

TAOSV: Trajectory Reconstruction Assisted Chinese Offline Signature Verification

Shaoxin Ma, Lizhong Zheng, and Xu Zhao*

School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, China
{msx-2020, zhenglz0108, zhaoxu}@sjtu.edu.cn

Abstract. Chinese signatures are difficult to verify from static images alone because dense stroke overlap, connected radicals, and personal abbreviations can hide local order and continuity cues that are useful for distinguishing genuine signatures from forgeries. We propose TAOSV, a trajectory-assisted framework for Chinese offline signature verification that uses reconstructed handwriting trajectory as auxiliary guidance. TAOSV first estimates a pseudo-online handwriting-trajectory sequence from each offline image, converts the sequence into fixed-length order points, and uses these points to sample multi-scale image patches before ViT-based global embedding extraction. On MSDS and ChiSig, TAOSV reduces EER from 24.18% to 19.60% and from 11.64% to 8.89% over the corresponding image-only verifier, respectively, and achieves state-of-the-art performance under the reported protocols. Further ablations show that reconstructed order cues provide complementary verification evidence for Chinese offline signature verification.

Keywords: Offline Signature Verification · Chinese Signatures · Vision Transformer · Handwriting Trajectory Reconstruction

1 Introduction

Offline Chinese signature verification (OSV) determines whether two static signature images were written by the same person. This task is important for document authentication and identity verification, but Chinese signatures are difficult to compare because they often contain compressed character shapes, connected radicals, dense stroke overlap, and highly individualized local writing patterns [7, 12, 16, 33]. Recent OSV systems based on convolutional networks, Siamese learning, and vision transformers learn strong image representations [10, 15]. However, most of these systems process a signature as a regular image grid or as visual tokens without an explicit representation of the order in which local stroke evidence should be read. For Chinese signatures, this missing structure matters because visually similar signatures can differ in stroke continuity, junction placement, breakpoint organization, and locally consistent writing order.

* Corresponding author.

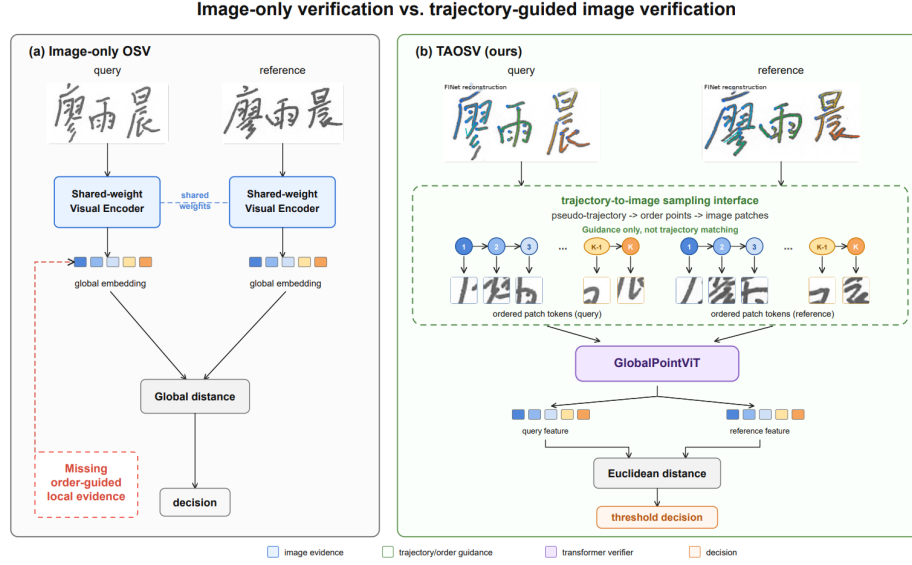


Fig. 1: Image-only verification versus trajectory-guided image verification. Conventional image-only OSV compares global visual embeddings without explicit order-guided local evidence. In contrast, TAOSV reconstructs a pseudo-online trajectory, converts it into order points, and uses these points to guide image patch sampling before global embedding comparison.

Introducing writing order into offline OSV is useful but inherently uncertain. A static signature image does not record pen motion, and different writing orders can produce similar visible shapes [32]. For this reason, this paper does not treat the estimated sequence as the original online path of a signature. Instead, we study a verification-oriented use of reconstructed handwriting structure: an estimated pseudo-online trajectory is treated as an auxiliary cue that helps the verifier decide where to sample image evidence and how to order the resulting visual tokens.

We propose TAOSV, a trajectory-assisted framework for Chinese OSV. Given an offline signature image, a reconstruction front-end estimates a handwriting-sequence represented by normalized coordinates and categorical structural states. The estimated sequence is then converted into fixed-length order points. Rather than matching trajectory directly, the verifier uses these order points to sample multi-scale patches from the original image, combines visual tokens with sequential and spatial positional encodings, and extracts an image-grounded global embedding. The trajectory acts as guidance for reading image evidence rather than a standalone biometric signal. Figure 1 contrasts this design with conventional image-only verification.

The main claim of this paper is that reconstructed order guidance provides an effective verification cue for Chinese offline signature verification. TAOSV

provides a concrete interface between offline-to-online reconstruction and image-based signature verification: reconstruction supplies ordered sampling anchors, and verification tests whether these anchors improve the comparison of static signatures. Experiments on MSDS and ChiSig show that this design improves over image-only verification and achieves state-of-the-art EERs of 19.60% and 8.89%, respectively. Our contributions are threefold:

- We formulate trajectory-assisted Chinese OSV as an image-grounded verification problem in which a reconstructed trajectory provides auxiliary structural guidance, not a ground-truth online path.
- We introduce a reconstruction-to-verification interface that converts a pseudo-online handwriting-structure sequence into fixed-length order points for local visual sampling and positional encoding.
- We provide a verification-centered evaluation with image-only, trajectory-assisted, oracle-trajectory, and interface ablations, demonstrating that reconstructed trajectories improve downstream Chinese offline signature verification on MSDS and ChiSig.

2 Related Work

Offline signature verification. Offline signature verification compares static signature images to determine writer identity consistency. Traditional methods rely on handcrafted contour, texture, graphometric, or local shape descriptors combined with statistical classifiers, whereas modern approaches learn discriminative representations with CNNs, Siamese networks, and metric-learning objectives [6, 10, 18, 28]. Later systems improve image-based verification by modeling sparse stroke cues [27, 29], local deformations [35], discriminative local regions [20], shallow-to-deep visual features [34], and multi-scale global-regional representations [11]. Recent work further explores self-supervised or reconstruction-based pretraining [3, 14, 21], Transformer-based local-global relation modeling [15, 26], cross-attention multi-task interaction [22], and OCR-aware signature verification [16]. Unlike these image-grounded methods, TAOSV introduces estimated handwriting-structure cues to organize where and in what order local image evidence is read.

Handwriting trajectory reconstruction of Chinese characters. Handwriting trajectory reconstruction estimates an ordered writing sequence from an offline handwriting image. Early methods use contour tracing, skeleton analysis, graph traversal, or rule-based stroke ordering to infer plausible pen trajectories from static characters [13, 25]. Recent neural approaches formulate the task as image-to-sequence prediction, where a visual encoder extracts appearance features and a sequence decoder predicts ordered coordinate or state tokens [1, 2, 5, 24]. For Chinese handwriting, this problem is especially relevant because characters contain complex stroke layouts, dense local structures, and multiple plausible writing orders. FINet further shows that ordered point prediction can be learned from offline Chinese character images in a font-imitation setting [36], while recent

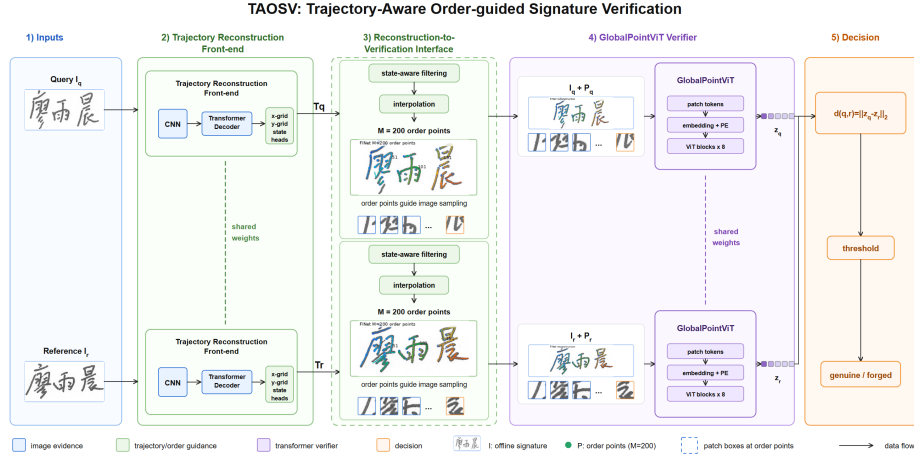


Fig. 2: Overview of TAOSV. A shared trajectory reconstruction front-end estimates pseudo-online sequences for the query and reference signatures. The reconstruction-to-verification interface converts each sequence into order points, which guide multi-scale image patch sampling before the shared GlobalPointViT verifier computes image-grounded embeddings and a final distance.

trajectory Transformer models emphasize global radical context for complex Chinese handwriting trajectory recovery [17]. These studies provide technical foundations for estimating structured handwriting sequences, but their primary objective is reconstruction. In contrast, our work uses reconstruction as an intermediate representation for Chinese OSV, and its value must be tested through downstream verification.

Online cues and offline limitations. Unlike online signature verification, offline OSV has no direct access to temporal order, pen-up transitions, velocity, or pressure [4, 8]. TAOSV therefore treats reconstructed trajectories as auxiliary structural hypotheses that guide image-grounded verification, rather than as true writing paths or standalone matching signals.

3 Method

3.1 Overview

Given a query-reference pair of offline Chinese signature images, offline signature verification determines whether the two images were produced by the same writer. Let I_q and I_r denote the query and reference images, respectively. The proposed framework keeps the offline image as the primary source of verification evidence and uses the reconstructed handwriting trajectory as auxiliary structural guidance. The final verification score is therefore computed from image-grounded embeddings rather than by comparing reconstructed trajectories as standalone signals.

The pipeline contains two modules, as summarized in Figure 2. The trajectory reconstruction front-end first estimates a pseudo-online handwriting-structure sequence from each offline image,

$$\hat{\mathcal{T}} = G_\phi(I) = \{(\hat{x}_t, \hat{y}_t, \hat{s}_t)\}_{t=1}^N, \quad (1)$$

where N denotes the decoded sequence length, $(\hat{x}_t, \hat{y}_t)$ are normalized coordinate hypotheses, and $\hat{s}_t$ is an estimated categorical structural state. The trajectory-assisted verifier then converts the estimated sequence into fixed-length order points, samples local visual evidence from the original image, and produces an image-grounded embedding

$$z = F_\theta(I, \hat{\mathcal{T}}). \quad (2)$$

For a query-reference pair, the verification distance is measured by the Euclidean distance between the image-grounded embeddings of the query and reference signatures, as formally defined later in Eq. (12). A smaller distance indicates stronger evidence that the two signatures belong to the same writer.

3.2 Trajectory Reconstruction as Auxiliary Structure Estimation

Since a static image does not specify a unique writing order, the reconstructed sequence serves as an estimated handwriting-structure cue. Given an offline signature image I , the reconstruction front-end predicts coordinate-state tokens $\hat{u}_t = (\hat{x}_t, \hat{y}_t, \hat{s}_t)$ in an autoregressive manner. A visual encoder maps the image to memory tokens, and a Transformer decoder predicts each token conditioned on the image memory and the previously decoded prefix:

$$p(\hat{u}_t \mid I, \hat{u}_{<t}) = D_\phi(E_\phi(I), \hat{u}_{<t}). \quad (3)$$

Decoding is initialized with a beginning-of-sequence token and terminates when an end-of-sequence token is emitted or when the maximum sequence length is reached.

The output space of the reconstruction front-end consists of normalized image coordinates and categorical structural states. The state vocabulary supports sequence control and stroke organization, including beginning-of-sequence, end-of-sequence, pen-up, pen-down and padding tokens. These states are auxiliary labels and should not be interpreted as directly observed physical pen events. In implementation, the coordinate heads predict discrete image-grid classes for x and y , while the state head predicts a categorical state:

$$o_t = [o_t^x, o_t^y, o_t^s], \quad o_t^x \in \mathbb{R}^W, \quad o_t^y \in \mathbb{R}^H, \quad o_t^s \in \mathbb{R}^K. \quad (4)$$

W and H are coordinate-class counts, equal to the resized image width and height in our grid-classification implementation, K is the number of state classes.

$$\mathcal{L}_{\text{rec}} = \sum_t m_t [\text{CE}(o_t^x, g_t^x) + \text{CE}(o_t^y, g_t^y) + \lambda_s \text{CE}(o_t^s, s_t)], \quad (5)$$

where g_t^x and g_t^y are discretized coordinate targets, s_t is the state target, $m_t \in \{0, 1\}$ masks padded positions, and λ_s controls the relative weight of the state-classification loss.

The resulting sequence is used only as an auxiliary input to the verifier. Reconstruction diagnostics may quantify sequence quality, but they are not sufficient to establish the contribution to OSV. The central question is whether the estimated sequence contributes to image-grounded Chinese offline signature verification under a controlled verification protocol.

3.3 Trajectory Assisted Signature Verification

The verifier uses the estimated handwriting-structure sequence as auxiliary guidance for image-grounded signature comparison. Given a query image I_q , a reference image I_r , and their corresponding order-point sequences P_q and P_r , the verifier computes a pairwise distance that measures whether the two signatures are likely to have been produced by the same writer. The order points are not matched directly. Instead, they specify where visual evidence is sampled from each offline image and define the sequence order in which the sampled tokens are processed before global embedding extraction.

Each estimated trajectory is first converted into a fixed-length ordered point set

$$P = \{(x_i, y_i)\}_{i=1}^M, \quad (6)$$

where M is fixed across signatures. The implementation filters out invalid or special state tokens, interpolates along stroke-continuation segments, preserves stroke-start or move-to points, and uniformly samples the resulting sequence into M order points. Each point is represented by normalized image coordinates. The state channel is used to construct the order map, whereas the minimal verifier uses only the coordinate channels of P as input anchors.

For each order point, the verifier extracts local image evidence at multiple spatial scales. Let $p_i = (x_i, y_i)$ denote an order point in normalized coordinates, and let $\mathcal{P}$ denote the set of patch scales. For each scale $\rho \in \mathcal{P}$, a point-centered patch is extracted from the original signature image by bilinear sampling:

$$X_{i,\rho} = \text{Sample}(I, x_i, y_i, \rho). \quad (7)$$

The sampled patches provide the visual content of the verifier tokens. Thus, the trajectory is not used to produce a separate matching score; it defines the image locations to be sampled and the order in which the resulting visual tokens are presented to the Transformer.

Each scale-specific patch is flattened and projected into a common token space,

$$e_{i,\rho} = \text{Proj}_\rho(X_{i,\rho}), \quad (8)$$

and the projected features are fused across scales by learned image-conditioned weights:

$$v_i = \sum_{\rho \in \mathcal{P}} \alpha_\rho(I) e_{i,\rho}, \quad \alpha_\rho(I) = \sigma(a_\rho(I)). \quad (9)$$

The scalar weight $\alpha_\rho(I) \in (0, 1)$ is produced by a scale-specific linear head from the input image and is shared by all order points of the same image. This construction allows the verifier to combine fine local stroke details with broader neighborhood context while keeping the evidence grounded in the original offline image.

The point-centered tokens are augmented with sequential and spatial positional information before Transformer encoding:

$$h_i = v_i + g_i^{1D} \text{PE}_{1D}(i) + g_i^{2D} \text{PE}_{2D}(x_i, y_i), \quad (10)$$

where PE_{1D} denotes a sinusoidal positional encoding over the point order, and $\text{PE}_{2D}(x_i, y_i)$ denotes a sinusoidal two-dimensional positional encoding sampled at the point coordinate (x_i, y_i) . The scalar gates $g_i^{1D} = \sigma(\beta_{1D})$ and $g_i^{2D} = \sigma(\beta_{2D})$ are learnable parameters constrained to $(0, 1)$, allowing the model to adaptively balance sequential order cues and spatial coordinate cues.

Following the class-token aggregation strategy of Vision Transformers [9], a learnable CLS token is prepended to the ordered point-token sequence, and the resulting sequence is processed by a global Transformer encoder. The output CLS token is used as the global signature embedding $z \in \mathbb{R}^D$:

$$z = F_\theta(I, P)_{\text{cls}}. \quad (11)$$

For a query-reference pair, the verifier computes the global embedding distance

$$d(q, r) = \|z_q - z_r\|_2, \quad (12)$$

where $\|\cdot\|_2$ denotes the Euclidean norm and z_q and z_r are the global embeddings of the query and reference signatures, respectively. A smaller distance indicates stronger evidence that the two signatures belong to the same writer.

3.4 Training and Inference

Training follows the two-stage pipeline. The reconstruction front-end is first trained with the supervised reconstruction loss in Eq. (5), and then supplies estimated trajectories for verifier training. Given a labeled query-reference pair, $y_{qr} = 1$ denotes a genuine pair and $y_{qr} = 0$ denotes a forgery pair. Let $d = d(q, r)$ denote the embedding distance in Eq. (12). The verifier is optimized with the focal contrastive loss used in TransOSV [15]:

$$\begin{aligned} \mathcal{L}_{\text{ver}} = & y_{qr} \sigma(\kappa_+(d - \alpha_+)) [d - m_+]_+^2 \\ & + (1 - y_{qr}) \sigma(\kappa_-(\alpha_- - d)) [m_- - d]_+^2, \end{aligned} \quad (13)$$

where κ_+ and κ_- control the focal weighting sharpness for genuine and forgery pairs, respectively; α_+ and α_- are focal transition distances; and m_+ and m_- are the genuine and forgery margins.

Inference uses only offline signature images. For each query-reference pair, TAOSV computes the verification distance defined in Sec. 3.3. A binary decision is then obtained by thresholding this distance:

$$\hat{y}_{qr} = \mathbb{I}[d(q, r) \leq \tau], \quad (14)$$

where τ is selected on the validation set and kept fixed for testing. Thus, online trajectories are not required at test time.

4 Experiments

4.1 Experimental Setup

Datasets for Reconstruction. The reconstruction network is trained with complementary handwriting and signature data. CASIA-HWDB1.1 [19] is used for character-structure pretraining, where online trajectories are converted into ordered coordinate-state sequences and arranged into synthetic signature-like layouts. For signature-domain reconstruction training and adaptation on MSDS [31] and ChiSig [30], we use only the training writers defined by the downstream verification protocol. MSDS is used because it provides paired offline signature images and online trajectory records. For ChiSig, pseudo-trajectories are generated by InkSight [23] only for the training split used for adaptation, followed by manual quality control to remove incomplete, highly jittery, or structurally invalid reconstructions. No validation or test writers from MSDS or ChiSig are used to train or adapt the reconstruction front-end.

Datasets for Verification. We evaluate signature verification on MSDS and ChiSig using writer-disjoint protocols. For MSDS, writers 0–287 are used for training, writers 288–319 are used for validation and threshold selection, and writers 320–401 are used for testing. This results in 5,760 genuine and 5,760 forgery pairs for training, 640 genuine and 640 forgery pairs for validation, and 15,580 genuine and 15,580 forgery pairs for testing. For ChiSig, we construct a writer-disjoint train/validation/test protocol with 311, 44, and 89 writers, respectively. It contains 2,149 genuine and 2,149 forgery pairs for training, 387 genuine and 387 forgery pairs for validation, and 799 genuine and 799 forgery pairs for testing. The validation and test writers are not used to train either the reconstruction front-end or the verifier. In both datasets, the decision threshold is selected only on the validation set, and the final test writers are disjoint from both training and validation writers. Unless otherwise specified, the verifier receives only offline signature images at test time. Ground-truth online trajectories are used only in the MSDS oracle variant to estimate an upper bound. During verifier training, each mini-batch samples anchor signatures from the training writers and pairs each anchor with one genuine sample and one forgery sample, which keeps the batch balanced and stabilizes the focal contrastive loss.

Implementation Details. The reconstruction model follows the architecture described in Sec. 3. W is 256, H is 128, λ_s is 1.5, the maximum sequence length is 512, and the reconstruction network uses a CNN spatial encoder, a 2-layer Transformer encoder, a 4-layer Transformer decoder, 4 attention heads, and a 256-dimensional embedding. The verifier uses 200 order points, 17×17 and 33×33 point-centered patches, 768-dimensional patch tokens, and the first 10 layers of google/vit-base-patch16-224 [9]. It is trained with $\mathcal{L}_{\text{ver}}$ using $\alpha_+ = 0.6$,

Table 1: MSDS reconstruction diagnostics. Lower is better.

Method	DTW ↓	LDTW ↓	RMSE ↓
InkSight [23]	2079.80	6.02	84.70
TAOSV (ours)	562.08	2.58	30.68

$\alpha_- = 0.8$, $m_+ = 0.3$, $m_- = 0.9$, and $\kappa_+ = \kappa_- = 10$, for 20 epochs and optimized using AdamW with a learning rate of $5 * 10^{-5}$, a weight decay of 10^{-2} , and a batch size of 16.

All architectural hyperparameters are selected exclusively on the validation writers, and test writers are used only for final reporting. Ablation tables are reported as post-hoc analyses and are not used to tune the final model.

4.2 Handwriting Reconstruction Analysis

We first diagnose whether the reconstruction front-end provides a usable pseudo-online handwriting-structure sequence before using it for verification. On MSDS, each offline signature has a recorded online trajectory, which allows direct reconstruction evaluation. Following the signature reconstruction literature, we report dynamic time warping (DTW), length-normalized DTW (LDTW), and root mean square error (RMSE). These metrics are used only as reconstruction diagnostics, lower values indicate closer agreement with the recorded sequence, but the downstream verification claim is tested separately in the following comparison and ablation sections.

Table 1 compares our reconstruction front-end with InkSight [23] on the MSDS diagnostic set. TAOSV reduces DTW by 73.0%, LDTW by 57.1%, and RMSE by 63.8% relative to InkSight. This result indicates that the specialized reconstruction front-end is better aligned with the recorded signature trajectories under the same diagnostic metrics, providing a stronger structural cue for the order-point interface used by the verifier.

Qualitative examples in Figure 3 show the role of this reconstruction in the full pipeline. The reconstructed trajectories follow the visible stroke layout and provide ordered anchors over both MSDS and ChiSig signatures, including cleaner digital trajectories and noisier scanned signatures. Mismatches remain around ambiguous junctions and long pen-up transitions, so the trajectory is not used as a standalone biometric signal. Instead, it supplies auxiliary structural guidance: the verifier still reads evidence from the original offline image, but the reconstructed order points define where local patches are sampled and how the patches are ordered.

4.3 Verification Results

The main comparison tests whether TAOSV improves over strong offline signature verification methods. All baselines are reproduced under the same writer-

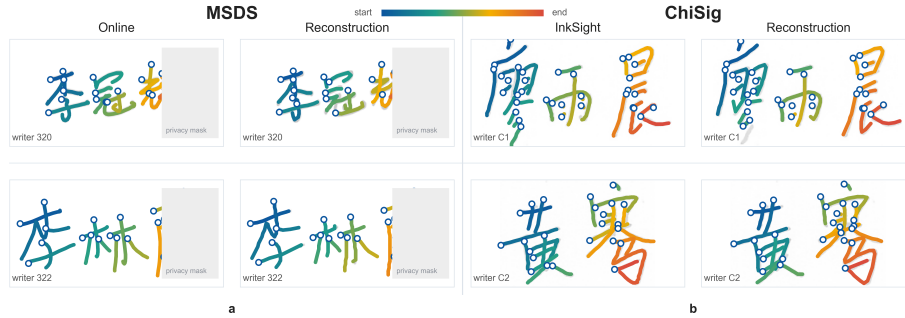


Fig. 3: Qualitative trajectory reconstruction examples on MSDS and ChiSig. (a) MSDS examples comparing ground-truth online trajectories with reconstructed trajectories from offline signature images. The third character is masked to comply with the MSDS privacy policy. (b) ChiSig examples comparing InkSight-generated pseudo-online trajectories with our reconstructed trajectories. In all panels, the faint gray background denotes the offline signature image, colored curves indicate the temporal writing order from start to end, and open circles mark stroke-start points.

Table 2: Comparison with state-of-the-art methods on MSDS and ChiSig datasets.

Method	MSDS (%)				ChiSig (%)			
	Acc $\uparrow$	EER $\downarrow$	FAR $\downarrow$	FRR $\downarrow$	Acc $\uparrow$	EER $\downarrow$	FAR $\downarrow$	FRR $\downarrow$
TAOSV	80.18	19.60	20.27	18.88	90.11	8.89	14.64	5.13
SDINet	79.13	20.49	16.85	24.90	89.83	9.67	9.17	11.14
2C2S	77.52	21.80	15.42	29.54	78.32	21.43	18.48	24.78
DetailSemNet	76.51	23.44	25.14	<u>21.84</u>	86.40	12.78	<u>11.24</u>	15.89
TransOSV	76.36	24.39	22.19	26.69	83.29	16.77	19.27	14.14
MS-SigNet	74.31	25.53	28.12	23.27	68.85	30.51	26.74	35.42
IDN	70.49	29.52	35.13	23.88	66.88	33.19	34.63	31.66
SigNet	67.89	32.12	33.92	30.31	83.53	17.23	22.61	<u>10.51</u>

disjoint split, pair construction, validation-threshold rule, and input preprocessing. We evaluate the proposed method on MSDS and ChiSig and report the same cross-dataset configuration, i.e., $M = 200$ order points, patch scales of 17 and 33, and 10 ViT layers.

On MSDS, TAOSV achieves the best overall verification performance under the reported protocol. It obtains the highest accuracy (80.18%), the lowest EER (19.60%), and the lowest FRR (18.88%), improving over the strongest baseline SDINet on the main accuracy–error balance. Its FAR is not the lowest, since 2C2S and SDINet make more conservative decisions and therefore produce fewer false acceptances. This indicates that the advantage of TAOSV is not uniform dominance over every error type, but a better balanced operating point. In particular, reconstructed trajectory guidance improves equal-error behavior and reduces false rejections, while FAR remains sensitive to the decision threshold.



Fig. 4: Qualitative examples showing how reconstructed trajectory guidance improves offline signature verification. Each row presents a reference–query pair, reconstructed pseudo-online trajectories, and verification distances from the image-only baseline and TAOSV. GT denotes the ground-truth label, d the embedding distance, and τ the validation-selected threshold. Green and red bars indicate genuine and forgery predictions, respectively.

Table 3: Information-source ablation.

Method	MSDS					ChiSig				
	Acc ↑	EER ↓	AUC ↑	FAR ↓	FRR ↓	Acc ↑	EER ↓	AUC ↑	FAR ↓	FRR ↓
Image-only	76.95	24.18	0.8372	21.11	27.13	88.61	11.64	0.9554	8.51	14.27
Reconstructed	80.18	19.60	0.8889	20.27	18.88	90.11	8.89	0.9655	14.64	5.13
Oracle	82.02	18.39	0.9014	17.37	19.26	—	—	—	—	—

The same trend is also observed on ChiSig under the identical cross-dataset configuration. TAOSV achieves the highest accuracy (90.11%), the lowest EER (8.89%), and the lowest FRR (5.13%). Although SDINet still obtains a lower FAR, TAOSV provides a better overall verification balance without dataset-specific tuning. These results show that reconstructed trajectory-guided sampling provides complementary features beyond the regular offline image representation, and that the benefit is consistent across two Chinese signature datasets.

4.4 Ablation Studies

The ablations test whether the gains of TAOSV come from trajectory-assisted verification rather than incidental hyperparameter choices. We first separate the information sources by comparing an image-only verifier, a verifier assisted by reconstructed trajectories, and an MSDS oracle variant that uses recorded online trajectories. This comparison asks whether reconstructed trajectories add verification evidence beyond the offline image, and whether more accurate trajectories leave further headroom. We then ablate the reconstruction-to-verification interface by varying the number of order points, the Transformer depth, the patch scales, and the 1D ordering positional encoding.

Table 4: Number of order points ablation.

Points	MSDS					ChiSig				
	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$
100	<u>80.18</u>	20.39	0.8828	18.61	22.36	88.80	11.14	0.9554	13.27	9.14
200	<u>80.18</u>	<u>19.60</u>	<u>0.8889</u>	<u>20.27</u>	<u>18.88</u>	90.11	8.89	0.9655	14.64	5.13
300	80.22	19.47	0.8892	20.30	18.69	90.36	<u>9.57</u>	<u>0.9654</u>	<u>12.89</u>	<u>6.38</u>
400	78.86	20.62	0.8755	22.05	19.22	<u>90.24</u>	9.70	0.9644	12.14	7.38

Information source. This ablation asks whether order information is a real source of verification evidence rather than a different parameterization of the same image. The image-only baseline removes reconstructed order guidance, while the reconstructed-order-guided variant uses the configuration throughout, i.e., $M = 200$ order points, patch scales of 17 and 33, and 10 ViT layers. Adding the reconstructed trajectory improves MSDS accuracy from 76.95% to 80.18% and reduces EER from 24.18% to 19.60%; on ChiSig, it improves accuracy from 88.61% to 90.11% and reduces EER from 11.64% to 8.89%. The oracle row is only an MSDS upper-bound comparison because it replaces the predicted trajectory with the recorded online trajectory, reducing MSDS EER further to 18.39%. ChiSig has no recorded online trajectories, so no oracle claim is made for that dataset.

Number of order points. The final $M = 200$ setting is selected on the validation writers. Table 4 reports a post-hoc sensitivity analysis on the test writers. The results show that $M = 200$ provides balanced cross-dataset behavior, whereas increasing the number of order points to $M = 300$ mainly improves MSDS but degrades ChiSig. This suggests that denser order-point sampling is not uniformly beneficial and may increase sensitivity to dataset-specific stroke density or reconstruction artifacts.

Depth of ViT layers. The 10-layer ViT is the validation-selected encoder used in the main experiments. Table 5 examines the effect of encoder depth after fixing the remaining configuration. The post-hoc results show that increasing the depth to 12 layers does not yield consistent cross-dataset gains and can degrade ChiSig performance. This indicates that additional Transformer capacity alone does not necessarily improve the use of reconstructed order cues and may increase dataset-specific sensitivity.

Patch scale fusion. The 17/33 patch-scale pair is selected on the validation writers and used throughout the main experiments. Table 6 analyzes test-set sensitivity to the neighborhood scale around each order point. The post-hoc results show that combining a small local patch with a moderately larger contextual patch yields more stable cross-dataset behavior than using a single neighborhood or a substantially larger context. This supports the design intuition that each order point should encode both fine stroke evidence and limited surrounding shape context.

Table 5: ViT-depth ablation.

Layers	MSDS					ChiSig				
	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$
4	<u>80.35</u>	20.00	0.8861	18.95	21.14	89.80	9.89	0.9593	12.02	<u>8.39</u>
6	80.01	20.01	0.8865	19.95	20.08	<u>90.30</u>	9.51	0.9632	10.64	8.76
8	80.10	19.73	0.8878	20.21	<u>19.26</u>	90.49	<u>9.20</u>	0.9673	<u>10.51</u>	8.51
10	80.18	<u>19.60</u>	0.8889	20.27	18.88	90.11	8.89	<u>0.9655</u>	14.64	5.13
12	80.65	19.38	0.8914	<u>19.29</u>	19.48	87.05	11.83	0.9538	8.76	17.15

Table 6: Patch-scale ablation.

Patch scale			MSDS					ChiSig				
17	33	49	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$
✓	–	–	79.34	20.58	0.8791	20.81	20.34	88.05	10.89	0.9554	16.02	7.88
✓	✓	–	<u>80.18</u>	19.60	0.8889	20.27	18.88	90.11	8.89	0.9655	14.64	5.13
✓	✓	✓	79.49	20.95	0.8787	<u>19.83</u>	21.95	86.80	12.89	0.9392	18.77	7.63
–	✓	–	80.36	<u>19.75</u>	<u>0.8871</u>	19.48	19.97	<u>89.61</u>	<u>10.51</u>	<u>0.9618</u>	9.14	11.64
–	✓	✓	79.18	20.52	0.8799	21.39	<u>19.61</u>	88.17	11.39	0.9470	<u>10.01</u>	13.64
–	–	✓	79.25	20.62	0.8800	20.99	20.26	87.80	11.83	0.9508	14.39	10.01

1D ordering positional encoding. This ablation isolates whether the sequential index of each order point contributes beyond PE_{2D} . The variant without PE_{1D} keeps the same order points and PE_{2D} but removes the PE_{1D} term in Eq. 10. Table 7 shows that removing the PE_{1D} term degrades the main operating balance: accuracy, EER, AUC, and FRR are worse on both datasets, with a larger EER/FRR gap on ChiSig. The variant has slightly lower FAR, so the supported conclusion is not that PE_{1D} improves every metric, but that it improves the overall verification balance and especially the ChiSig generalization behavior.

4.5 Runtime Analysis

We measure runtime to separate the cost of pairwise verification from the cost of generating the reconstructed trajectory used by TAOSV. Table 8 reports verification latency for all compared methods and additionally reports the reconstruction latency of TAOSV per signature image. The verification rows are measured per query-reference comparison, while the reconstruction row is measured per single offline signature image.

TAOSV has a verification latency of 22.211 ms per pair. This is slower than lightweight CNN-style baselines such as SigNet, SDINet, IDN, and MS-SigNet,

Table 7: Ablation of PE_{1D} with PE_{2D} retained.

Variant	MSDS					ChiSig				
	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$	Acc $\uparrow$	EER $\downarrow$	AUC $\uparrow$	FAR $\downarrow$	FRR $\downarrow$
Full model	80.18	19.60	0.8889	<u>20.27</u>	18.88	90.11	8.89	0.9655	<u>14.64</u>	5.13
w/o PE_{1D}	<u>79.83</u>	20.13	<u>0.8848</u>	20.25	20.00	<u>89.30</u>	<u>10.39</u>	<u>0.9638</u>	12.02	9.39

Table 8: Runtime comparison. All values are in ms.

Method	Stage	Mean ↓	Std	Median	Min	Max
SigNet	Verification	1.581	0.016	1.574	1.561	1.640
SDINet	Verification	4.230	0.062	4.212	4.188	4.580
IDN	Verification	6.223	0.123	6.187	6.146	6.766
MS-SigNet	Verification	6.934	0.025	6.930	6.894	6.986
TransOSV	Verification	18.616	0.291	18.513	18.421	19.817
TAOSV (ours)	Verification	22.211	0.587	21.965	21.885	24.223
2C2S	Verification	25.846	0.590	25.667	25.578	28.793
DetailSemNet	Verification	48.546	1.009	48.189	48.027	52.936
TAOSV (ours)	Reconstruction	1436.314	280.376	1350.735	1053.111	2032.444

and also slower than the reproduced TransOSV baseline, but it remains faster than heavier image-based baselines such as 2C2S and DetailSemNet. The main runtime bottleneck is therefore not the pairwise verifier but the reconstruction front-end: the mean reconstruction latency is about $64.7\times$ the mean TAOSV verification latency. If reconstructed trajectories are cached, deployment uses the 22.211 ms pairwise verification stage after preprocessing. The online cold pair latency is approximately 2894.8 ms. Thus, TAOSV trades additional preprocessing cost for stronger trajectory-guided verification, and reducing reconstruction latency is an important direction for practical deployment.

5 Discussion and Conclusion

TAOSV suggests that reconstructed handwriting trajectories can provide useful structural guidance for offline Chinese signature verification when they remain grounded in the original image. Because a static signature image does not uniquely determine pen motion, the reconstructed pseudo-online trajectory is best understood as an auxiliary hypothesis rather than as the true writing path. In TAOSV, this trajectory guides how the verifier reads local image evidence, helping it use stroke continuity, junctions, pen-up transitions, and stroke endings more explicitly than a regular image-grid representation. This interpretation also defines the method’s main boundary: ambiguous crossings, long pen-up transitions, heavy scanning noise, and unfamiliar writing styles may reduce trajectory reliability. The reconstruction stage further introduces non-negligible preprocessing cost, making faster and uncertainty-aware trajectory reconstruction an important direction for practical deployment.

In conclusion, this paper revisits Chinese offline signature verification from a handwriting-trajectory perspective and proposes TAOSV, which uses reconstructed trajectories as image-grounded guidance rather than standalone matching signals. Experiments on MSDS and ChiSig show consistent gains over image-only verification, reducing EER from 24.18% to 19.60% and from 11.64% to 8.89%, respectively. On MSDS, oracle trajectories further reduce EER to 18.39%, indicating additional potential from improved reconstruction. Future work will focus on faster reconstruction, uncertainty-aware guidance, and evaluation on noisier scans and broader writing systems.

References

1. Archibald, T., Poggemann, M., Chan, A., Martinez, T.: TRACE: A differentiable approach to line-level stroke recovery for offline handwritten text. In: Lladós, J., Lopresti, D., Uchida, S. (eds.) 16th International Conference on Document Analysis and Recognition, ICDAR 2021, Lausanne, Switzerland, September 5-10, 2021, Proceedings, Part III. Lecture Notes in Computer Science, vol. 12823, pp. 414–429. Springer (2021). https://doi.org/10.1007/978-3-030-86334-0_27, https://doi.org/10.1007/978-3-030-86334-0_27
2. Bhunia, A.K., Bhowmick, A., Bhunia, A.K., Konwer, A., Banerjee, P., Roy, P.P., Pal, U.: Handwriting trajectory recovery using end-to-end deep encoder-decoder network. In: 24th International Conference on Pattern Recognition, ICPR 2018, Beijing, China, August 20-24, 2018. pp. 3639–3644. IEEE Computer Society (2018). <https://doi.org/10.1109/ICPR.2018.8546093>, <https://doi.org/10.1109/ICPR.2018.8546093>
3. Chattopadhyay, S., Manna, S., Bhattacharya, S., Pal, U.: SURDS: self-supervised attention-guided reconstruction and dual triplet loss for writer independent offline signature verification. In: 26th International Conference on Pattern Recognition, ICPR 2022, Montreal, QC, Canada, August 21-25, 2022. pp. 1600–1606. IEEE (2022). <https://doi.org/10.1109/ICPR56361.2022.9956442>, <https://doi.org/10.1109/ICPR56361.2022.9956442>
4. Chauhan, M., Satbhai, A., Hashemi, M.A., Ali, M.B., Ramamurthy, B., Gao, M., Lyu, S., Srihari, S.N.: Vision-language model based handwriting verification. CoRR **abs/2407.21788** (2024). <https://doi.org/10.48550/ARXIV.2407.21788>, <https://doi.org/10.48550/ARXIV.2407.21788>
5. Chen, Z., Yang, D., Liang, J., Liu, X., Wang, Y., Peng, Z., Huang, S.: Complex handwriting trajectory recovery: Evaluation metrics and algorithm. In: Wang, L., Gall, J., Chin, T., Sato, I., Chellappa, R. (eds.) Computer Vision - ACCV 2022 - 16th Asian Conference on Computer Vision, Macao, China, December 4-8, 2022, Proceedings, Part II. Lecture Notes in Computer Science, vol. 13842, pp. 58–74. Springer (2022). https://doi.org/10.1007/978-3-031-26284-5_4, https://doi.org/10.1007/978-3-031-26284-5_4
6. Dey, S., Dutta, A., Toledo, J.I., Ghosh, S.K., Lladós, J., Pal, U.: Signet: Convolutional siamese network for writer independent offline signature verification. CoRR **abs/1707.02131** (2017), <http://arxiv.org/abs/1707.02131>
7. Díaz, M., Ferrer, M.A., Vessio, G.: Explainable offline automatic signature verifier to support forensic handwriting examiners. Neural Comput. Appl. **36**(5), 2411–2427 (2024). <https://doi.org/10.1007/S00521-023-09192-7>, <https://doi.org/10.1007/s00521-023-09192-7>
8. Díaz, M., Mendoza-García, A., Ferrer, M.A., Sabourin, R.: A survey of handwriting synthesis from 2019 to 2024: A comprehensive review. Pattern Recognit. **162**, 111357 (2025). <https://doi.org/10.1016/J.PATCOG.2025.111357>, <https://doi.org/10.1016/j.patcog.2025.111357>
9. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N.: An image is worth 16x16 words: Transformers for image recognition at scale. In: 9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021. OpenReview.net (2021), <https://openreview.net/forum?id=YicbFdNTTy>

10. Hafemann, L.G., Sabourin, R., Oliveira, L.S.: Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognit.* **70**, 163–176 (2017). <https://doi.org/10.1016/J.PATCOG.2017.05.012>, <https://doi.org/10.1016/j.patcog.2017.05.012>
11. Huang, F., Lu, H.: Multiscale global and regional feature learning using co-tuplet loss for offline handwritten signature verification. *CoRR* **abs/2308.00428** (2023). <https://doi.org/10.48550/ARXIV.2308.00428>, <https://doi.org/10.48550/arXiv.2308.00428>
12. Ji, L., Wang, H., Hou, J., Chen, Z., Li, Z.: Signature authenticity verification using a cross-path four-stream network for preventing disguising frauds. *Comput. Electr. Eng.* **122**, 109998 (2025). <https://doi.org/10.1016/J.COMPELECENG.2024.109998>, <https://doi.org/10.1016/j.compeleceng.2024.109998>
13. Lee, S., Pan, J.C.: Offline tracing and representation of signatures. *IEEE Trans. Syst. Man Cybern.* **22**(4), 755–771 (1992). <https://doi.org/10.1109/21.156588>, <https://doi.org/10.1109/21.156588>
14. Li, H., Wei, P., Hu, P.: Static-dynamic interaction networks for offline signature verification. In: Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021. pp. 1893–1901. AAAI Press (2021). <https://doi.org/10.1609/AAAI.V35I3.16284>, <https://doi.org/10.1609/aaai.v35i3.16284>
15. Li, H., Wei, P., Ma, Z., Li, C., Zheng, N.: Transosv: Offline signature verification with transformers. *Pattern Recognit.* **145**, 109882 (2024). <https://doi.org/10.1016/J.PATCOG.2023.109882>, <https://doi.org/10.1016/j.patcog.2023.109882>
16. Lin, C., Wang, D., Su, Y., Ge, D., Zhang, X.: OCR4HSV: A multi-task learning approach for handwritten signature verification. In: Antonacopoulos, A., Chaudhuri, S., Chellappa, R., Liu, C., Bhattacharya, S., Pal, U. (eds.) *Pattern Recognition - 27th International Conference, ICPR 2024, Kolkata, India, December 1-5, 2024, Proceedings, Part XXXI. Lecture Notes in Computer Science*, vol. 15331, pp. 287–302. Springer (2024). https://doi.org/10.1007/978-3-031-78119-3_20, https://doi.org/10.1007/978-3-031-78119-3_20
17. Lin, J., Chen, Z., Liang, L., Peng, W., Huang, S.: Handwriting trajectory recovery via trajectory transformer with global radical context-aware module. In: Antonacopoulos, A., Chaudhuri, S., Chellappa, R., Liu, C., Bhattacharya, S., Pal, U. (eds.) *Pattern Recognition - 27th International Conference, ICPR 2024, Kolkata, India, December 1-5, 2024, Proceedings, Part XX. Lecture Notes in Computer Science*, vol. 15320, pp. 182–195. Springer (2024). https://doi.org/10.1007/978-3-031-78498-9_13, https://doi.org/10.1007/978-3-031-78498-9_13
18. Ling, L., Zhang, H., Yin, F., Liu, C.: Deep metric learning with cross-writer attention for offline signature verification. In: Smith, E.H.B., Liwicki, M., Peng, L. (eds.) *Document Analysis and Recognition - ICDAR 2024 - 18th International Conference, Athens, Greece, August 30 - September 4, 2024, Proceedings, Part II. Lecture Notes in Computer Science*, vol. 14805, pp. 250–267. Springer (2024). https://doi.org/10.1007/978-3-031-70536-6_15, https://doi.org/10.1007/978-3-031-70536-6_15
19. Liu, C., Yin, F., Wang, D., Wang, Q.: CASIA online and offline chinese handwriting databases. In: 2011 International Conference on Document Analysis and Recognition, ICDAR 2011, Beijing, China, September 18-21, 2011. pp. 37–41.

- IEEE Computer Society (2011). <https://doi.org/10.1109/ICDAR.2011.17>, <https://doi.org/10.1109/ICDAR.2011.17>
20. Lu, X., Huang, L., Yin, F.: Cut and compare: End-to-end offline signature verification network. In: 25th International Conference on Pattern Recognition, ICPR 2020, Virtual Event / Milan, Italy, January 10-15, 2021. pp. 3589–3596. IEEE (2020). <https://doi.org/10.1109/ICPR48806.2021.9412377>, <https://doi.org/10.1109/ICPR48806.2021.9412377>
 21. Manna, S., Chattopadhyay, S., Bhattacharya, S., Pal, U.: SWIS: self-supervised representation learning for writer independent offline signature verification. In: 2022 IEEE International Conference on Image Processing, ICIP 2022, Bordeaux, France, 16-19 October 2022. pp. 1411–1415. IEEE (2022). <https://doi.org/10.1109/ICIP46576.2022.9897562>, <https://doi.org/10.1109/ICIP46576.2022.9897562>
 22. Meng, H., Lin, X., Ubul, K., Aysa, A.: Multi-task interaction network based on a cross-attention fusion mechanism for offline signature verification. In: Antonacopoulos, A., Chaudhuri, S., Chellappa, R., Liu, C., Bhattacharya, S., Pal, U. (eds.) Pattern Recognition - 27th International Conference, ICPR 2024, Kolkata, India, December 1-5, 2024, Proceedings, Part XXV. Lecture Notes in Computer Science, vol. 15325, pp. 17–31. Springer (2024). https://doi.org/10.1007/978-3-031-78389-0_2, https://doi.org/10.1007/978-3-031-78389-0_2
 23. Mitrevski, B., Rak, A., Schnitzler, J., Li, C., Maksai, A., Berent, J., Musat, C.C.: Inksight: Offline-to-online handwriting conversion by teaching vision-language models to read and write. *Trans. Mach. Learn. Res.* **2025** (2025), <https://openreview.net/forum?id=pSyUfV5BqA>
 24. Nguyen, H.T., Nakamura, T., Nguyen, C.T., Nakawaga, M.: Online trajectory recovery from offline handwritten japanese kanji characters of multiple strokes. In: 25th International Conference on Pattern Recognition, ICPR 2020, Virtual Event / Milan, Italy, January 10-15, 2021. pp. 8320–8327. IEEE (2020). <https://doi.org/10.1109/ICPR48806.2021.9413294>, <https://doi.org/10.1109/ICPR48806.2021.9413294>
 25. Plamondon, R., Privitera, C.M.: The segmentation of cursive handwriting: an approach based on off-line recovery of the motor-temporal information. *IEEE Trans. Image Process.* **8**(1), 80–91 (1999). <https://doi.org/10.1109/83.736691>, <https://doi.org/10.1109/83.736691>
 26. Shih, M., Huang, T., Shih, Y., Shuai, H., Liu, H., Yeh, Y., Huang, C.: Detailsemnet: Elevating signature verification through detail-semantic integration. In: Leonardis, A., Ricci, E., Roth, S., Russakovsky, O., Sattler, T., Varol, G. (eds.) Computer Vision - ECCV 2024 - 18th European Conference, Milan, Italy, September 29-October 4, 2024, Proceedings, Part XXV. Lecture Notes in Computer Science, vol. 15083, pp. 449–466. Springer (2024). https://doi.org/10.1007/978-3-031-72698-9_26, https://doi.org/10.1007/978-3-031-72698-9_26
 27. Wei, P., Li, H., Hu, P.: Inverse discriminative networks for handwritten signature verification. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019. pp. 5764–5772. Computer Vision Foundation / IEEE (2019). <https://doi.org/10.1109/CVPR.2019.00591>, http://openaccess.thecvf.com/content_CVPR_2019/html/Wei_Inverse_Discriminative_Networks_for_Handwritten_Signature_Verification_CVPR_2019_paper.html

28. Xiao, W., Ding, Y.: A two-stage siamese network model for offline handwritten signature verification. *Symmetry* **14**(6), 1216 (2022). <https://doi.org/10.3390/SYM14061216>, <https://doi.org/10.3390/sym14061216>
29. Xiong, Y., Cheng, S., Ren, J., Zhang, Y.: Attention-based multiple siamese networks with primary representation guiding for offline signature verification. *Int. J. Document Anal. Recognit.* **27**(2), 195–208 (2024). <https://doi.org/10.1007/S10032-023-00455-6>, <https://doi.org/10.1007/s10032-023-00455-6>
30. Yan, K., Zhang, Y., Tang, H., Ren, C., Zhang, J., Wang, G., Wang, H.: Signature detection, restoration, and verification: A novel chinese document signature forgery detection benchmark. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2022, New Orleans, LA, USA, June 19–20, 2022*. pp. 5159–5168. IEEE (2022). <https://doi.org/10.1109/CVPRW56347.2022.00564>, <https://doi.org/10.1109/CVPRW56347.2022.00564>
31. Zhang, P., Jiang, J., Liu, Y., Jin, L.: MSDS: A large-scale chinese signature and token digit string dataset for handwriting verification. In: Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., Oh, A. (eds.) *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022* (2022), http://papers.nips.cc/paper_files/paper/2022/hash/eca896a8cf6363c9573a701c8c5c9cc5-Abstract-Datasets_and_Benchmarks.html
32. Zhang, P., Jin, L.: Online writer retrieval with chinese handwritten phrases: A synergistic temporal-frequency representation learning approach. *IEEE Trans. Inf. Forensics Secur.* **19**, 10387–10399 (2024). <https://doi.org/10.1109/TIFS.2024.3493594>, <https://doi.org/10.1109/TIFS.2024.3493594>
33. Zhang, P., Liu, Y., Lai, S., Li, H., Jin, L.: Privacy-preserving biometric verification with handwritten random digit string. *IEEE Trans. Pattern Anal. Mach. Intell.* **47**(4), 3049–3066 (2025). <https://doi.org/10.1109/TPAMI.2025.3529022>, <https://doi.org/10.1109/TPAMI.2025.3529022>
34. Zhang, Y., Wang, J., Chen, Z., Li, Z.: Cross layer weakly supervised data augmentation network for offline signature verification. In: *International Joint Conference on Neural Networks, IJCNN 2024, Yokohama, Japan, June 30 - July 5, 2024*. pp. 1–8. IEEE (2024). <https://doi.org/10.1109/IJCNN60899.2024.10650739>, <https://doi.org/10.1109/IJCNN60899.2024.10650739>
35. Zheng, Y., Iwana, B.K., Malik, M.I., Ahmed, S., Ohyama, W., Uchida, S.: Learning the micro deformations by max-pooling for offline signature verification. *Pattern Recognit.* **118**, 108008 (2021). <https://doi.org/10.1016/J.PATCOG.2021.108008>, <https://doi.org/10.1016/j.patcog.2021.108008>
36. Zhu, Y., Li, S., Wang, H., Wei, F.: Finet: Handwriting trajectory reconstruction of chinese characters based on the font imitate network. *Pattern Recognit.* **157**, 110949 (2025). <https://doi.org/10.1016/J.PATCOG.2024.110949>, <https://doi.org/10.1016/j.patcog.2024.110949>