DetailSemNet: Elevating Signature Verification through Detail-Semantic Integration (Supplementary Material)

Meng-Cheng Shih[1], Tsai-Ling Huang[1], Yu-Heng Shih[1], Hong-Han Shuai[1], Hsuan-Tung Liu[2], Yi-Ren Yeh[3], and Ching-Chun Huang[1]

[1] National Yang Ming Chiao Tung University, Taiwan
{mcshih.ee11,christina.ii12,ra890927.cs12,hhshuai,chingchun}@nycu.edu.tw
[2] E.SUN Financial Holding Co., Ltd, Taiwan ahare-18342@esunbank.com.tw
[3] National Kaohsiung Normal University, Taiwan yryeh@nknu.edu.tw

1 More Comparison Results

In addition to the related works listed in Tab. 1 and Tab. 2, we extend our comparison by including several state-of-the-art (SOTA) methods, namely SET [21], SWIS [17], SigCNN [7], and Consensus [1], using the BHSig-H, BHSig-B, and CEDAR Datasets. These additional comparisons are presented in Tab. A1 and Tab. A2, enhancing the comprehensiveness of our performance evaluation. The results further demonstrate that our method surpasses previous approaches.

Table A1: More OSV comparison on BHSig-H and BHSig-B Datasets.

Method	BHSig-H				BHSig-B			
	FAR	FRR	Acc ↑	EER ↓	FAR	FRR	Acc ↑	EER ↓
SET [21]	8.93	10.94	90.06	9.32	5.67	10.83	91.79	8.21
SWIS [17]	10.40	59.80	72.43	-	36.70	11.60	72.04	-
Ours	**1.07**	**3.59**	**98.24**	**2.07**	**0.95**	**4.04**	**98.19**	**2.11**

Table A2: More OSV comparison on CEDAR Dataset.

Method	FAR	FRR	Acc ↑	EER ↓
SigCNN [7]	-	-	-	6.41
Consensus [1]	5.27	17.45	92.29	11.36
Ours	**0.36**	**0.58**	**99.53**	**0.58**

Table A3: Datasets and protocol description. ([a]: number of signers. [b]: number of signatures per signer)

Datasets	**Language**	**Total**[a]	**Genuine**[b]	**Forged**[b]	**Training**[a]	**Testing**[a]
CEDAR [8]	English	55	24	24	50	5
BHSig-H [20]	Hindi	160	24	30	100	60
BHSig-B [20]	Bengali	100	24	30	50	50
ChiSig [27]	Chinese	500	3-5	above 10	80	20
GPDS Synthetic [5]	Western	4000	24	30	3200	800
MCYT-75 [18]	Western	75	15	15	-	-
SigComp 2011 [13]	Multi	64	23-24	8-20	-	-

	Training pairs		**Testing pairs**	
	pos	**neg**	**pos**	**neg**
CEDAR [8]	13800	28800	1380	2880
BHSig-H [20]	27600	72000	16560	43200
BHSig-B [20]	13800	36000	13800	36000

2 Overview of Major OSV Datasets

We have compiled detailed information on several prominent OSV datasets in Tab. A3. These datasets have been extensively utilized and referenced by previous OSV methods. While certain datasets may not be currently accessible, we have conducted our experiments using the datasets currently available. Below, we overview these OSV datasets.

2.1 ChiSig Dataset

The ChiSig Dataset [27] represents a new Chinese document offline signature dataset. It consists of 10,242 images that feature 500 distinct signed names. Given the intricate complexity of Chinese characters, this dataset provides an ideal testbed to evaluate the effectiveness of signature verification models. In our study, we used a subset of 100 signed names from the ChiSig dataset for training and testing purposes. Specifically, 80 of these names were used for the training phase, while the remaining 20 were reserved for testing.

2.2 CEDAR Dataset

The CEDAR Dataset [8] has a preprocessing detail that has been overlooked by some papers, including Longjam et al. [14], 2C2S [22] and DeepHSV [9]. As pointed out by CaC [15] and HybridFE [26], the image normalization is not performed during preprocessing, leading to unfair comparison (achieving 100% accuracy on CEDAR). We show some raw samples in Fig. A1. The difference in backgrounds between genuine and forged samples is clearly visible.

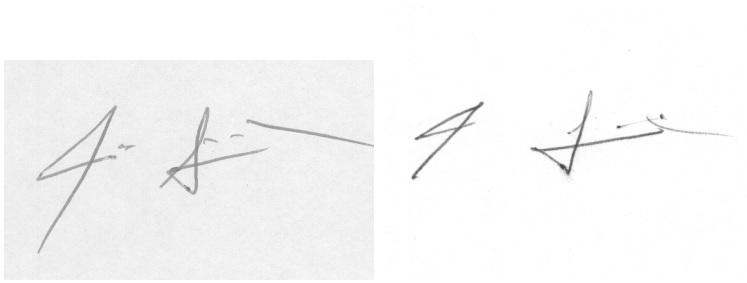

Fig. A1: A genuine (left) and a forged (right) sample of CEDAR. Without image normalization, differences in the background can make the model easily perform verification based on the background rather than the strokes of the signature.

2.3 GPDS Synthetic Dataset

The GPDS Synthetic Dataset [5] is a synthetic dataset comprising a substantial volume of image data. Consequently, training models on it requires a significant investment of time and computational resources. Moreover, despite its size, the synthetic nature of the dataset introduces a domain gap between the synthetic and real-world cases. Consequently, only a limited number of studies have utilized it for experimentation. Additionally, this dataset has several predecessors, such as GPDS-960 and GPDS-300, which are no longer accessible.

2.4 MCYT-75 Dataset

The MCYT-75 Dataset [18] is a widely recognized dataset within the OSV domain. It has served as a crucial resource in various prior studies, including those by Maergner et al. [16], Viana et al. [24], and Qian Wan and Qin Zou [25]. Unfortunately, public access to the dataset is currently unavailable. Therefore, seeking assistance from the research groups associated with previous studies becomes necessary.

2.5 SigComp 2011 Dataset

The SigComp 2011 Dataset [13] is one of the few OSV datasets that contain multilingual data. It includes signature data in both Dutch and Chinese. One of its drawbacks is the limited amount of data that it provides. The Dutch dataset, for instance, consists of only 362 signatures for training and 1932 signatures for testing, while the Chinese dataset includes 575 signatures for training and 602 signatures for testing. Additionally, its protocol differs from ours as it uses 12 signatures as references, unlike our method, which employs only one signature as a reference. For these reasons, recent methods tend to use other datasets mentioned earlier rather than the SigComp 2011 Dataset for experimental comparisons. Thus, our method is also less straightforward to compare directly with results obtained using this dataset.

3 Preprocessing

In the preprocessing stage of signature images, we aim to apply normalization and binarization techniques to render them more suitable for input into our model. The process begins with applying a Gaussian filter to the input image, which helps eliminate minor background noise. Subsequently, we employ Otsu's algorithm [19] for image binarization. This step involves finding the center of mass of the image and determining an appropriate threshold to remove background noise effectively. Once the center of mass is identified, it is used to centrally align the signature within the image. The final step in the preprocessing procedure involves resizing the image to a specific dimension. For our purposes, we resize the image to a standard size of 224×224 pixels. This standardized sizing ensures uniformity and consistency in the input data for our model.

4 Revisiting Earth Mover's Distance (EMD)

In our study, we utilize the Earth Mover's Distance (EMD) [28] as an auxiliary metric to evaluate the dissimilarity between two signatures, focusing on their local structure. Let $\mathcal{R} = \{(r_0, w_{r_0}), (r_1, w_{r_1}), ..., (r_{N-1}, w_{r_{N-1}})\}$ includes N pair data, where r_i denotes the i_{th} feature token of the i_{th} image patch in the reference signature, with w_{r_i} representing its corresponding weight. Similarly, $\mathcal{Q} = \{(q_0, w_{q_0}), (q_1, w_{q_1}), ..., (q_{M-1}, w_{q_{M-1}})\}$, having M pair data, signifies the pairs of feature tokens and their weights in the Query signature. These feature tokens are derived from a *Reference* image and a *Query* image, respectively, based on our DetailSemNet model. It is important to note that the weights assigned to these tokens can be pre-defined.

To measure the Earth Mover's Distance (EMD) between the two token sets $\mathcal{R}$ and $\mathcal{Q}$, we first define the *flow* between them as a matrix $F = [f_{ij}] \in \mathbb{R}^{N \times M}$. Intuitively, f_{ij} means the amount of importance of the i_{th} token in *Rererence* "flowing" to the j_{th} token in *Query*. Therefore, f_{ij} implicitly indicates the confidence level of local token matching. Next, we employ $D = (d_{ij}) \in \mathbb{R}^{N \times M}$ to denote the ground distance matrix. The distance element d_{ij} between (r_i, q_j) is calculated as defined in Eq. (6). This equation measures the cosine distance between pairs of token embeddings.

With the matrices (F) and (D), our objective is to identify the optimal flow F^* that minimizes the constrained objective function outlined in Eq. (A1). This involves calculating the total weighted distances (i.e., $d_{ij} f_{ij}$) between pairs of patch feature tokens from the two sets.

$$COST(\mathcal{R}, \mathcal{Q}, F) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} d_{ij} f_{ij} \tag{A1}$$

subject to the following constraints:

$$f_{ij} \geq 0, 0 \leq i \leq N-1, 0 \leq j \leq M-1 \tag{A2}$$

$$\sum_{i=0}^{N-1} f_{ij} \leq w_{q_j}, 0 \leq i \leq N-1 \tag{A3}$$

$$\sum_{j=0}^{M-1} f_{ij} \leq w_{r_i}, 0 \leq j \leq M-1 \tag{A4}$$

$$\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} f_{ij} = \min(\sum_{i=0}^{N-1} w_{r_i}, \sum_{j=0}^{M-1} w_{q_j}) \tag{A5}$$

After finding the optimal flow F^*, the Earth Mover's Distance is determined by the following operation:

$$EMD(\mathcal{R}, \mathcal{Q}) = \frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} d_{ij} f_{ij}^*}{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} f_{ij}^*} \tag{A6}$$

It is crucial to recognize that the Earth Mover's Distance (EMD) functions effectively as a metric only when two key conditions are met: first, the two distributions under comparison must have the same total weights, and second, the ground distance function must satisfy metric properties, as outlined in [23]. In our work, we assign uniform weights to each feature token and implement weight normalization by Eq. (A7) and Eq. (A8) to satisfy these requirements. This approach ensures that the cumulative weight of all features amounts to 1.

$$w_{r_i} = \frac{1}{N}, 0 \leq i \leq N-1 \tag{A7}$$

$$w_{q_j} = \frac{1}{M}, 0 \leq j \leq M-1 \tag{A8}$$

Finally, to effectively address the optimization problem outlined in Eq. (A1), we employ the Sinkhorn algorithm [2]. This algorithm enhances the traditional Earth Mover's Distance (EMD) by incorporating entropic regularization, which introduces a smoothing term. The inclusion of this term allows the Sinkhorn algorithm to provide an efficient and effective solution to the EMD optimization problem, balancing computational efficiency with the precision required for our analysis.

Table A4: Comparing different λ_0 on BHSig-H dataset.

λ_0	EER(%) ↓	Acc(%) ↑	AUC ↑
0.1	2.52	97.98	0.997
1.0	**2.07**	**98.24**	**0.998**
10.	2.77	97.59	0.997
100.	3.47	96.99	0.995

5 The Fusion Ratio between Global Distance and Local Structural Distance

Our model is specifically designed to calculate the "distance" between two input images, a critical factor in determining the final verification result. A key aspect of this process is the appropriate adjustment of the fusion ratio between the Global Distance and the Local Structural Distance. As outlined in Eq. (2), the hyperparameter λ_0 is instrumental in balancing the contributions of these two types of distances. Fine-tuning λ_0 directly influences the model's overall performance. The experimental findings regarding the optimization of λ_0 are comprehensively detailed in Tab. A4. Our experiments indicate that setting λ_0 to 1.0 yields the best results in terms of both accuracy and reliability. Therefore, this setting (λ_0=1.0) has been consistently applied across all our experimental configurations.

Table A5: Comparing Various Training Strategies: Our training involves using the BHSig-H dataset and testing on BHSig-B. In this context, the term "upper bound" denotes training the model on BHSig-B and testing it on the same dataset, serving as a reference point.

Setting	EER(%) ↓
zero-shot	7.46
finetuning	4.70
upper bound	2.11

Table A6: Ablation study of w/wo SemanticsAttend on EER (%). The performance significantly drops without the attention module.

	BHSig-H	BHSig-B	CEDAR	ChiSig
×	21.06	13.29	22.14	29.01
✓	**2.07**	**2.11**	**0.58**	**5.85**

6 Cross-dataset Evaluation with Different Training Strategies

In the context of cross-dataset scenarios, in addition to assessing the model's robustness in a zero-shot setting, we also explore the effectiveness of fine-tuning our model for performance evaluation on different datasets. Fine-tuning enables the model to adapt to specific characteristics of a new dataset, which can lead to

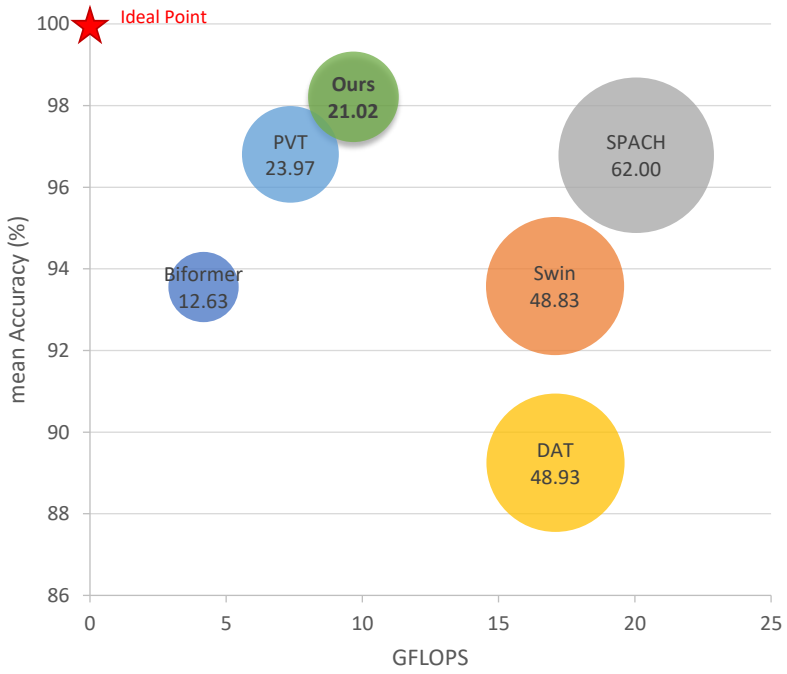

Fig. A2: Compare our method with different transformer-based architecture. The X-axis represents computational complexity, the Y-axis represents verification accuracy, and the size of the bubbles represents model size (M).

improved performance, especially in cases where the training and testing datasets vary significantly.

As detailed in Tab. A5, we present the results obtained using different training strategies. Here, the "zero-shot" testing approach, which is consistent with the methodology described in the sections (Tab. 3) of the main paper, involves training the model on the BHSig-H dataset and subsequently evaluating its performance on the BHSig-B dataset directly without any additional fine-tuning. As for the term "upper bound" mentioned in Tab. A5, it represents the performance achieved in the specific experiment (Tab. 3), where both training and testing occur on the same dataset (BHSig-B in this instance). For the "fine-tuning" process, we unfreeze only the parameters of $Conv_{fuse}$, meaning that the adjustments involve only a limited number of parameters. The results in Tab. A5 demonstrate that even with minimal parameter adjustments, the model's performance can closely approach the upper bound. This outcome underscores the model's strong

generalization capabilities, suggesting its potential effectiveness across different dataset scenarios.

7 More Ablation Studies on the SemanticsAttend Branch

Account for the impact of the SemanticsAttend Branch. In terms of results with and without the SemanticsAttend Branch, we present them in Tab. A6. This highlights the effectiveness of our method.

In order to further compare the possibility of using CNNs as backbones in our method. In experiments using CNNs as the backbone, including SM in CNN contributes to worse performance. In the case of using ResNet50 [6] as the backbone and evaluated on the ChiSig dataset, Acc dropped from 87.79% to 87.28%. In summary, CNN backbone performs less effectively on OSV than ViT.

8 GFLOPs and Model Size

According to Section 4.5, we compared the EER and Acc between various transformer architectures with our method, and here we further consider model size and computational complexity. If our model is to be further used in real-world scenarios, these indicators remain equally important. Detailed results can be seen from Fig. A2, where our method achieves the highest accuracy (Y-axis) while maintaining relatively good performance in terms of model parameter size (bubble size, M) and computational complexity (X-axis). From this experiment, we can further observe that our model not only outperforms other methods in verification accuracy, but also exhibits good computational efficiency. This is evidently crucial for real-world deployment scenarios.

9 Add Classifier

In the *synthetic dataset* [5], we can additionally integrate a classifier directly into our model, allowing it to learn the common genuine/ forged differences between inter-classes. Specifically, we concatenate two features from the third stage of the siamese backbone. Using a multi-layer convolutional classifier, the final classifier outputs two logits, which are supervised through cross-entropy loss.

To train the classifier part, we employ a cross-entropy loss, defined as follows:

$$p_i = \frac{e^{z_i}}{e^{z_0} + e^{z_1}} \quad for\ i = 0, 1 \tag{A9}$$

$$Loss_{CE} = ylog(p_0) + (1 - y)log(p_1). \tag{A10}$$

Here, the supervised label y is assigned a value of 1 for positive pairs of signature (genuine-genuine) and 0 for negative (genuine-forged) signature pairs. p_0 and p_1 represent the two logits, z_0 and z_1, output of the classifier after softmax.

In the model, which integrates a classifier, training is performed using $Loss_{DM}$ combined with $Loss_{CE}$, defined as

$$Loss_{Total} = Loss_{DM} + Loss_{CE}. \tag{A11}$$

During inference time, we introduce the CE distance, $dis_{CE} = z_0 - z_1$. The total distance, if the model includes a classifier, is calculated as

$$dis' = dis + dis_{CE}. \tag{A12}$$

Table A7: Signature verification comparison on GPDS Synthetic Dataset (%).

Method	FAR	FRR	Acc ↑	EER ↓
SigNet [3]	22.24	22.24	77.76	22.24
Correlated Features [4]	28.34	27.62	88.79	-
SDINet [10]	12.37	8.32	89.66	10.43
AVN [11]	11.78	7.58	90.32	9.77
TransOSV [12]	10.64	10.64	-	10.64
Ours *#ref=1*	11.56	11.81	88.33	11.69
Ours *#ref=2*	8.60	9.81	90.89	9.15
Ours *#ref=3*	7.96	8.20	91.95	8.06
Ours *classifier*	9.17	8.59	91.12	8.92

10 Results on GPDS Synthetic Dataset

The GPDS Synthetic Dataset [5] is a synthetic dataset comprising a substantial volume of image data. The previous methods used 3200 signers for training and the remaining 800 signers for testing. Consequently, training models on it requires a significant investment of time and computational resources. Moreover, despite its size, the synthetic nature of the dataset introduces a domain gap between the synthetic and real-world cases. Our model focuses on detecting subtle differences between signature pairs, making it less suitable for the synthetic dataset. However, with minor adjustments to the testing methods or model architecture, our model can still outperform the state-of-the-art methods on the synthetic dataset. First, we can show the testing strategy to use only one reference pair (*#ref=1*), resulting in an EER of 11.69. By increasing the number of reference pairs, we can mitigate synthetic bias and improve verification performance. Using two reference pairs (*#ref=2*), the EER is 9.15, and with three reference pairs (*#ref=3*), the EER is 8.06, surpassing the previous methods, such as SigNet [3], Correlated Features [4], SDINet [10], AVN [11], and TransOSV [12]. We can also enhance our original model with a simple classification architecture, allowing it to directly learn the authenticity features of the synthetic data. With this modification, our model achieves an EER of 8.92,

surpassing the performance of the previous best methods. Our method can outperform by using multiple pairs or simply adding a classifier while testing. The results of these comparative methods on the GPDS Synthetic Dataset are detailed in Tab. A7. In the following, we will describe the details of our two minor adjustments.

10.1 Multiple References

In the GPDS dataset [5], we improve verification accuracy by increasing the number of reference images. Specifically, each class in the GPDS dataset contains 24 genuine signatures and 30 forged signatures. From the genuine signatures, a specific number (1, 2, or 3) of samples are selected as reference images, and the remaining samples are used as test images. In other words, when the number of reference images is set to 1, 2, and 3, each class contains 23, 44, and 63 positive pairs and 30, 60, and 90 negative pairs, respectively. In this setup, we average the distances obtained from multiple pairs for evaluation.

10.2 Add Classifier

In Sec. 9, we describe how a simple classifier can be used to improve the accuracy of our model on synthetic datasets. Here, we perform the evaluation using Eq. (A12). The experimental results are shown in the last row of Tab. A7, labeled **Ours** *classifier*, demonstrating that it surpasses the precision of previous methods. Achieving an accuracy (Acc) of 91.12%, a False Acceptance Rate (FAR) of 9.17%, a False Rejection Rate (FRR) of 8.59%, and an Equal Error Rate (EER) of 8.92%.

References

1. Brimoh, P., Olisah, C.C.: Consensus-threshold criterion for offline signature verification using convolutional neural network learned representations. ArXiv **abs/2401.03085** (2024), https://api.semanticscholar.org/CorpusID:266844275
2. Cuturi, M.: Sinkhorn distances: Lightspeed computation of optimal transport. In: Burges, C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K. (eds.) Advances in Neural Information Processing Systems. vol. 26. Curran Associates, Inc. (2013), https://proceedings.neurips.cc/paper_files/paper/2013/file/af21d0c97db2e27e13572cbf59eb343d-Paper.pdf
3. Dey, S., Dutta, A., Toledo, J.I., Ghosh, S.K., Llados, J., Pal, U.: Signet: Convolutional siamese network for writer independent offline signature verification. arXiv preprint arXiv:1707.02131 (2017)
4. Dutta, A., Pal, U., Lladós, J.: Compact correlated features for writer independent signature verification. In: 2016 23rd International Conference on Pattern Recognition (ICPR). pp. 3422–3427 (2016). https://doi.org/10.1109/ICPR.2016.7900163

5. Ferrer, M.A., Diaz-Cabrera, M., Morales, A.: Static signature synthesis: A neuromotor inspired approach for biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence **37**, 667–680 (2015), https://api.semanticscholar.org/CorpusID:16293664
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 770–778 (2015), https://api.semanticscholar.org/CorpusID:206594692
7. Jiang, J., Lai, S., Jin, L., Zhu, Y., Zhang, J., Chen, B.: Forgery-free signature verification with stroke-aware cycle-consistent generative adversarial network. Neurocomputing **507**, 345–357 (2022), https://api.semanticscholar.org/CorpusID:251420764
8. Kalera, M., Xu, A.: Offline signature verification and identification using distance statistics. IJPRAI **18**, 1339–1360 (11 2004). https://doi.org/10.1142/S0218001404003630
9. Li, C., Lin, F., Wang, Z., Yu, G., Yuan, L., Wang, H.: Deephsv: User-independent offline signature verification using two-channel cnn. In: 2019 International Conference on Document Analysis and Recognition (ICDAR). pp. 166–171 (2019). https://doi.org/10.1109/ICDAR.2019.00035
10. Li, H., Wei, P., Hu, P.: Static-dynamic interaction networks for offline signature verification. In: AAAI Conference on Artificial Intelligence (2021), https://api.semanticscholar.org/CorpusID:235306077
11. Li, H., Wei, P., Hu, P.: Avn: An adversarial variation network model for handwritten signature verification. IEEE Transactions on Multimedia **24**, 594–608 (2022). https://doi.org/10.1109/TMM.2021.3056217
12. Li, H., Wei, P., Ma, Z., Li, C., Zheng, N.: Transosv: Offline signature verification with transformers. Pattern Recognition **145**, 109882 (2024). https://doi.org/https://doi.org/10.1016/j.patcog.2023.109882, https://www.sciencedirect.com/science/article/pii/S0031320323005800
13. Liwicki, M., Malik, M.I., Heuvel, C.E.v.d., Chen, X., Berger, C., Stoel, R., Blumenstein, M., Found, B.: Signature verification competition for online and offline skilled forgeries (sigcomp2011). In: 2011 International Conference on Document Analysis and Recognition. pp. 1480–1484 (2011). https://doi.org/10.1109/ICDAR.2011.294
14. Longjam, T., Kisku, D.R., Gupta, P.: Writer independent handwritten signature verification on multi-scripted signatures using hybrid cnn-bilstm: A novel approach. Expert Syst. Appl. **214**(C) (mar 2023). https://doi.org/10.1016/j.eswa.2022.119111
15. Lu, X., Huang, L., Yin, F.: Cut and compare: End-to-end offline signature verification network. In: 2020 25th International Conference on Pattern Recognition (ICPR). pp. 3589–3596 (2021). https://doi.org/10.1109/ICPR48806.2021.9412377
16. Maergner, P., Pondenkandath, V., Alberti, M., Liwicki, M., Riesen, K., Ingold, R., Fischer, A.: Combining graph edit distance and triplet networks for offline signature verification. Pattern Recognition Letters **125**, 527–533 (2019). https://doi.org/https://doi.org/10.1016/j.patrec.2019.06.024, https://www.sciencedirect.com/science/article/pii/S0167865519301850
17. 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). pp. 1411–1415 (2022). https://doi.org/10.1109/ICIP46576.2022.9897562

18. Ortega-Garcia, J., Fierrez, J., Simon, D., Gonzalez, J., Faundez-Zanuy, M., Espinosa, V., Satue, A., Hernáez, I., Igarza, J., Vivaracho-Pascual, C., Escudero-Mancebo, D., Moro-Sancho, Q.: Mcyt baseline corpus: a bimodal biometric database. iee proc vis image signal process spec issue biom internet. IEE Proceedings - Vision Image and Signal Processing pp. 395 – 401 (12 2003). https://doi.org/10.1049/ip-vis:20031078
19. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics **9**(1), 62–66 (1979). https://doi.org/10.1109/TSMC.1979.4310076
20. Pal, S., Alaei, A., Pal, U., Blumenstein, M.: Performance of an off-line signature verification method based on texture features on a large indic-script signature dataset. In: 2016 12th IAPR Workshop on Document Analysis Systems (DAS). pp. 72–77 (2016). https://doi.org/10.1109/DAS.2016.48
21. Ren, J.X., Chen, J., Xiong, Y.J.: Set: a squeeze-and-excitation transformer for offline signature verification. 2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta) pp. 1812–1816 (2022), https://api.semanticscholar.org/CorpusID:260254654
22. Ren, J.X., Xiong, Y.J., Zhan, H., Huang, B.: 2c2s: A two-channel and two-stream transformer based framework for offline signature verification. Engineering Applications of Artificial Intelligence **118**, 105639 (2023). https://doi.org/https://doi.org/10.1016/j.engappai.2022.105639, https://www.sciencedirect.com/science/article/pii/S0952197622006297
23. Rubner, Y., Tomasi, C., Guibas, L.: The earth mover's distance as a metric for image retrieval. International Journal of Computer Vision **40**, 99–121 (11 2000). https://doi.org/10.1023/A:1026543900054
24. Viana, T.B., Souza, V.L., Oliveira, A.L., Cruz, R.M., Sabourin, R.: A multi-task approach for contrastive learning of handwritten signature feature representations. Expert Systems with Applications **217**, 119589 (2023). https://doi.org/https://doi.org/10.1016/j.eswa.2023.119589, https://www.sciencedirect.com/science/article/pii/S0957417423000908
25. Wan, Q., Zou, Q.: Learning metric features for writer-independent signature verification using dual triplet loss. 2020 25th International Conference on Pattern Recognition (ICPR) pp. 3853–3859 (2021), https://api.semanticscholar.org/CorpusID:233877659
26. Xiong, T., Zhang, X.: Hybrid feature extraction based deep learning model for offline signature verification. 2023 6th International Conference on Software Engineering and Computer Science (CSECS) pp. 1–6 (2023), https://api.semanticscholar.org/CorpusID:267703979
27. 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: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. pp. 5163–5172 (June 2022)
28. Zhang, C., Cai, Y., Lin, G., Shen, C.: Deepemd: Differentiable earth mover's distance for few-shot learning. IEEE Transactions on Pattern Analysis and Machine Intelligence p. 1–17 (2022). https://doi.org/10.1109/tpami.2022.3217373, http://dx.doi.org/10.1109/TPAMI.2022.3217373