of values, *e.g.*, pen position, tilt, pressure, and other factors. Feature extraction from dynamic handwriting can be approached in two main ways: global-based and stroke-based. The global-based approach, exemplified by the Gaussian Mixture Model-Universal Background Model [22], is effective in capturing general handwriting characteristics. Other methods, such as the Color Coherence Vector [23], also provided useful dynamic features and demonstrated their effectiveness. On the other hand, the stroke-based approach focuses on the finer details of individual strokes. Novel techniques, *e.g.*, sparse coding [24] and time-series coding [25], have been successfully applied in writer identification. Recurrent Neural Networks (RNNs), known for their ability to model temporal dependencies, have also been explored to enhance dynamic handwriting verification. For instance, Zhang *et al.* [26] introduced a set of random hybrid strokes, while Lai *et al.* [27] proposed a Gated Autoregressive Unit (GARU) to extract dynamic features for fine-grained handwriting verification.

Despite their effectiveness, dynamic feature-based methods face significant challenges. Dynamic data is highly sensitive to variations in writing conditions, such as pen speed or pressure, leading to substantial intra-class variability. This makes it difficult to achieve consistent recognition across different writing samples from the same individual. Furthermore, dynamic handwriting authentication primarily focuses on the temporal aspects of writing, which may overlook the critical visual elements of handwriting that are also essential for identity verification.

Therefore, while dynamic features provide valuable insights, they are not sufficient on their own. By integrating dynamic features with static ones, these systems can significantly reduce intra-class variability and enhance forgery detection. This fusion leverages both the temporal and spatial characteristics of handwriting, offering a more comprehen-

sive and robust solution. Attackers would find it considerably harder to replicate both the handwriting style (static features) and the dynamic writing behavior (dynamic features), thereby strengthening the overall authentication system's performance.

B. Multi-modal handwriting verification

The primary contribution of previous work on combining dynamic and static handwriting features is the enhancement of the dataset, addressing the problem of limited handwriting data. A common approach is to synthesize offline handwriting from online handwriting samples. Naouel *et al.* [28] proposed using a Gaussian model to synthesize offline signatures from real dynamic handwritten information. These synthetic samples perform similarly to real static handwritten signatures, thereby improving the identification of skilled forgery handwriting. Melo *et al.* [29] employed a deep convolutional neural network to generate and synthesize offline signatures, with the goal of enlarging the offline training dataset and improving offline signature verification performance. Furthermore, pioneer studies in [30], [31], and [32] have demonstrated that multi-modal handwriting feature fusion can significantly reduce the error rate of verification. Zhang *et al.* [33] also explored the complementarity between online and offline handwriting in handwritten mathematical expression recognition, combining both modes of features to effectively improve recognition performance.

While these works have advanced the research on dynamic and static handwriting feature fusion, they overlook the critical issue of the correlation between static and dynamic handwriting features. In particular, the fusion methods used in these studies focus primarily on enhancing dataset size and feature diversity but fail to explicitly address how dynamic and static features interact to capture handwriting style characteristics more effectively. Furthermore, most existing handwriting verification systems remain content-dependent, meaning they do not adequately mitigate the influence of the handwritten content on the verification process. Nevertheless, content-independent handwriting verification, by contrast, focuses primarily on extracting handwriting style characteristics while weakening the impact of the written content. This approach significantly helps reduce intra-class similarity and significantly enhances the ability to resist skilled forgery attacks.

Despite the promise of content-independent methods, most pioneering works [34]–[36] require large samples of handwritten text to accurately extract the writer's style characteristics, making data collection more difficult [37]. Therefore, developing a method that can effectively extract handwriting style features from a smaller set of handwritten characters is of great importance. This would not only simplify data collection but also enhance the accuracy and reliability of content-independent handwriting verification systems in real-world applications.

III. Multi-path handwriting style extraction network

A. System overview

In this section, we introduce the handwriting feature extractor across multiple modalities and detail the process of multi-modal fusion for handwriting style extraction.

(a) (b) (c)

Fig. 3. The preprocessing of images. (a) Original static image. (b) Preprocessed static image. (c) Pressure image converted from dynamic pressure data.

A writer's online handwriting comprises both static images, which visually represent the handwriting, and dynamic time-series data, such as horizontal and vertical coordinates (*x,y* coordinates) and pressure information. To fully capture the handwriting characteristics, our approach employs a multi-path feature fusion network, consisting of several independent feature extraction paths. These paths are designed to extract static and dynamic handwriting features as well as the complementary information between them. Recognizing the inherent relationship between static appearance and dynamic motion in handwriting, the proposed network integrates data-based fusion, network-based fusion, and feature-based fusion to achieve comprehensive multi-modal feature integration.

As illustrated in Fig. 2, the *static_img* path processes a preprocessed static handwritten image, normalized to 224×224 pixels. Standard image enhancement techniques, such as graying and dilation (Fig. 3(a), Fig. 3(b)), are applied to reduce noise and highlight the handwriting traces. To extract texture features from the static image, we employ a Residual Convolutional Network (ResNet-18) [38], which exhibits rotation invariance and consists of four convolution modules, each containing four convolution layers. Batch normalization (*BN*) layers and ReLU activations [39] are applied throughout the network to stabilize and enhance feature learning. After the convolutional modules and a fully connected layer, we finally obtain the feature representation of the static handwriting.

Style Channel Attention: To extract discriminative handwriting style features from static images, we introduce the SCA module into this path. The SCA module enhances channel-wise stylistic representations by decoupling and reweighting positive and negative channel responses. This polarity-aware modeling enables the network to focus on style-specific features and improves the extraction of individual writing styles critical for authentication.

The pressure image (*press_img*) is constructed by mapping the time-series pressure signals onto a spatial grid aligned with the handwriting trajectory. Specifically, the dynamic handwriting sequence (x_t, y_t, p_t) is first normalized and projected onto a 224×224 grid. Pressure values are aggregated at each pixel location and normalized to $[0, 1]$, resulting in a single-channel pressure image that preserves the spatial structure of handwriting while encoding pressure intensity. This pressure image, derived from dynamic pressure data, encodes the pen pressure intensity at each handwriting position, which provides valuable structural information that bridges both dynamic and static handwriting features. The *press_img* path shares the same network architecture as the *static_img* path to extract complementary image-level features, as shown in Fig. 2.

Moreover, the *dynamic_data* path processes the dynamic handwriting sequence, which includes *x,y* coordinates. The time-series data is first segmented into 16 parts based on

TABLE I
PRE-EXTRACTION FEATURES OF DYNAMIC HANDWRITING

Features	Descriptors
X-coordinate	x_t
Y-coordinate	y_t
X-velocity	$\dot{x}_t$
Y-velocity	$\dot{y}_t$
X-acceleration	$\ddot{y}_t$
Y-acceleration	$\ddot{y}_t$
Path-tangent angle	$\theta_t = arctan\,(\dot{y}_t/\dot{x}_t)$
Velocity magnitude	$v_t = \sqrt{\dot{x}_t^2 + \dot{y}_t^2}$
Log curvature radius	$\rho_t = log\left(v_t/\dot{\theta}_t\right)$
Total acceleration magnitude	$a_t = \sqrt{\dot{v}_t^2 + v_t \cdot \dot{\theta}_t}$
First order derivatives of the above signals	$\dot{\theta}_t$ $\dot{v}_t$ $\dot{\rho}_t$ $\dot{a}_t$

stroke information, with each segment comprising 224 data units. Each unit contains a 14-dimensional feature vector calculated from the *x,y* coordinates, as summarized in Tab. I. If the stroke count or length is insufficient, zero padding is applied. These segments are then temporally aligned to form a two-dimensional input of size $(14 \times 16) \times (14 \times 16)$. This path employs the same ResNet-18 structure as the *static_img* path to extract dynamic handwriting features related to stroke position and trajectory.

Trajectory Spatial Attention: To further capture style-related spatial patterns in the writing process, we incorporate the TSA module after the convolutional layers. TSA enhances the spatial sensitivity to writing trajectories by modeling polarity-separated spatial activations and propagating them over a graph-based spatial adjacency structure. This mechanism highlights trajectory regions with distinctive directional transitions and structural handwriting flows, helping to extract stable handwriting style features that are less influenced by content.

Building upon this multi-path feature extractor, we design a multi-level fusion strategy that operates at the data level, network layer level, and feature level. The creation of the *press_img* represents the data-level fusion of pressure information and static handwriting appearance. This synthesis, and its correlation with the corresponding static handwriting, is illustrated in Fig. 3(c) and Fig. 3(a). Through these fusion mechanisms and the proposed attention modules, our network is able to extract handwriting style features that are crucial for accurate and robust handwriting authentication.

B. Network layer fusion

In handwriting style extraction, stylistic cues are often expressed as directional asymmetries, such as preferred stroke orientations, pressure emphasis, and motion transitions, which are inherently different from content-dependent character structures. Conventional attention mechanisms aggregate activations symmetrically and thus tend to preserve dominant content-related patterns.

To address this limitation, we design two polarity-aware attention modules, *i.e.*, SCA, and TSA, which are integrated into intermediate layers of the network to guide cross-modal feature fusion between *static_img* path, *press_img* path, and *dynamic_data* path. The proposed SCA and TSA explicitly

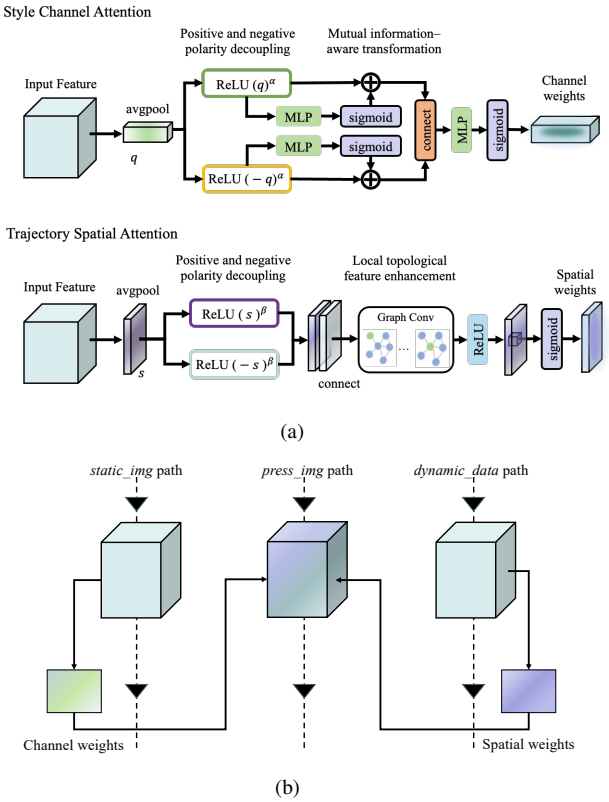


Fig. 4. Network_based fusion: (a) Style Channel Attention (SCA) and Trajectory Spatial Attention (TSA). (b) Processing.

decouple positive and negative activations. This polarity-aware design enables the network to emphasize directional and trajectory-sensitive features that are stable across different contents but distinctive across writers, making it particularly suitable for content-independent handwriting style abstraction. Fig. 4 depicts the essential principle of SCA and TSA and the specific processing flow.

Technical Details of Style Channel Attention: Specifically, SCA is a polarity-aware mechanism designed to emphasize channel-wise stylistic cues in the static handwriting feature map. Given an intermediate feature map $x \in \mathbb{R}^{B\times C\times H\times W}$ from the *static_img* path, we first conduct global average pooling to obtain a compact channel descriptor as follows:

$$q = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{:,:,i,j} \in \mathbb{R}^{B\times C}, \tag{1}$$

where B denotes the batch size, C the number of channels, and H, W the height and width of the feature map, respectively. $x \in \mathbb{R}^{B\times C\times H\times W}$ represents the intermediate feature map extracted from the static image branch. Next, to capture directional style tendencies, we perform positive and negative polarity decoupling using:

$$q^+ = \mathrm{ReLU}(q)^{\alpha_p}, \quad q^- = \mathrm{ReLU}(-q)^{\alpha_p}, \tag{2}$$

where $q \in \mathbb{R}^{B\times C}$ is the global average pooled descriptor across spatial dimensions. q^+ and q^- denote the positive and negative polarity-enhanced components, modulated by a non-linearity parameter α_p.

Then, each polarity stream is passed through an independent MLP and sigmoid branch to generate preliminary

polarity-aware activations as follows:

$$\gamma^+ = \sigma(\text{MLP}(q^+)), \quad \gamma^- = \sigma(\text{MLP}(q^-)). \tag{3}$$

Instead of simply treating these two activations as two independent channels, these activations are further combined via element-wise addition: $\gamma_{\text{pre}} = \gamma^+ + \gamma^-$. To better capture higher-order interactions between the polarities, the combined activation vector γ_{pre} is concatenated with the original q, and fed into a second transformation module composed of another MLP and sigmoid, forming a mutual information–aware transformation:

$$\gamma = \sigma(\text{MLP}(\text{Concat}(q, \gamma_{\text{pre}}))) \in \mathbb{R}^{B\times C}. \tag{4}$$

Finally, to effectively extract the handwriting style and ignore the influence of irrelevant factors, this channel attention map γ is broadcast and used to reweight the original feature map as follows:

$$x_{\text{sca}} = x \cdot \gamma[:, :, None, None]. \tag{5}$$

This process enables SCA to highlight direction-sensitive channel activations that are discriminative to individual writer styles. Moreover, in our fusion framework, the SCA-generated weights from the *static_img* path are used to reweight the corresponding feature responses in the same layer of the *press_img* path, facilitating consistent cross-modal style enhancement.

By leveraging the polarity decomposition and mutual information modeling at the channel level, SCA provides a theoretically grounded means to disentangle style-specific features from *static_img* path, thus effectively extracting the handwriting style that corresponds to writer identity in static images, which is beneficial for more accurate handwriting-based identity authentication.

Technical Details of Trajectory Spatial Attention: To simultaneously extract the handwriting style from the dynamic handwriting features, TSA focuses on enhancing spatial sensitivity to writer-specific trajectory patterns by modeling polarity-aware spatial activations. Given an intermediate feature map $x \in \mathbb{R}^{B\times C\times H\times W}$ from the *dynamic_data* path, we first perform channel-wise average pooling, which can be denoted as:

$$s = \frac{1}{C}\sum_{c=1}^{C} x_{:,c,:,:} \in \mathbb{R}^{B\times H\times W}, \tag{6}$$

where $s \in \mathbb{R}^{B\times H\times W}$ is the spatially aggregated map obtained by averaging over channels. Then, a positive–negative polarity decomposition is applied over the spatial plane, which can be denoted as follows:

$$s^+ = \text{ReLU}(s)^\beta, \quad s^- = \text{ReLU}(-s)^\beta, \tag{7}$$

where s^+ and s^- represent the polarity-enhanced spatial responses modulated by exponent β, and these two polarity maps are concatenated to form a polarity-enhanced spatial tensor: $s_{\text{cat}} = \text{Concat}(s^+, s^-) \in \mathbb{R}^{B\times 2\times H\times W}$.

To first obtain a coarse attention map, this tensor is passed through a convolutional layer followed by a sigmoid gate, which can be expressed as:

$$\gamma_{\text{tsa}} = \sigma(\text{Conv}(s_{\text{cat}})) \in \mathbb{R}^{B\times 1\times H\times W}. \tag{8}$$

To further incorporate local spatial structure, we construct a grid graph $\mathcal{G}$ over the spatial field and apply depthwise graph convolution to simulate trajectory propagation, which can be expressed as follows:

$$\hat{s} = \text{DW-GraphConv}(\gamma_{\text{tsa}}, \mathcal{G}) \in \mathbb{R}^{B\times 1\times H\times W}. \tag{9}$$

The generated refined spatial attention $\hat{s}$ is used to reweight the pressure pathway feature maps to better obtain the style features in the corresponding layer, which can be represented as:

$$x_{\text{tsa}} = x \cdot \hat{s}. \tag{10}$$

This design enables TSA to highlight spatial regions that are rich in directionality, curvature, and stroke transitions. Apparently, these attributes are critical for modeling the topological handwriting style. By integrating spatial polarity decomposition and graph-based topological filtering, TSA provides a principled mechanism to enhance trajectory-sensitive features while mitigating superficial content dependency, thereby further extracting features related to the user's handwriting style and reducing the reliance on handwritten content to a certain extent.

Design Rationale of Cross-modal Reweighting: In our multi-path architecture, the *press_img* path is intentionally designed as a modality-bridging representation between static appearance and dynamic motion. While the *static_img* path focuses on visual stroke morphology and the *dynamic_data* path encodes temporal trajectory patterns, the *press_img* path preserves spatial alignment while embedding dynamic pressure cues, which makes it a natural anchor for cross-modal interaction.

Specifically, the SCA module extracts channel-wise, direction-sensitive style cues from the *static_img* path that are strongly correlated with stroke orientation and structural emphasis. Reweighting the *press_img* features using SCA-generated weights encourages pressure-based representations to focus on stylistically salient channels identified from static appearance. Similarly, the TSA module captures trajectory-aware spatial patterns from the *dynamic_data* path, which highlights regions associated with distinctive writing motions. Apparently, applying TSA-generated spatial weights to the *press_img* path aligns pressure features with motion-sensitive trajectory cues.

Through this design, the *press_img* path serves as a style-aligned fusion anchor, enabling coherent integration of complementary static and dynamic style information, rather than simple feature concatenation.

C. Feature layer fusion

To extract a wealth of information that can be used to distinguish the authenticity of handwriting style, the proposed method concatenates the feature vectors obtained by different paths before the first full connection layer (FC1). Then, each feature vector is given a reasonable weight allocation to obtain more useful information, which can effectively represent the features of handwriting style. Different from the simple fusion of final decision results [40], the proposed feature-based fusion provides flexibility for the prediction of the task network, and the feature-based fusion can be expressed as:

$$FV = \sum_{i=1}^{n_v}\sum_{j=1}^{n_o} \lambda_j^i f v_j^i, \tag{11}$$

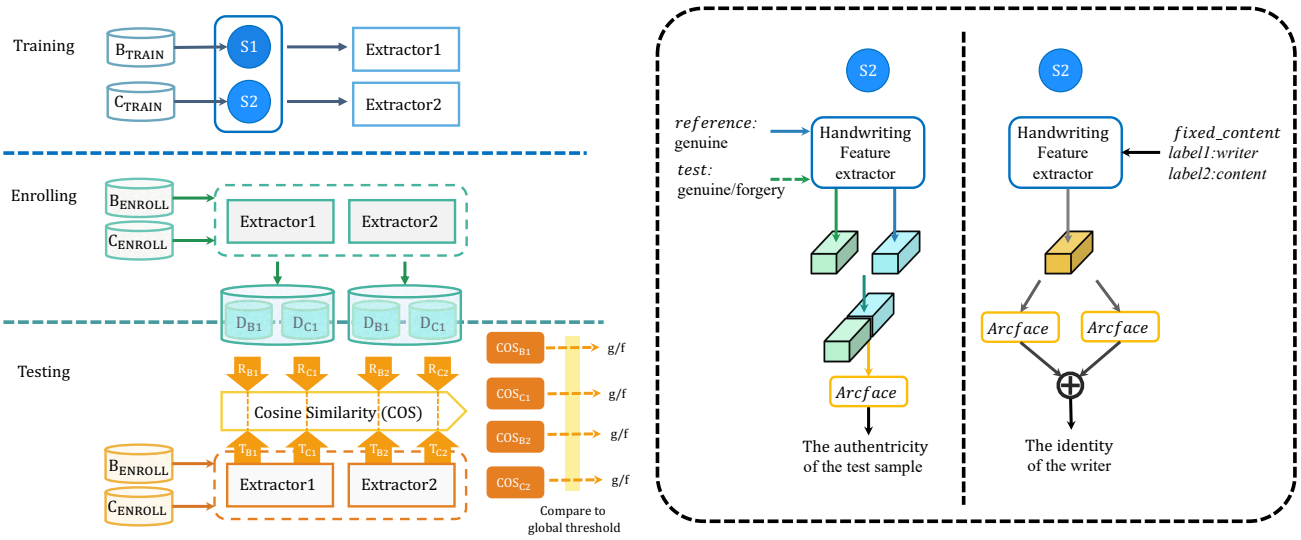


Fig. 5. The workflow of the handwriting verification system. In the *Training* process, we use two datasets to train the proposed model to identify authenticity and alleviate the influence of content by *S1* and *S2*, respectively. In the *Enrolling* process, we randomly selected r samples to generate the writer's handwriting feature template. In the *Testing* process, we calculate the cosine similarity between the *FV* of the test sample and the reference template. If it is higher than the threshold, the test sample will be judged as genuine(g). Otherwise, this ample will be judged as forgery(f).

where fv_j^i and λ_j^i are the value and the weight of the j-th component in i-th feature vector obtained by corresponding path, respectively. n_v and n_o denote the number of feature vectors participating in fusion and the number of components of each feature vector, respectively.

By concatenating features extracted by the three paths [41] and obtaining the final result after passing through two fully connected layers, the proposed model effectively fuses multiple features. In particular, we first choose the embedding features of the first fully connected layer as *FV*. Then, to better measure the loss of whether the input sample is correctly classified, the process of *Addictive angular margin* (*Arcface*) [42] is employed in the second fully connected layer due to it introducing intervals between classes in the angular space.

IV. Content-Independent Online Handwriting Verification

To effectively counter skilled forgery attacks, we propose a content-independent online handwriting verification scheme that minimizes the impact of handwritten content on the verification process. The core idea behind this approach is to train the proposed handwriting feature extractor using handwriting samples from various writers, each containing different specific content. This training process encourages the feature extractor to ignore the content of the writing and instead focus on the distinctive handwriting style. As a result, the system becomes more robust, enabling it to accurately perform identity verification even when faced with various types of forged handwriting.

A. Dataset

To comprehensively validate the effectiveness and generalizability of the proposed handwriting authentication system, we conduct experiments on both a widely used public handwriting dataset and a self-curated dataset. The public dataset enables us to assess the performance of our approach under standardized evaluation protocols and diverse writing scenarios, while the self-built dataset allows us to specifically evaluate the system's ability to extract content-independent handwriting style features in real-world user-centric applications. By validating the system from these two complementary perspectives, we aim to demonstrate its robustness across different writing environments, handwriting contents, and participant groups.

The BiosecurID database, a frequently employed resource for handwriting verification, comprises handwritten signature samples from 132 writers gathered across four different sessions. Specifically, each writer contributes a total of 16 genuine samples and 12 skilled forgery samples. These samples encompass two handwriting modalities: a grayscale static image and its corresponding dynamic data, which includes *x, y* coordinates as well as pressure information.

Furthermore, to investigate the extraction of content-independent handwriting style features, we construct a content-independent Chinese Electronic Handwriting Dataset (CIEHD) sourced from 50 individuals. This dataset contains both genuine and skilled forgery handwriting samples across two modalities, all captured using a mobile phone-based handwriting pad application. Particularly, the genuine handwriting samples in CIEHD cover three distinct word categories: *name-content*, *fixed-content*, and *random-content*. The *name-content* refers to each participant's personal name, the *fixed-content* includes five common words written by all participants, and the *random-content* comprises 250 words randomly assigned to the 50 writers, with each writer receiving five distinct words. Each participant is asked to write five samples of their *name-content* and the five *fixed-content* words, along with two samples of each of the five *random-content* words. This writing process is repeated in four different positional scenarios to increase writing variability. As for skilled forgery handwriting, it focuses exclusively on the *name-content* words. Four different imitators observe and replicate the genuine *name-content* handwriting of each writer. As a result, each writer contributes 20 skilled forgery handwriting samples, leading to a comprehensive dataset containing approximately 9000 handwriting samples in total.

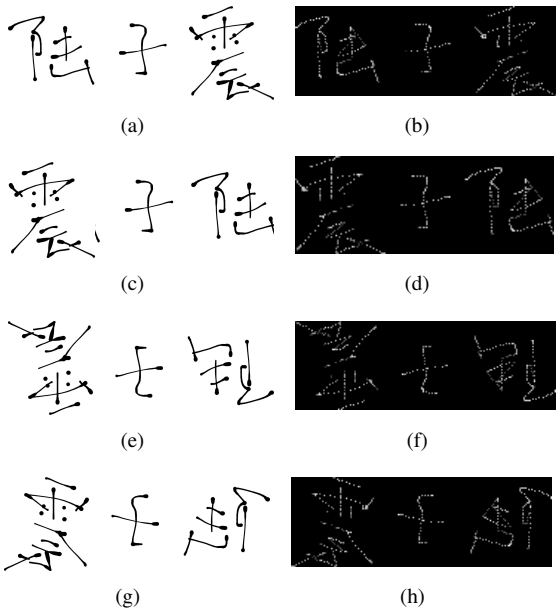


Fig. 6. The implementation of data augmentation. (a) static handwritten image. (b) press_image. (c) static handwritten image after 'cutting and splicing'. (d) press_image after 'cutting and splicing'. (e) static handwritten image after 'rotting'. (f) press_image after 'rotting'. (g) static handwritten image after 'flipping'. (h) press_image after 'flipping'.

B. Implementation of content-independent online handwriting verification

The proposed content-independent handwriting verification system seeks to determine the authenticity of handwriting by mitigating the influence of handwritten content. This is achieved through a combination of data augmentation, a handwriting style attention mechanism, and a specially designed loss function. The system operates in three distinct stages, as illustrated in Fig. 5.

To align with these stages, this dataset is divided into three parts. Approximately 70% of both genuine and forged samples from BiosecurID, as well as 70% of *fixed-content* samples from CIEHD, comprise B_{TRAIN} and C_{TRAIN}. Here, B and C represent the samples from BiosecurID and CIEHD, respectively. The remaining 30% of genuine handwriting samples are used to randomly select r samples, forming B_{ENROLL} and C_{ENROLL}. The rest of the samples are designated as B_{TEST} and C_{TEST}.

1) Data augmentation: Data collection conditions often impose constraints, leading to small handwriting datasets [43]. Consequently, trained network models are prone to overfitting, which limits their generalization performance. A common strategy to address overfitting is the use of data augmentation techniques. These techniques involve transformations such as altering image shapes and other operations that expand the dataset and increase its diversity, typically applied in general image recognition tasks. However, in the context of accurately identifying distinct handwriting or verifying its authenticity, it is crucial to maintain the handwriting style while also accounting for spatial and temporal alignment between the transformed static images and dynamic data. Therefore, our proposed method applies identical processing to both static handwriting and corresponding dynamic data. This includes operations such as flipping, rotating, cutting, and splicing, as shown in Fig. 6.

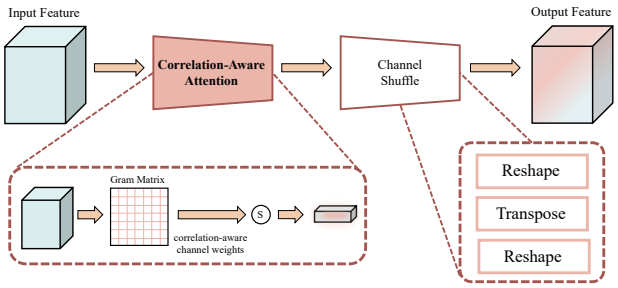


Fig. 7. The implementation of the handwriting Correlation-Aware Attention mechanism.

2) Handwriting Correlation-Aware Attention: To improve feature-level consistency and reduce reliance on fixed channel positions, the proposed handwriting CAA mechanism models inter-channel dependencies within handwriting representations. This mechanism provides the task network with enhanced structural cues for extracting discriminative identity information, as shown in Fig. 7. Each feature map in the input tensor encodes a specific aspect of the handwriting, *e.g.*, straight or curved textures, where the activation intensity indicates the presence of that feature.

By using the Gram matrix of the input features, CAA employs a covariance-based approach to compute the inter-channel correlation structure. The Gram matrix quantifies the pairwise similarity between channel responses and identifies the feature maps with significant structural relevance. This correlation structure is then used to derive a set of importance scores across channels, referred to as correlation-aware channel weights.

Subsequently, the original feature tensor is reweighted according to these attention weights and subjected to a channel shuffle operation. The channel shuffle operation is introduced as a regularization mechanism to disrupt fixed channel semantics after correlation-aware reweighting, encouraging the network to rely on stable inter-channel correlation patterns rather than content-specific channel activations.

In a supervised learning framework, the task network, which is responsible for extracting identity-relevant features, is incentivized to overcome this perturbation and focus more effectively on stable, consistent representations.

Specifically, the $(k, k^{'})$-th element of the Gram matrix in layer l (denoted as G^l) can be computed as:

$$G^l_{kk'} = \sum_{i=1}^{n^l_H} \sum_{j=1}^{n^l_W} a^l_{ijk} \cdot a^l_{ijk'}, \tag{12}$$

where n^l_c, n^l_H, and n^l_W represent the number of channels, height, and width of the feature maps in layer l, respectively. a^l_{ijk} indicates the activation value at position (i, j) of the k-th channel. Then, to derive the relative importance of each channel within the feature map, we compute the importance vector s^l and the normalized weight vector W^l as:

$$s^l_i = \sum_{j=1}^{n^l_c} G^l_{ij}, \tag{13}$$

$$W^l_i = \frac{e^{s^l_i}}{\sum_{j=1}^{n^l_c} e^{s^l_j}}, \tag{14}$$

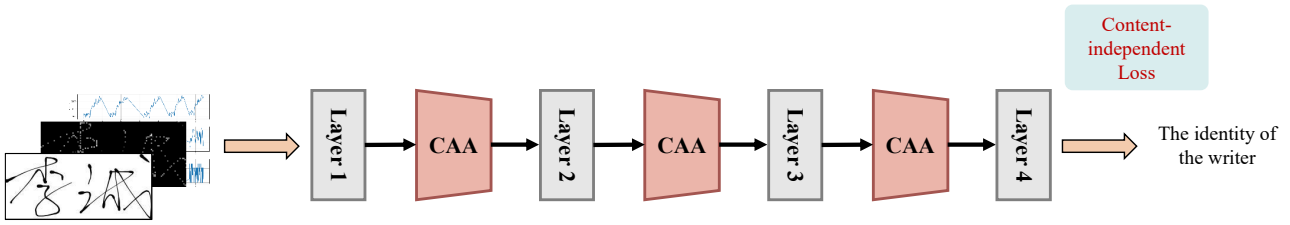


Fig. 8. The workflow of S2. The inputs include static images, dynamic data, and pressure images. Correlation-Aware Attention mechanism module is added to each network layer in the model. The loss function also has the effect of reducing the impact of content. The training task of S2 is to recognize different handwriting, that is, different writers.

where $W^l, s^l \in \mathbb{R}^{n_c^l \times 1}$ are used to reweight the input feature maps along the channel dimension.

3) Loss function based on Arcface: ArcFace, originally proposed in [42] to enhance the discriminative power of features in face recognition, is adopted in our handwriting verification task due to the conceptual similarity between the two domains. Both face recognition and handwriting verification aim to distinguish between individuals based on subtle biometric patterns, making techniques that improve inter-class separability and intra-class compactness transferable across these modalities. Therefore, ArcFace serves as an effective loss function in our setting by enforcing angular margin constraints that strengthen the identity discrimination capability of the learned handwriting features. Specifically, this loss function can be represented as:

$$L_{Arcface}(y,n) = -\frac{1}{N}\sum_{i=1}^{N} log\frac{e^{s\left(cos\left(\theta_{y_i}+m\right)\right)}}{e^{s\left(cos\left(\theta_{y_i}+m\right)\right)}+\sum_{j=1,j\neq y_i}^{n} e^{s\,cos\,\theta_j}}, \quad (15)$$

where s denotes the radius of the hypersphere on which the learned embedding features are distributed. x_i is the feature representation of the i-th sample belonging to class y_i, and W_{y_i} is the weight vector corresponding to class y_i. The angle between x_i and W_{y_i} is denoted as θ_{y_i}, while θ_j represents the angle between x_i and W_j, the weight vector of class j. The scalar m is an additive angular margin penalty introduced to simultaneously improve intra-class compactness and inter-class separability. We denote the batch size as N and the number of classes as n. Based on these definitions, we further design two margin-based penalties tailored to content-independent handwriting verification, formulated as:

$$L = L_{Arcface}(u,p) - \alpha_l L_{Arcface}(c,q), \quad (16)$$

where p and q denote the numbers of writers and *fixed-content* words, respectively. For the i-th sample, u_i and c_i indicate its corresponding writer and content labels. The two loss terms share the same ArcFace formulation defined in Eq. (16), but are applied to different classification objectives: writer identity discrimination and content classification, respectively. The negative sign encourages the model to extract writer-specific features while suppressing content-related information. The scalar α_l is a hyperparameter that balances the contributions of the two objectives. In particular, the hyperparameter α_l is selected based on validation performance, which balances writer discrimination and content suppression. Extremely small or large values lead to insufficient suppression or degraded identity separability, respectively.

We design two distinct training stages, denoted *S1* and *S2*, as illustrated in Fig.5 and Fig.8. In *S1* stage, the primary goal is to pre-train the feature extractor, enabling it to capture handwriting style features specific to the writer. During this stage, the focus is on distinguishing between genuine and forged handwriting by learning the stylistic traits that are unique to each writer. The feature extractor is built upon our multi-path architecture and integrates the proposed SCA and TSA modules to enhance polarity-aware stylistic modeling and spatial trajectory sensitivity. The training inputs consist of *reference-test* pairs, where the *reference* is a genuine sample from the same writer, and the *test* is either a genuine or forged sample from the same writer, excluding the *reference*. Using the handwriting feature extractor, we extract two feature vectors (*FVs*) from the *reference-test* pair. These feature vectors are then fused and passed through the *ArcFace* module, which effectively computes the loss based on the accurate classification of the *test* samples as either genuine or a forgery.

In *S2* stage, the focus shifts towards fine-tuning the pre-trained feature extractor to minimize the influence of handwriting content. The objective here is to decouple the writer's style from the content of the writing, thereby enhancing the coupling between the handwriting style and user identity. To further enhance identity-consistent representations and suppress content-related variations, we incorporate the CAA module during training, which regularizes the extractor through channel-wise correlation modeling and structure-preserving perturbation. This stage involves training on C_{TRAIN}, where each sample is associated with both *label1* and *label2*, representing the writer and content categories, respectively. Each sample is first passed through the pre-trained handwriting feature extractor to obtain a feature vector, which is then fed into two ArcFace-based classification heads corresponding to writer identity and content categories, respectively. The losses $L_{Arcface}(u,p)$ and $L_{Arcface}(c,q)$ are subsequently computed for optimization. These modules compute the losses related to correctly classifying the sample into the *label1* and *label2* categories. Finally, the total loss in *S2* is derived by combining these two individual losses.

During the S2 fine-tuning stage, only fixed-content samples are used to supervise the content-suppression loss $L_{\text{ArcFace}}(c,q)$. This design choice is intentional rather than a limitation. When all writers produce identical lexical content, variations in handwriting are dominated by style-related factors, allowing the model to explicitly identify and suppress content-discriminative cues in a controlled and stable manner. In contrast, directly using random-content samples for content suppression would introduce excessive lexical and structural variability, making it difficult to disen-

tangle content-related factors from writer-specific style under supervised training.

Training: This process focuses on training handwriting feature extractors. Through the implementation of both *S1* and *S2*, the proposed model effectively extracts content-independent handwriting style features, enabling accurate handwriting authenticity identification. Specifically, *Extractor1* is obtained through the *S1* stage, while *Extractor2* is derived by combining both *S1* and *S2*. The goal of this process is to train the model to better capture and differentiate the stylistic features inherent to individual writers, which ensures that the extracted features can be used for effective handwriting verification.

Enrolling: This process involves creating handwriting feature templates for each writer. Feature vectors (*FVs*) are extracted from r registered samples in either B_{ENROLL} or C_{ENROLL} using *Extractor1* or *Extractor2*. The writer's handwriting feature template is then computed as the average of these *FVs*. Consequently, handwriting feature template databases are established, denoted as D_{B1}, D_{C1}, D_{B2}, and D_{C2}. The goal of the enrolling process is to generate a reliable and stable feature template for each writer, which will be used as a reference for later comparisons during the testing process. This template helps in associating each individual's handwriting style with their unique identity.

Testing: In this process, the trained handwriting feature extractors are used for testing. A test sample is randomly selected from either B_{TEST} or C_{TEST}, and its feature vector (*FV*) is extracted using either *Extractor1* or *Extractor2*. This results in four distinct cases: T_{B1}, T_{C1}, T_{B2}, and T_{C2}. The cosine similarity between the *FV* of the test sample and the reference template of the claimed writer, *i.e.*, either R_{B1} from D_{B1}, R_{C1} from D_{C1}, R_{B2} from D_{B2}, or R_{C2} from D_{C2}, is computed. If the similarity score exceeds a predefined threshold, the test sample is classified as genuine; otherwise, it is deemed a forgery. The purpose of the testing phase is to verify the authenticity of a given handwriting sample by comparing it with the pre-established feature templates. This step enables the model to distinguish between genuine and forged handwriting based on the extracted stylistic features.

V. Evaluation

We conduct a series of experiments to showcase the effectiveness of multi-modal fusion in handwriting verification and to demonstrate the practicality of our proposed content-independent approach, which effectively mitigates the influence of handwritten content. Our experiments are carried out using the BiosecurID and CIEHD datasets for both training and testing, with the registered sample size set at three based on our experimental analysis. Each experiment was repeated five times, and the reported results represent the average of these five runs.

A. Data preprocessing results

As depicted in Fig. 3(a) and Fig. 3(b), the preprocessing of static handwriting through grayscale conversion and scaling enhances the visibility of the handwriting skeleton while also creating smoother stroke boundaries. This process effectively retains keystroke information, such as the starting stroke, closing stroke, and turning stroke. For dynamic handwriting, our proposed method incorporates pre-extraction of time function features utilizing information from handwriting track locations and pressure data, as detailed in Tab. I. Following this feature pre-extraction, we obtain a 14-dimensional signal vector that adequately represents the pre-extracted dynamic handwriting characteristics.

TABLE II
Comparison of *static*, *dynamic*, and *fusion* on BiosecurID.

Modal	Method	EER_SF (%)
static	FINet	16.49
	SherlockNet	20.73
	Fusion at the decision level	25.87
	The proposed method (S1)	14.14
dynamic	SherlockNet	6.43
	Fusion at the decision level	5.42
	The proposed method (S1)	5.06
fusion	FINet	15.74
	SherlockNet	4.92
	Fusion at the decision level	3.52
	The proposed method (S1)	3.64
	The proposed method (S1+S2)	2.45

B. Metrics

We employ several evaluation metrics to comprehensively assess the system's performance, including Equal Error Rate (EER), False Rejection Rate (FRR), False Acceptance Rate (FAR), Classification Accuracy (CA), and Verification Accuracy (VA). These metrics are selected to evaluate various aspects of the system: CA measures the model's ability to distinguish between different writers, while VA evaluates the model's capability to verify handwriting authenticity. FRR is defined as the ratio of false rejections to the total number of genuine samples, whereas FAR is the ratio of false acceptances to the total number of forgery samples. EER, which represents the point where FRR equals FAR, serves as a key indicator of system error. Lower EER values indicate better system performance. Additionally, to provide a visual assessment of the system's performance, we utilize the Detection Error Tradeoff (DET) curve. For binary verification, a global decision threshold is determined based on the EER criterion on the validation set, and the same threshold is applied to compute FAR, FRR, EER, and VA on the test set.

Accuracy rates, such as CA and VA, represent the proportion of correctly classified samples out of the total samples. Higher accuracy rates reflect improved system performance. These accuracy metrics are further categorized based on the content type of the samples used, resulting in the following detailed categories: CA_FC, VA_SF, VA_FC, VA_RC, EER_SF, EER_FC, EER_RC, FAR_RC, and FRR_RC. In this classification, 'SF' refers to *name-content* with skilled forgery samples, 'FC' indicates *fixed-content* with forgery samples from other writers' content, and 'RC' denotes *random-content* with forgery samples from other writers' *random-content*. In particular, for name-content, the dataset includes skilled forgeries, where human imitators are asked to carefully observe and imitate the genuine handwriting of a target writer. These samples are designed to simulate realistic skilled forgery attacks. On the other hand, for fixed-

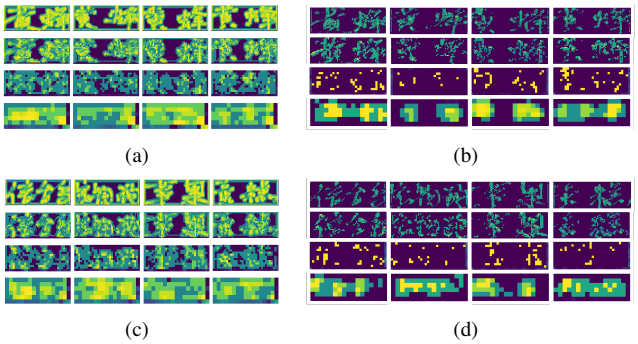


Fig. 9. The visualization results of the feature maps obtained by *Extractor1* and *Extractor2*. From top to bottom are the results of each layer of the proposed network. (a) and (b) are the results of different writers (writer 1, writer 2, writer 3, writer 4) by *Extractor1* and *Extractor2*, respectively. (c) and (d) are the results of different handwritten contents of writer 1 by *Extractor1* and *Extractor2*, respectively.

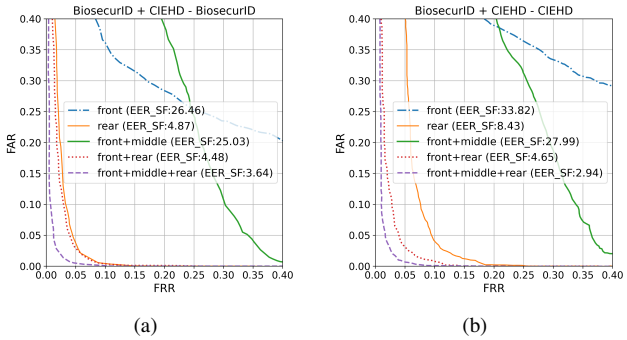


Fig. 10. The DET curves comparison of different fusion methods on two datasets. (a) BiosecurID. (b) CIEHD.

content and random-content evaluations, no skilled imitation process is conducted. Instead, the forgery samples correspond to random forgeries, constructed by treating genuine handwriting samples from other writers as impostor samples for the target writer. By evaluating the system with these diverse metrics, we gain a comprehensive understanding of its strengths and weaknesses in handling different types of handwriting verification tasks.

C. Single-modal vs multi-modal

In Tab. II, we present a comparison of *static*, *dynamic*, and *fusion* methods applied to the BiosecurID dataset, alongside various state-of-the-art techniques, including the Font Imitate Network (FINet) [44] and the Image-based Freeform Handwriting Authentication SherlockNet [45]. Additionally, we compare our approach with a method that combines static images and dynamic information at the decision level, which is denoted as Fusion at the decision level [1].

Our analysis demonstrates that the *fusion* approach achieves the highest performance, with the *dynamic* method closely following. This highlights the effectiveness of combining both dynamic and static data in improving handwriting verification performance. Notably, the proposed method outperforms SOTA techniques such as those in [44] and [45], illustrating the advantages of our approach in fusing multi-modal data without relying on handwriting synthesis.

Moreover, after incorporating *S2*, our method is fine-tuned to focus on handwriting style characteristics, which leads to superior performance compared to the Fusion at the decision level [1]. This reinforces the significance of combining both static and dynamic modalities to enhance the robustness and accuracy of handwriting verification.

D. Feasibility analysis of the content-independent handwriting verification

To evaluate the proposed method's ability to capture the distinctive handwriting style features of different writers, we design two key performance indicators. First, we assess whether the model can effectively differentiate between different writers based on their unique handwriting styles when the same handwriting content is provided. Second, we investigate whether the model can still accurately capture a writer's unique handwriting style when presented with distinct handwriting contents.

For this analysis, we utilize the *fixed-content* and *random-content* sample sets from CIEHD, conducting the evaluation under two distinct scenarios: *S1+S2* and *S1*, as summarized in Tab. III. The results show a substantial improvement in verification performance with the addition of *S2* stage. Under the *S1+S2* scenario, CA_FC achieves an impressive 99%, highlighting the method's ability to enhance inter-class differences among writers, even when they share the same handwriting content. Furthermore, the False Rejection Rate (FRR_RC) drops significantly to 5.88% under *S1+S2*, demonstrating the model's capability to accurately extract unique handwriting characteristics from a writer's work, irrespective of the content.

Despite the higher absolute error rate under random-content evaluations, the consistent improvement from S1 to S1+S2 demonstrates that the learned content-suppression mechanism generalizes beyond fixed-content supervision. This observation indicates that S2 fine-tuning effectively weakens content-dependent biases, even when evaluated on previously unseen and unconstrained content.

Fig. 9 illustrates the feature maps generated by both *Extractor1* and *Extractor2*, emphasizing their effectiveness in reducing the influence of handwritten content. Fig. 9(a) and Fig. 9(b) visualize feature activations from different writers writing the same content. It clarifies that before content alleviation, the activations are concentrated in similar spatial regions, indicating that the learned features are still strongly influenced by content-dependent structural patterns. After the content-alleviating process, the activation maps become more spatially diversified and writer-specific, revealing clearer inter-writer differences that correspond to individual handwriting styles rather than shared content structure.

On the other hand, Fig. 9(c) and Fig. 9(d) demonstrate feature activations from the same writer writing different contents. Prior to content suppression, the feature maps exhibit noticeable variations aligned with content-specific stroke layouts. In contrast, after applying the proposed content-alleviating mechanisms, the activation patterns show higher spatial consistency across different contents, indicating that the extracted features are dominated by stable writer-specific style cues instead of content variations.

E. Comparison of fusion methods

Fig. 10 presents the Detection Error Tradeoff (DET) curves for various fusion schemes and their combinations, tested on both the BiosecurID and CIEHD datasets. The fusion schemes include 'front' fusion, which exclusively

TABLE III
THE COMPARISON OF THE PERFORMANCE OF HANDWRITING VERIFICATION OBTAINED BY *S1* AND *S1+S2* ON BIOSECURID AND CIEHD BASED ON DIFFERENT HANDWRITTEN CONTENTS (%).

Training mode	Testset	EER_SF	VA_SF	EER_FC	VA_FC	CA_FC	EER_RC	FAR_RC	FRR_RC	VA_RC
S1	BiosecurID	3.64	93.24	-	-	-	-	-	-	-
	CIEHD	2.94	98.96	3.49	97.07	98.56	8.76	6.48	8.72	93.66
S1+S2	BiosecurID	**2.45**	**96.49**	-	-	-	-	-	-	-
	CIEHD	**0.19**	**100**	**1.70**	**98.65**	**99.95**	**5.71**	**6.13**	**5.88**	**94.38**

utilizes data-based fusion; 'rear' fusion, relying solely on feature-based fusion; 'front+middle' fusion, which combines data-based and network-based fusion; 'front+rear' fusion, integrating both data-based and feature-based fusion; and 'front+middle+rear' fusion, where all three fusion methods are applied within the proposed model.

The results from Fig. 10 highlight that, among the individual fusion methods, 'front' fusion demonstrates the lowest performance, while 'rear' fusion stands out as the most effective. This emphasizes the significant contribution of 'feature_fusion' in the high-dimensional feature space, which plays a key role in enhancing the multi-modal fusion handwriting verification system.

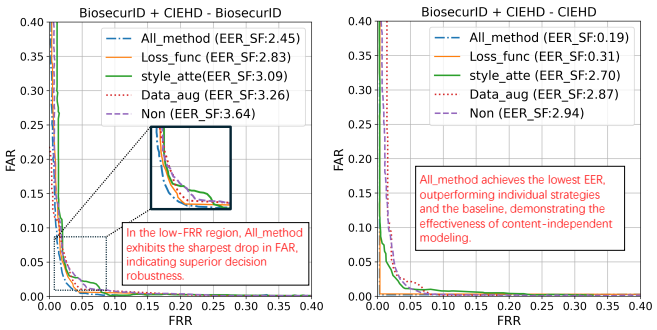
Furthermore, 'front+middle+rear' fusion outperforms all other fusion methods, showcasing the optimal performance achieved when incorporating all three fusion techniques. This suggests that the proposed model's ability to extract handwriting features and determine handwriting authenticity is significantly improved by the combined effect of 'data_fusion', 'network_fusion', and 'feature_fusion'.

The observed improvement from 'front+rear' to 'front+middle+rear' fusion further underscores the effectiveness of our network-based fusion strategy. Specifically, the introduction of the SCA and TSA modules enables polarity-aware channel modeling and trajectory-sensitive spatial alignment, which strengthens the feature interaction across modalities. This structural enhancement allows the network to better capture identity-relevant handwriting patterns, contributing directly to the overall performance improvement.

F. Performance of content-alleviating methods

To evaluate the impact of various content-independence methods on handwriting verification, we conduct experiments using both the BiosecurID and CIEHD datasets. The training process consisted of pre-training (*S1*) and fine-tuning (*S2*), incorporating three content-alleviating techniques: data augmentation (*Data_aug*), CAA mechanism (*Style_att*), and a specially designed loss function (*Loss_func*), along with a combination of all three (*All_methods*). To make a complete comparison, we also include a baseline model (*Non*), which is pre-trained without any content-alleviating methods.

The results, shown in the color-coded DET curves of Fig. 11, are further summarized in Tab. IV. Models 'F0', 'F1', 'F2', 'F2 w/o CS', 'F3', and 'F4' correspond to the settings with different content-alleviating modules: *Non*, *Data_aug*, *Style_att* (attention-based style modeling, *i.e.*, SCA + TSA + CAA), *Style_att* without the channel shuffle, *Loss_func*, and *All_methods*, respectively.



(a) (b)

Fig. 11. The DET curves of different content-reducing schemes based on Biosecurid+CIEHD training set (%). (a) BiosecurID. (b) CIEHD.

From Fig. 11, we can observe that all content-alleviating methods significantly enhance the performance of the pre-trained model compared to the baseline model without any content-alleviating techniques ('F0'). Notably, *Loss_func* stands out as the most effective method in improving the performance of the proposed model. Furthermore, combining all three methods results in a substantial performance boost compared to using any single method alone. This combination leads to a reduced error rate, increased accuracy, and greater stability across multiple test results, demonstrating the synergistic effect of integrating all three methods.

The ablation results further validate the necessity of the proposed components. Simple feature fusion without polarity-aware attention yields limited improvement, indicating that content-dependent correlations are still preserved. Introducing SCA and TSA consistently improves performance under cross-content evaluations, demonstrating their effectiveness in extracting style cues invariant to textual content. Moreover, incorporating CAA and the dual ArcFace objective leads to the most stable gains, especially in random-content settings, highlighting the importance of both feature-level and objective-level decoupling for content-independent handwriting verification.

Furthermore, the effectiveness of the channel shuffle operation is validated by the comparison between F2 and F2 w/o CS in Tab. IV. While both configurations employ the same attention-based style modeling framework, removing the channel shuffle consistently leads to inferior performance, especially under cross-content evaluations. This observation indicates that channel shuffle contributes to improved robustness by mitigating content-dependent overfitting, rather than merely serving as an architectural detail. Although the absolute performance gap is moderate, the consistent gains demonstrate the regularization effect of channel shuffling in content-independent handwriting verification.

TABLE IV
THE PERFORMANCE OF DIFFERENT CONTENT-INDEPENDENCE SCHEMES ON THE BIOSECURID AND CIEHD TEST SET: MEAN (VARIANCE) (%)

Method	BiosecurID				CIEHD			
	FAR	FRR	EER	ACC	FAR	FRR	EER	ACC
F0	3.66 (0.0195)	3.65 (0.0121)	3.64 (0.0189)	96.49 (1.72)	2.62 (0.0082)	2.95 (0.0070)	2.94 (0.0073)	98.22 (2.54)
F1	3.27 (0.0022)	3.28 (0.0026)	3.26 (0.0026)	96.50 (0.73)	2.98 (0.0029)	2.87 (0.0025)	2.87 (0.0022)	98.38 (0.37)
F2 w/o CS	3.18 (0.0070)	3.21 (0.0074)	3.19 (0.0075)	97.03 (1.64)	2.84 (0.0032)	2.82 (0.0027)	2.83 (0.0035)	99.07 (0.037)
F2	3.07 (0.0061)	3.09 (0.0066)	3.09 (0.0062)	97.73 (1.54)	2.70 (0.0009)	2.71 (0.0006)	2.70 (0.0006)	99.92 (0.032)
F3	2.94 (0.0054)	3.12 (0.0070)	2.83 (0.0059)	97.22 (0.29)	0.29 (0.0005)	0.30 (0.0003)	0.31 (0.0007)	100 (0)
F4	2.45 (0.0053)	2.45 (0.0052)	2.45 (0.0075)	99.11 (0.17)	0.24 (0.0000)	0.26 (0.0005)	0.19 (0.0009)	100 (0)

TABLE V
TESTING ON CIEHD, THE INFLUENCE OF THE CHANGE IN THE REGISTERED SAMPLE SIZE ON THE VERIFICATION SYSTEM PERFORMANCE IS SHOWN AS THE MEAN (VARIANCE) OF THE RESULTS OF MULTIPLE EXPERIMENTS (%).

Registered sample size	EER_SF	VA_SF	EER_FC	VA_FC	EER_RC	VA_RC
1	0.324 (0.005)	99.69 (0.062)	2.536 (0.828)	97.92 (0.035)	7.780 (3.366)	92.53 (3.042)
3	0.221 (0.001)	99.92 (0.043)	1.771 (0.067)	98.31 (0.099)	6.818 (0.532)	93.34 (0.869)
5	0.210 (0)	99.92 (0.033)	1.703 (0.133)	98.55 (0.087)	6.239 (2.182)	93.71 (2.251)
7	0.181 (0)	100 (0)	0.952 (0.029)	98.74 (0.043)	5.308 (1.887)	95.64 (3.071)

G. Impact of registered sample size

In practical authentication scenarios, it is crucial to strike an optimal balance between user comfort and system performance [46], [47]. To achieve this, we investigate the impact of the registered sample size on verification performance. As shown in Tab. V, the trends are clear: as the registered sample size increases, the EER generally decreases, while the VA tends to improve.

In our study, we carefully consider both the system's performance and the operational feasibility in real-world user authentication scenarios. Consequently, we choose a registered sample size of 3, as it provides an ideal trade-off between ensuring satisfactory system performance and meeting practical user authentication requirements.

VI. CONCLUSION

This paper introduces an innovative handwriting verification system, *Ph-Wri*, which integrates both static images and dynamic data with a content-independent learning strategy for accurate identity authentication. The system leverages a multi-path feature fusion mechanism to capture and integrate writer-specific features across different handwriting modalities. Notably, polarity-aware attention modules, *i.e.*, SCA and TSA, enhance channel-wise and spatial trajectory modeling, leading to more discriminative handwriting style representations. To mitigate the influence of handwritten content, we propose several content-independent strategies, *i.e.*, data augmentation, the style attention mechanism based on inter-channel dependencies, and a dual-objective loss function. These methods collectively strengthen the system's ability to extract identity-relevant features independent of content variations. Extensive experiments show that the fusion of static and dynamic data outperforms single-modal approaches. Additionally, the application of content-independent strategies significantly improves both verification accuracy and robustness. Unlike traditional static signature systems, *Ph-Wri* enables the creation of electronic signatures that incorporate both visual and temporal handwriting cues, offering enhanced security. This approach holds great potential for use in financial, commercial, and forensic applications.

ACKNOWLEDGMENT

This work is supported by National Natural Science Foundation of China (Grant No. 62302145, No. U23A20303, No. 62372149, No. 62572168, and No. 62422213), Fundamental Research Funds for the Central Universities (Grant No. JZ2025HGTB0225), Key Laboratory of Space-Air-Ground Integrated Network Security, Ministry of Industry and Information Technology, and Major Scientific and Technological Project of Anhui Provincial Science and Technology Innovation Platform (Grant No. 202305a12020012).

REFERENCES

[1] Z. Shi, F. Li, D. Hao, and Q. Sun, "Handwritten signature verification via multimodal consistency learning," *IEEE Transactions on Information Forensics and Security*, 2025.
[2] Q. Zeng, F. Li, Z. Zhao, Y. Li, and Y. Wang, "Acouwrite: Acoustic-based handwriting recognition on smartphones," *IEEE Transactions on Mobile Computing*, 2024.
[3] Y. Feng, D. Dai, J. Huang, P. Yang, X.-Y. Li, F. Han, and L. Yang, "Battery-free monitoring of micron-level vibrations with sub-hertz frequency accuracy: Toward robust and accurate industrial sensing," *IEEE Transactions on Mobile Computing*, pp. 1–18, 2025.
[4] H. Wang, K. Fan, H. Chen, Q. Chen, K. Zhang, F. Li, H. Li, and Y. Yang, "Joint biological id: A secure and efficient lightweight biometric authentication scheme," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 3, pp. 2578–2592, 2022.
[5] P. Zhang and L. Jin, "Online writer retrieval with chinese handwritten phrases: A synergistic temporal-frequency representation learning approach," *IEEE Transactions on Information Forensics and Security*, 2024.
[6] P. Zhang, Y. Liu, S. Lai, H. Li, and L. Jin, "Privacy-preserving biometric verification with handwritten random digit string," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.
[7] C. Lin, J. He, C. Shen, Q. Li, and Q. Wang, "Crossbehaauth: Cross-scenario behavioral biometrics authentication using keystroke dynamics," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 3, pp. 2314–2327, 2022.

[8] J. Huang, J.-X. Bai, X. Zhang, Z. Liu, Y. Feng, J. Liu, X. Sun, M. Dong, and M. Li, "Keystrokesniffer: An off-the-shelf smartphone can eavesdrop on your privacy from anywhere," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 6840–6855, 2024.

[9] M. Hu, D. Wang, C. Li, Y. Xu, and B. Tu, "Behavioral biometrics-based continuous authentication using a lightweight latent representation masked one-class autoencoder," *IEEE Transactions on Dependable and Secure Computing*, 2024.

[10] S. Li, L. Fei, B. Zhang, X. Ning, and L. Wu, "Hand-based multimodal biometric fusion: A review," *Information Fusion*, p. 102418, 2024.

[11] J. Huang, B. Liu, C. Miao, X. Zhang, J. Liu, L. Su, Z. Liu, and Y. Gu, "Phyfinatt: An undetectable attack framework against phy layer fingerprint-based wifi authentication," *IEEE Transactions on Mobile Computing*, vol. 23, no. 7, pp. 7753–7770, 2024.

[12] Y. Zhang, W. Sun, and M. Li, "Wirite: General and practical wi-fi based hand-writing recognition," *IEEE Transactions on Mobile Computing*, 2023.

[13] X. Zhang, H. Yan, J. Huang, B. Liu, Y. Feng, J. Liu, M. Li, F. Zhang, and Z. Liu, "Beyond physical labels: Redefining domains for robust wifi-based gesture recognition," *arXiv preprint arXiv:2601.03825*, 2026.

[14] Y. Cao, C. Cai, F. Li, Z. Chen, and J. Luo, "Enabling passive user authentication via heart sounds on in-ear microphones," *IEEE Transactions on Dependable and Secure Computing*, 2024.

[15] C. Lin, Z. Sun, A. Ahmad, X. Fan, Y. Wang, L. Wang, X. Fan, and G. Wu, "Airwrite: An aerial handwriting trajectory tracking and recognition system with mmwave," *IEEE Transactions on Mobile Computing*, 2024.

[16] J. Lei, Q. Pei, Y. Wang, W. Sun, and X. Liu, "Privface: Fast privacy-preserving face authentication with revocable and reusable biometric credentials," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 5, pp. 3101–3112, 2021.

[17] S. Singh, N. K. Garg, and M. Kumar, "Feature extraction and classification techniques for handwritten devanagari text recognition: a survey," *Multimedia Tools and Applications*, vol. 82, no. 1, pp. 747–775, 2023.

[18] C. Fahn, C. Lee, and H. Chen, "A text independent handwriting forgery detection system based on branchlet features and gaussian mixture models," in *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, 2016, pp. 690–697.

[19] S. Ghosh, P. Shivakumara, P. Roy, U. Pal, and T. Lu, "Graphology based handwritten character analysis for human behaviour identification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 55–65, 2020.

[20] V. Kumar and S. Sundaram, "Utilization of information from cnn feature maps for offline word-level writer identification," *Expert Systems with Applications*, vol. 238, p. 121709, 2024.

[21] J. Jiang, S. Lai, L. Jin, Y. Zhu, J. Zhang, and B. Chen, "Forgery-free signature verification with stroke-aware cycle-consistent generative adversarial network," *Neurocomputing*, vol. 507, pp. 345–357, 2022.

[22] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Characterizing and evaluating adversarial examples for offline handwritten signature verification," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 2153–2166, 2019.

[23] Y. Liu, F. B. Khalid, L. Wang, Y. Zhang, and C. Wang, "Elegantly written: Disentangling writer and character styles for enhancing online chinese handwriting," in *European Conference on Computer Vision*. Springer, 2024, pp. 409–425.

[24] A. Ghaffari, M. Kafaee, and V. Abolghasemi, "Smooth non-negative sparse representation for face and handwritten recognition," *Applied Soft Computing*, vol. 111, p. 107723, 2021.

[25] C. Taleb, L. Likforman-Sulem, C. Mokbel, and M. Khachab, "Detection of parkinson's disease from handwriting using deep learning: a comparative study," *Evolutionary intelligence*, pp. 1–12, 2023.

[26] X. Zhang, G. Xie, C. Liu, and Y. Bengio, "End-to-end online writer identification with recurrent neural network," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 2, pp. 285–292, 2017.

[27] S. Lai and L. Jin, "Recurrent adaptation networks for online signature verification," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 6, pp. 1624–1637, 2019.

[28] N. Arab, H. Nemmour, and Y. Chibani, "A new synthetic feature generation scheme based on artificial immune systems for robust offline signature verification," *Expert Systems with Applications*, vol. 213, p. 119306, 2023.

[29] V. K. Melo, B. L. D. Bezerra, D. Impedovo, G. Pirlo, and A. Lundgren, "Deep learning approach to generate offline handwritten signatures based on online samples," *IET Biometrics*, vol. 8, no. 3, pp. 215–220, 2018.

[30] M. Diaz, M. A. Ferrer, D. Impedovo, M. I. Malik, G. Pirlo, and R. Plamondon, "A perspective analysis of handwritten signature technology," *Acm Computing Surveys (Csur)*, vol. 51, no. 6, pp. 1–39, 2019.

[31] R. Tolosana, R. Vera-Rodriguez, C. Gonzalez-Garcia, J. Fierrez, A. Morales, J. Ortega-Garcia, J. C. Ruiz-Garcia, S. Romero-Tapiador, S. Rengifo, M. Caruana *et al.*, "Svc-ongoing: Signature verification competition," *Pattern Recognition*, vol. 127, p. 108609, 2022.

[32] S. K. Singh and A. Chaturvedi, "An efficient multi-modal sensors feature fusion approach for handwritten characters recognition using shapley values and deep autoencoder," *Engineering Applications of Artificial Intelligence*, vol. 138, p. 109225, 2024.

[33] J. Zhang, J. Du, and L. Dai, "Track, attend, and parse (tap): An end-to-end framework for online handwritten mathematical expression recognition," *IEEE Transactions on Multimedia*, vol. 21, no. 1, pp. 221–233, 2018.

[34] V. Pippi, S. Cascianelli, L. Baraldi, and R. Cucchiara, "Evaluating synthetic pre-training for handwriting processing tasks," *Pattern Recognition Letters*, 2023.

[35] L. Nandanwar, P. Shivakumara, H. A. Jalab, R. W. Ibrahim, R. Raghavendra, U. Pal, T. Lu, and M. Blumenstein, "A conformable moments-based deep learning system for forged handwriting detection," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[36] F. A. Khan, F. Khelifi, M. A. Tahir, and A. Bouridane, "Dissimilarity gaussian mixture models for efficient offline handwritten text-independent identification using sift and rootsift descriptors," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 289–303, 2019.

[37] M. Mayr, M. Dreier, F. Kordon, M. Seuret, J. Zöllner, F. Wu, A. Maier, and V. Christlein, "Zero-shot paragraph-level handwriting imitation with latent diffusion models," *International Journal of Computer Vision*, pp. 1–22, 2025.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[39] X. Han, J. Huang, M. Li, C. Cai, and T. Zheng, "Loki: Physical-world adversarial attacks on wireless indoor localization via differentiable object placement," *IEEE Transactions on Information Forensics and Security*, pp. 1–1, 2025.

[40] M. K. Fallah, M. Najafi, S. Gorgin, and J.-A. Lee, "Abstraction and decision fusion architecture for resource-aware image understanding with application on handwriting character classification," *Applied soft computing*, vol. 162, p. 111813, 2024.

[41] W. Zhou, Y. Xiao, W. Yan, and L. Yu, "Cmpffnet: Cross-modal and progressive feature fusion network for rgb-d indoor scene semantic segmentation," *IEEE Transactions on Automation Science and Engineering*, 2023.

[42] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 4685–4694.

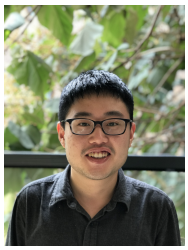
[43] Y. Bai, I. Shahid, H. Takawale, and N. Roy, "Scribe: Simultaneous voice and handwriting interface," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 7, no. 4, pp. 1–31, 2024.

[44] Y. Zhu, S. Li, H. Wang, and F. Wei, "Finet: Handwriting trajectory reconstruction of chinese characters based on the font imitate network," *Pattern Recognition*, vol. 157, p. 110949, 2025.

[45] J. Wang, L. Mou, C. Zheng, and W. Gao, "Image-based freeform handwriting authentication with energy-oriented self-supervised learning," *IEEE Transactions on Multimedia*, 2024.

[46] H. Zhu, M. Xiao, D. Sherman, and M. Li, "Soundlock: A novel user authentication scheme for vr devices using auditory-pupillary response." in *NDSS*, 2023.

[47] S. J. Wang, K. Pei, and J. Yang, "Smartinv: Multimodal learning for smart contract invariant inference," in *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2024, pp. 2217–2235.



Jinyang Huang is an Associate Professor at the School of Computer Science and Information Engineering, Hefei University of Technology (HFUT). His research interests include Multimodal Perception, Human-computer Interaction, Wireless Security, and Signal Processing. In this area, he has published 55 papers in international peer-reviewed journals and conferences, including ToN, TMC, TIFS, TAFFC, THMS, MobiCom, IEEE S&P, USENIX Security, Ubicomp, NeurIPS, Infocom, ACM MM, and ECCV. He has served as a TPC member for conferences, including ACM MM, IEEE ICME, and Globecom, and has the honor of becoming ACM MM 2024 Outstanding Reviewers. He is a Guest Editor for the Technical Committee of ICME 25 Special Session. He is the recipient of the Young Scientist of Anhui Computer Federation and IEEE HITC Distinguished PhD Dissertation Award.

Yuanhao Feng is now a Postdoctoral Fellow in the Department of Computing, The Hong Kong Polytechnic University. He received the Ph.D. degree from the University of Science and Technology of China in 2023. He is an outstanding graduate of the School of Computer Science in USTC 2023. Dr. Feng has long-term cooperation with the research group of Dr. Wang Meng of Hefei University of Technology. He has published many top journal and conference papers, such as MobiCom, INFOCOM, and TMC. His research interest is AIoT, which includes wireless sensing and communication.

Feng-Qi Cui (Student Member, IEEE) is currently pursuing M.E. degree in the Institute of Advanced Technology, University of Science and Technology of China (USTC), Hefei, China. He is also affiliated with Anhui Provincial Key Laboratory of Affective Computing and Advanced Intelligent Machines, Institute of Artificial Intelligence, Hefei Comprehensive National Science Center. His research interests include Computer Vision, Multimedia Information Processing and Affective Computing.

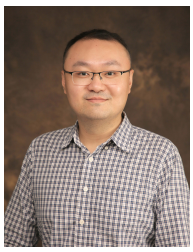
Xiang Zhang received the B.E. and D.E. degrees from Hefei University of Technology, Hefei, China, in 2017 and 2023, respectively. He is currently a tenured Associate Professor with the School of Cyber Security, Tianjin University. His research interests include IOT Security, Wireless Sensing and Cyber-Physical System Security. In these area, he has published 30+ papers in international peer-reviewed journals and conferences, including IEEE S&P, Usenix Security, ACM Mobicom and NeurIPS. Prof. Zhang is the recipient of the IEEE SMC Society Andrew P. Sage Best Transactions Paper Award, the IEEE HITC Distinguished Phd Dissertation Award, and the IEEE ICME 2025 Best Paper Award Nomination. He is a publication chair of IEEE ISPA 2025 and an AEC member of Usenix Security.

Zhi Liu (S'11-M'14-SM'19) received the Ph.D. degree in informatics in National Institute of Informatics. He is currently an Associate Professor at The University of Electro-Communications. His research interest includes video network transmission and mobile edge computing. He is now an editorial board member of Springer wireless networks and IEEE Open Journal of the Computer Society. He is a senior member of IEEE.

Xin Liu received the B.E. degree in Electronic Information Science and Technology from North University of China, China, in 2016. He is currently pursuing the Ph.D. degree in the School of Computer Science and Technology, Hefei University of Technology, China. His research interests include computer vision, deep learning, graph contrastive learning, and wireless sensing.

Jianchun Liu (Member, IEEE/ACM) received the Ph.D. degree in School of Data Science from the University of Science and Technology of China in 2022. He is currently an associate researcher in the School of Computer Science and Technology at University of Science and Technology of China. His main research interests are computer networks, federated learning and edge inference.

Fusang Zhang is a Professor in the School of Cyber Science and Technology at Beihang University. He received the MS and PhD degrees in computer science from the Institute of Software, Chinese Academy of Sciences, Beijing, China, in 2013 and 2017, respectively. He was awarded the ACM SIGMOBILE China Rising Star Award, CAAI Wu Wenjun Artificial Intelligence Excellent Youth Award, AIoTSys Young Scientist Award. His current research interests include mobile and pervasive computing, wireless sensing, mobile health, and quantum sensing. His work appeared at MobiCom, NSDI, UbiComp, Sensys, and CoNEXT. His work won MobiCom '22 Best Paper Award Runner-up, MobiCom '22 Best Community Paper Award Runner-up, UbiComp '21 (IMWUT '20) Distinguished Paper Award, CCF TPCI Best Paper Award'23, HHME PPC'20 Best Paper Award, MobiQutious'17 Best Paper Award Runner-up.



Meng Li is a Professor at the School of Computer Science and Information Engineering, Hefei University of Technology (HFUT), China. He was a Post-Doc Researcher at the Department of Mathematics and HIT Center, University of Padua, Italy, where he is with the Security and PRIvacy Through Zeal (SPRITZ) research group led by Prof. Mauro Conti (IEEE Fellow). He obtained his Ph.D. in Computer Science and Technology from the School of Computer Science and Technology, Beijing Institute of Technology (BIT), China, in 2019. He was sponsored by ERCIM 'Alain Bensoussan' Fellowship Programme (from 2020.10.1 to 2021.3.31) to conduct Post-Doc research supervised by Prof. Fabio Martinelli at CNR, Italy. He was sponsored by China Scholarship Council (CSC) as a Joint Ph.D. student (from 2017.9.1 to 2018.8.31) supervised by Prof. Xiaodong Lin (IEEE Fellow) in the Broadband Communications Research (BBCR) Lab at University of Waterloo and Wilfrid Laurier University, Canada. He is supported by CSC as a Visiting Scholar (from 2025.3.1 to 2025.6.30) collaborating with Prof. Mauro Conti (IEEE Fellow) at the HIT Center, University of Padua, Italy. He was promoted to a Professor from December 18, 2025. His research interests include security, privacy, applied cryptography, blockchain, TEE, and Internet of Vehicles. In this area, he has published 146 papers in topmost journals and conferences, including TIFS, TDSC, ToN, TMC, TKDE, TODS, TPDS, TSE, TSC, COMST, IEEE S&P, USENIX Security, MobiCom, INFOCOM, and ISSTA. He is a Senior Member of IEEE, CIE, CIC, and CCF. He is an Associate Editor for TIFS, TDSC, and TNSM. He has served as a TPC member for conferences, including ICDCS, Inscrypt, ICICS, and TrustCom. He is the recipient of 2024 IEEE HITC Award for Excellence (Early Career Researcher) and 2025 IEEE TCSVC Rising Star Award. He was selected into the IEEE Computer Society Computing's Top 30 Early Career Professionals for 2025.