Inverse Discriminative Networks for Handwritten Signature Verification

Ping Wei, Huan Li, Ping Hu*

Xi'an Jiaotong University, Xi'an, China

pingwei@xjtu.edu.cn, lh875056558@stu.xjtu.edu.cn, helenhu@xjtu.edu.cn

Abstract

Handwritten signature verification is an important technique for many financial, commercial, and forensic applications. In this paper, we propose an inverse discriminative network (IDN) for writer-independent handwritten signature verification, which aims to determine whether a test signature is genuine or forged compared to the reference signature. The IDN model contains four weight-shared neural network streams, of which two receiving the original signature images are the discriminative streams and the other two addressing the gray-inverted images form the inverse streams. Multiple paths of attention modules connect the discriminative streams and the inverse streams to propagate messages. With the inverse streams and the multi-path attention modules, the IDN model intensifies the effective information of signature verification. Since there was no proper Chinese signature dataset in the community, we collected a large-scale Chinese signature dataset with approximately 29,000 images of 749 individuals' signatures. We test our method on the Chinese signature dataset and other three signature datasets of different languages: CEDAR, BHSig-B, and BHSig-H. Experiments prove the strength and potential of our method.

1. Introduction

When a myriad of significant financial, commercial, and forensic documents are signed worldwide everyday, verifying the authenticity of the signatures is a critical issue to be concerned. Considering the huge amounts and wide applications of handwritten signatures, developing an automatic, accurate, and efficient signature verification technique is becoming particularly important and necessary.

This paper addresses the problem of writer-independent handwritten signature verification, which aims to determine whether a test signature is genuine or forged compared with the reference signature of any writer, as shown in Fig. 1. While the past decades have witnessed remarkable progress in signature verification [2, 12, 18, 35, 34], several existing challenges make it still an open problem. First, there was no proper Chinese signature dataset, which impedes the research and application of Chinese signature verification. Second, in a signature image the information of the signature is very sparse, because the signature strokes are often extremely thin and a large area of the image is the background. Third, most individual's signature styles are somewhat arbitrary, which makes the same individual's signatures on different occasions appear notably different. On the other hand, some skillfully forged signatures appear extremely similar to the genuine ones.

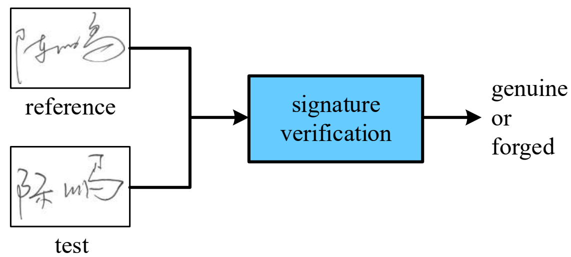

Figure 1. Illustration of handwritten signature verification.

In this paper, we propose a novel inverse discriminative network (IDN) model for writer-independent handwritten signature verification. This network contains four weight-shared streams, of which two streams are the discriminative streams and the other two are the inverse streams. The two discriminative streams respectively receive a reference signature image and a test signature image as inputs, and extract the signature features via four cascaded convolutional modules. The two inverse streams receive the inverse-gray reference and test signature images, respectively. The discriminative streams and the inverse streams are connected by multiple paths of attention modules which propagate messages at different scales to intensify the effective stroke information. The features from different discrimi-

*Ping Hu is the corresponding author.

native streams and inverse streams are merged to three different feature maps with convolutional modules, which are then fed to three fully-connected layers to make decisions. The whole IDN model is trained in an end-to-end way.

Our IDN model introduces two mechanisms which aim to resolve the sparse information issue of signatures. The first one is the inverse supervision mechanism which takes the inverse-gray reference signature and test signature as inputs and pushes the model to focus on the signature strokes rather than the image backgrounds. This mechanism is built on the fact that a model that focuses on the signature strokes rather the image should make the same verification decision when the gray values of the signature image are inverted. The second one is the multi-path attention mechanism which propagates messages between inverse streams and discriminative streams via multiple attention modules at different feature scales. The attention mechanism aims to enforce the model to learn and extract important features for signature verification.

Since there was no proper Chinese signature dataset in the community, we collected a large-scale and challenging Chinese signature dataset (CSD). We test our method on the collected Chinese signature dataset and other three public signature datasets of different languages: CEDAR Dataset [21], BHSig-B Dataset [27], and BHSig-H [27]. Extensive experiments demonstrate the effectiveness and strength of the proposed method.

1.1. Related Work

For the significance in financial, commercial, and forensic applications, signature verification has been extensively studied over the past decades [38, 34, 16, 18, 8], and many datasets were publicly released, such as CEDAR [21], MCYT-75 [14], BHSig [27], and GPDS [12, 13]. However, there was no large-scale Chinese signature dataset in the community, which impedes the research and applications on Chinese signature verification. This motivates us to collect a new Chinese signature dataset.

Geometric features in images are often used for signature verification [2, 10, 11, 35, 34, 30, 29], such as the signature heights, widths, areas, [2, 10, 11], or local patch features, such as LBP [35, 34, 30] and SIFT [29]. These features have laid a solid foundation for signature verification and performed well on some datasets. However, the hand-crafted features are vulnerable to noise and complex backgrounds, which makes them less effective on some complex data.

To overcome the drawbacks of hand-crafted features, neural network approaches are widely applied to signature verification [17, 18, 8, 1, 37, 28, 33, 22] and related tasks [36, 7, 23, 31]. Hafemann et al. [17] utilized convolutional neural networks to learn features in a writer-independent way, and presented a multi-task model [18] which both uses genuine signature and forgeries to train the networks. Dey et al. [8] modeled an offline writer independent signature verification with a Siamese convolutional network. Alvarez et al. [1] proposed a CNN-based architecture which combines a positive sample and a negative sample into a single image. Zhang et al. [37] presented an offline signature verification with deep convolutional generative adversarial networks [15]. Compared to the traditional methods, neural network methods have achieved impressive performance on signature verification. However, most existing approaches indeed address the problem of signature verification in the way of image classification rather than modeling the signature itself, which may lead to incorrect predictions on complex signature images.

We propose a four-stream network model which takes in two pairs of signature images: one pair contains the reference signature image and the test signature image, and another pair contains the inverse gray reference signature image and test signature image. With this strategy, our model not only extracts features from signature images but also specifically mines the signature stroke information.

Signature information is very sparse in images because signatures are often composed of thin strokes. Attention mechanism [32, 4, 19] is an effective way to enhance weak information and improves performance in object and image recognition. Chen *et al.* [6] designed a reverse attention method to detect salient objects. Huang *et al.* [20] utilized a reverse attention mechanism for semantic segmentation. Inspired by these attention models, we develop a multi-path attention approach which supervises the model to focus on and mine the signature stroke information.

2. Chinese Signature Dataset

Since there was no proper Chinese signature dataset, we collected a large scale and challenging Chinese signature dataset (CSD). Some examples are shown in Fig. 2. The dataset includes genuine signatures and forged signatures. To collect the genuine signatures, volunteers using Chinese wrote their names 20 times on a writing paper at different time. For forged signatures, each name has 10 simple forgeries and 10 skilled forgeries. The simple forgeries of each name were wrote by 10 different volunteers with their own writing styles and habits. The skilled forgeries of each name were wrote by calligraphers after they have carefully observed, learned, and imitated the genuine signatures.

All the writing papers with signatures were scanned to images from which all the handwritten signature patches were cropped and resized into image samples with the same size. With the OTSU algorithm [26] and non-standard Binarization, these signature images are preprocessed so that the background pixel values are 255 (white) and the signature strokes maintain the original gray values. In this way, each name has 20 genuine handwritten signature image samples and 20 forged handwritten signature samples. The dataset

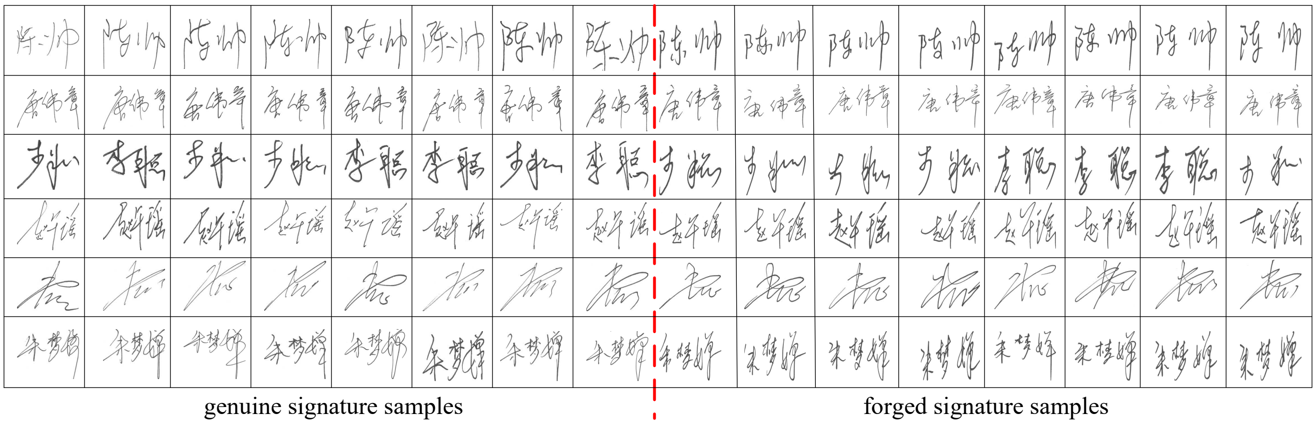

Figure 2. Samples of our Chinese Signature Dataset. In each row are the same name's signatures. The left eight samples are the genuine signatures and the right eight samples are the forged signatures.

includes a total of 749 names and approximately 29,000 signature image samples.

Our dataset has several characteristics which make it unique and challenging. First, our dataset is a large scale Chinese signature dataset, which we believe will contribute to the studies of Chinese signature verification and other related tasks. Second, it has a large number of individual's real handwritten signatures, which were all manually collected in natural settings. Third, since the genuine signatures were collected at different time and at different scenes, the same individual's signatures may appear notably different, as shown in Fig. 2. On the other hand, since the skilled forgeries were wrote by professional calligraphers, the forged signatures may appear extremely similar to the genuine ones. All these aspects make it a challenging and valuable signature dataset.

3. Inverse Discriminative Networks

Signature strokes are the decisive features to determine a signature's identity. However, a typical characteristic of a signature image is that the effective information for signature verification in the image is very sparse because the signature strokes are often extremely thin and most part of the signature image is the background. In this sparse information setting, if the signature stroke information was not effectively modeled and utilized, the background will be the dominant information and signature verification would become a common image classification problem. Thus, how to make the model focus on the signature strokes rather than the background is a key concern in signature verification.

To solve this problem, we propose a novel inverse discriminative network (IDN) model. The basic idea is that a robust model that captures the characteristics of signature strokes rather than the image itself will make the same verification decision when the gray values of the signature image are inverted. We use Fig. 3 to illustrate this point. The

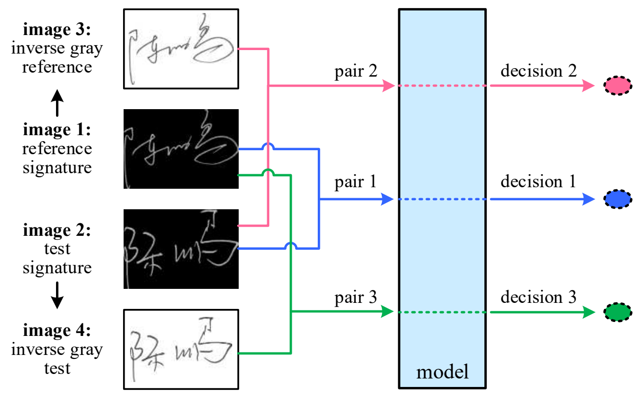

Figure 3. A robust signature verfication model should make the same decisions for the three reference-test pairs, i.e. the decision 1, the decision 2, and the decision 3 are the same.

image 1 and the *image* 2 are the reference signature and the test signature respectively, which have black backgrounds and gray signature strokes. The *image* 3 and the *image* 4 with white backgrounds are the inverse gray images of the *image* 1 and the *image* 2, respectively. The four signature images produce three reference-test pairs: the pair 1 (*image* 1, *image* 2), the pair 2 (*image* 2, *image* 3), and the pair 3 (*image* 1, *image* 4). The signature verification model taking the three pairs respectively as inputs would output three verification decisions: the decision 1, the decision 2, and the decision 3, respectively.

Since the three reference-test pairs originate from the same pair, if a signification verification model was well designed for characterizing the stroke information rather than the image color information, it should make the same verification decisions for the pair 1, the pair 2, and the pair 3, i.e. the decision 1, the decision 2, and the decision 3 are the same. For example, compared with pair 1, the reference signature in pair 2 is gray-inverse. If the model focuses on

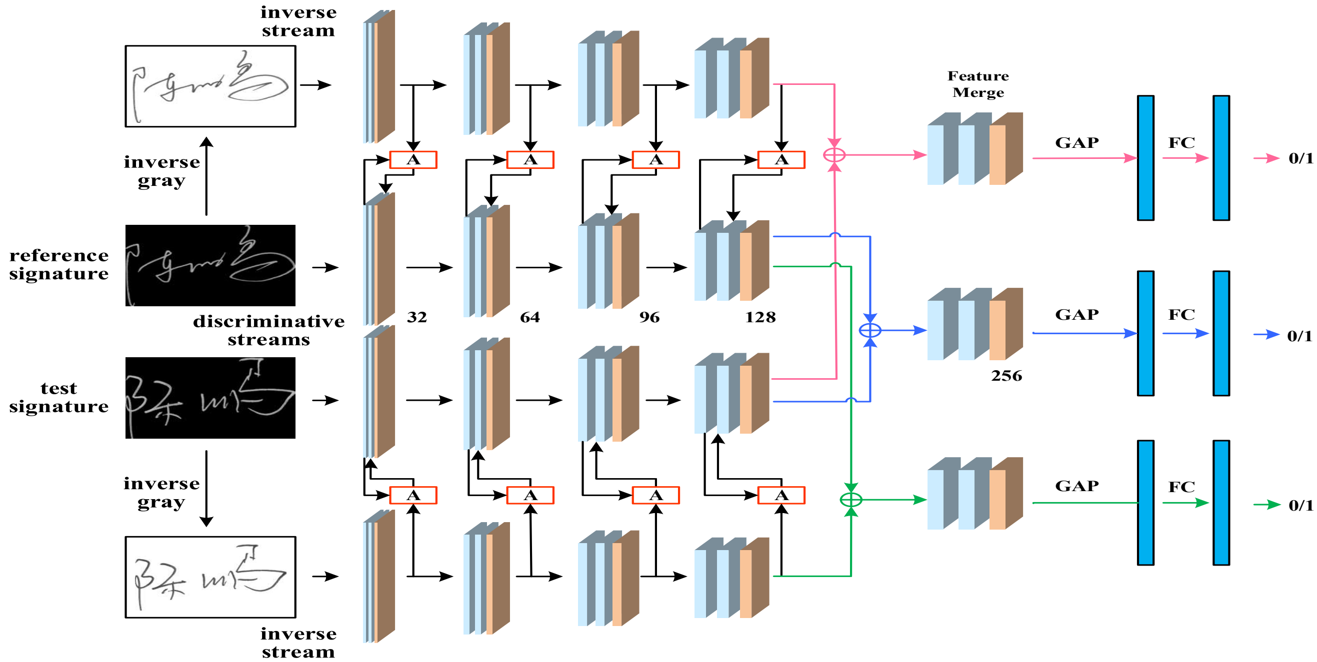

Figure 4. Architecture of the proposed inverse discriminative network. The discrininative and the inverse streams are connected by multi-path attention (A) modules, shown as the red boxes. Through a global average pooling (GAP) layer, the merged features are fed into the fully-connected (FC) layers to compute the verification results.

the signature stroke information, it will make the same decisions for the pair 1 and the pair 2 regardless of the image colors. Since the three pairs have different gray values, the common information should be related to the stroke information. Training the model with this strategy will force the model to focus on the signature strokes rather than the image colors.

Driven by this inverse supervision idea, we design the inverse discriminative network architecture.

3.1. Architecture

The proposed inverse discriminative network (IDN) is illustrated in Fig. 4. The input reference and test signature images to the model are with black backgrounds and gray signature strokes. The inverse images are with white backgrounds and gray strokes. The network contains four weight-shared streams, of which two are the discriminative streams and the other two are the inverse streams.

The two discriminative streams respectively take a reference signature image and a test signature image as inputs, and extract the signature features via cascaded convolutional modules [31]. Each convolutional module contains two convolutional layers (the kernel size is 3×3 and the strip is 1) activated by the ReLU function and one max-pooling layer (the kernel size is 2×2 and the strip is 2). The kernel numbers of the four modules in each stream are 32, 64, 96, and 128, respectively. The two inverse streams take the inverse-gray reference and test signature images as inputs, respectively. Each inverse stream has the same structure with the discriminative stream.

Between the discriminative and the inverse streams there are eight paths of attention modules connecting the convolutional modules of the two streams. As the red box shown in Fig. 4, each attention module is composed of a forward process and a backward process. The forward process receives features output from the first layer of the convolutional module in the discriminative stream. The backward process propagates attention information from the inverse stream to the second layer of the convolutional module in the discriminative stream. The inside structures of the attention module will be detailed in Section 3.2.

With three convolutional modules (two convolutional layers and a max-pooling layer with 256 kernels), the features from different streams are merged to three feature maps, which correspond to three pairs: the reference signature and the test signature, the inverse-gray reference signature and the test signature, the reference signature and the inverse-gray test signature. Through a global average pooling (GAP) layer, the three merged features are respectively input into three fully-connected layers to compute the verification results.

In the IDN architecture, the discriminative streams and the inverse streams are closely connected by the multi-path attention processes. With these connections, the whole IDN model is trained in an end-to-end way. This model uses two mechanisms to enforce the model to focus on the signature

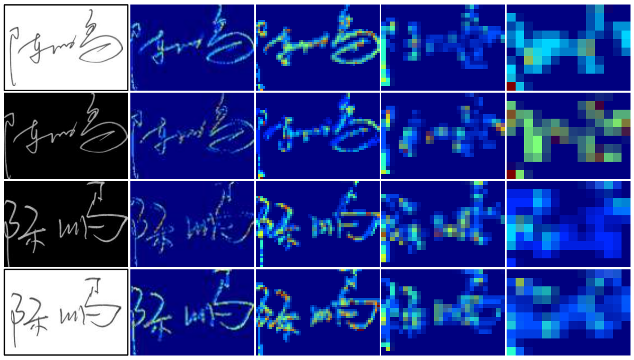

Figure 5. Feature maps output from the cascaded convolutional modules of the four streams.

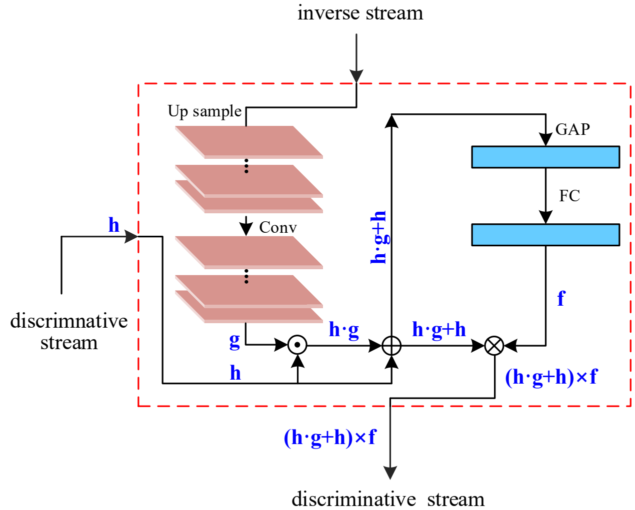

Figure 6. Attention module in the IDN framework. 'FC' denotes 'fully-connected' and 'GAP' indicates 'global average pooling'. '+' and '·' indicate element-wise addition and multiplication, respectively. '×' means multiplying each channel with a weight.

strokes rather than the whole image. The first is the inverse supervision mechanism. Based on the fact that inverting the gray values of signature images should not change the verification result, the inverse supervision mechanism will drive the feature extraction to focus on the signature strokes. The second one is the multi-path attention mechanism which enforces the model to extract important features for signature verification.

Fig. 5 shows some feature maps output from the cascaded convolutional modules of the four streams. This figure demonstrates that after cascaded attention and inverse supervision, the information for signature verification concentrates around signature strokes.

3.2. Multi-path Attention Modules

In the IDN framework, eight paths of attention modules propagate information between the discriminative streams and the inverse streams to force the model to extract important features for signature verification. Each attention module connects a convolutional module in the discriminative stream and a convolutional module in the inverse stream, as the red boxes shown in Fig. 4. Our attention module is inspired by the previous attention models in image-related tasks [32, 4, 19] but re-designed for connecting the discriminative streams and the inverse streams.

Fig. 6 shows the message flows inside an attention module. The feature map output from the convolutional module in the inverse stream is input into a up-sample structure which performs a up sampling with nearest neighbor algorithm and a convolution operation with sigmoid activation, as shown in the left side of Fig. 6. Let $\mathbf{g}$ be the output of the up-sample structure. Suppose $\mathbf{h}$ is the output from the first layer of the convolutional module in the discriminative stream. In the attention module, multiplying $\mathbf{h}$ by $\mathbf{g}$ element-wise and then adding $\mathbf{h}$ produce the intermediate attention measurement $\mathbf{h} \cdot \mathbf{g} + \mathbf{h}$, where '·' indicates element-wise multiplication. A following global average pooling (GAP) layer and a fully-connected layer (FC) with sigmoid activation receive the intermediate attention measurement and output the weight vector $\mathbf{f}$, as shown in the right side of Fig. 6. Multiplying each channel of the intermediate attention measurement by each element of $\mathbf{f}$ respectively generates the final attention mask $(\mathbf{h} \cdot \mathbf{g} + \mathbf{h}) \times \mathbf{f}$, which is fed back to the second layer of the convolutional module in the discriminative stream.

Since our attention module connects both the discriminative stream and the inverse stream, the final attention mask will guide the network to learn discriminative features for signature verification and restrain the misleading information. The whole IDN architecture has eight paths of attention modules connecting different convolutional modules, which applies the attention mechanism to different scales and resolution. With the multi-path attention mechanism, the important features for signature verification are enhanced.

3.3. Loss Function

As we discussed above, the signature verification decision should be independent of the signature image colors if the model correctly characterizes the signature stroke information. By inverting the gray values of the signature images, our model produces merged features for three pairs: the reference signature and the test signature, the inverse-gray reference signature and the test signature, the reference signature and the inverse-gray test signature, as shown in Fig. 4. In training, by forcing the model making the same decisions of signature verification for the merged features of the three pairs, the model will be guided to focus on the signature stroke information. We propose an inverse supervision loss function based on the cross entropy error.

Suppose y is a binary ground truth label of a test signature with respect to the reference signature, where 1 indicates the test signature is genuine and 0 indicates forged. $\hat{y}_i (i = 1, 2, 3)$ are the predicted probability values for the three pairs of the reference signature and the test signature, the inverse-gray reference signature and the test signature, the reference signature and the inverse-gray test signature, respectively. Based on the cross-entropy error function of binary classifiers, the inverse supervision loss for a single example is defined as:

$$L = -\sum_{i=1}^{3} \alpha_i [y \ln \hat{y}_i + (1 - y) \ln(1 - \hat{y}_i)], \qquad (1)$$

where α_i is a hyper-parameter which adjusts the three pairs' weights.

The inverse supervision loss has three components but with the same ground truth, which is different from the traditional cross entropy loss. Since the four streams of the network share the parameters, the model will be forced to focus on and mine the signature stroke information.

4. Experiments

We test our approach on four datasets: our Chinese Signature Dataset (CSD), CEDAR Dataset [21], BHSig-B Dataset [27], and BHSig-H [27], which belong to four different languages respectively: Chinese, English, Bengali, and Hindi. We also carry out the cross-language experiments, i.e. training on a dataset of a language and test on another dataset of a different language.

We train the model based on TensorFlow 1.4 platform with NIDIA 1080Ti and i7-8700 CPU. We use the mini-batch SGD with base learning rate 0.01.

4.1. Evaluation Metrics

We use Fasle Rejection Rate (FRR), False Acceptance Rate (FAR), Equal Error Rate (EER), Area Under the Curve (AUC), and Accuracy (Acc) to comprehensively evaluate our approach and compare it with other existing approaches.

FRR is defined as the ratio of the number of false rejections divided by the number of genuine samples and FAR is defined as the ratio of the number of false acceptances divided by the number of forged samples. Since FRR and FAR are mutually restricted, EER is applied to evaluate the equilibrium point where FRR equals to FAR. The lower EER is, the better model performance is. AUC is the area under the ROC curve, which is a comprehensive metric. Accuracy is the ratio of the number of correctly predictions divided by the number of all test samples.

4.2. Chinese Signature Dataset

Our Chinese Signature Dataset has 749 individuals' signature samples and each individual has 20 genuine samples and 20 forged samples. Among all the 749 individuals, we use 375 individuals' samples as training data, 187 individuals' samples as validation data, and the rest as testing data. For each individual, we have 190 ($20 \times 19/2$) pair samples of the reference and the genuine signature. We randomly select 10 genuine signatures as the references and 19 forgeries to form 190 pair samples of the reference and the forged signature. Thus, for each individual, we have a total of 380 pair samples, of which 190 are reference-genuine pairs and 190 are reference-forgery pairs. Since our forged samples include simple forgeries and skilled forgeries, we separated the simple forgeries and skilled forgeries in testing. The final performance is based on the average results of the simple forgeries and skilled forgeries.

Model	Acc	FRR	FAR	EER	AUC
CNN OSV[1]	82.75	10.51	19.05	19.63	88.63
Single Stream	88.06	10.98	13.07	15.57	92.28
Double Stream	88.26	8.99	12.64	15.78	91.85
Our IDN	**90.17**	**5.47**	**11.52**	**10.83**	**95.79**

Table 1. Comparison on Chinese Signature Dataset (%).

We compare our IDN method with other three approaches. The CNN OSV method [1] uses a convolutional neural network model to verify signatures in an offline way. The Single Stream method concatenates the reference signature and the test signature into one image and uses one stream of our IDN model to extract the features of the concatenated image and determine its label. The Double Stream method takes in the reference signature and the test signature in the two discriminative streams of our IDN model respectively, but without inverse streams and multi-path attention modules. Our IDN model have four streams which exploit the multi-path attention and the inverse mechanism for signature verification.

Table 1 shows the results of different approaches and Fig. 7 (a) shows the ROC curves of the Single Stream, the Double Stream, and our IDN. The results show that our IDN model outperforms other approaches by a large margin in all evaluation metrics. The reason why our IDN outperforms other approaches is that it takes advantage of the inverse supervision mechanism and the multi-path attention mechanism. This point is clearly demonstrated in the comparison with the Single Stream and the Double Stream approaches. The Single Stream method uses one stream of IDN to extract features and make decisions. The Double Stream method extracts the features of the reference and the test signatures respectively in two discriminative streams. Compared with these two baseline methods, the IDN has inverse streams and multi-path attention modules, which makes the IDN outperform the two baselines by a large margin. This proves the effectiveness of the inverse supervision and multi-path attention mechanisms.

Model	Type	FRR	FAR	EER
Morphology [24]	WI	12.39	11.23	11.59
Surroundness [25]	WI	8.33	8.33	-
Chain Code [3]	WD	9.36	7.84	-
Graph Matching [5]	WD	7.7	8.2	-
SigNet-F [18]	WD	-	-	4.63
Single Stream	WI	11.96	7.25	10.0
Double Stream	WI	3.04	8.19	4.86
Our IDN	WI	**2.17**	**5.87**	**3.62**

Table 2. Comparison on CEDAR Dataset (%).

Model	Type	FRR	FAR	Acc
SigNet [8]	WI	13.89	13.89	86.11
Correlated Feature [9]	WI	14.43	15.78	84.90
Texture Feature [27]	WD	33.82	33.82	66.18
Single Stream	WI	12.88	9.60	88.76
Double Stream	WI	6.49	11.23	91.14
Our IDN	WI	**5.24**	**4.12**	**95.32**

Table 3. Comparison on BHSig-B Dataset (%).

4.3. CEDAR Dataset

The CEDAR signature dataset [21] contains signature samples of English names. It is composed of 55 individuals' samples and each individual has 24 genuine and 24 forged signatures. Following other works, we use 50 individuals' samples for training and 5 individuals' samples for test. For each individual, we have 276 reference-genuine pairs and 276 reference-forgery pairs.

We compare our IDN method with other approaches: Morphology [24], Surroundness [25], Chain Code [3], Graph Matching[5], SigNet-F [18], Single Stream, and Double Stream. The Single Stream and Double Stream approaches are as the same definition in Section 4.2.

Table 2 shows the results of different approaches and Fig. 7 (b) shows the ROC curves of the Single Stream, the Double Stream, and our four stream IDN. In the table, WI indicates writer-independent methods which build one same model for any writers and WD means writer-dependent methods which train different models for each writer and often need more samples for training. It should be noted that the writer-dependent methods adopt different training methods from writer-independent methods. We list the writer-dependent methods here as references.

On this dataset, our IDN model outperforms other approaches in all reported evaluation metrics, which proves the strength of our method.

4.4. BHSig-B Dataset and BHSig-H Dataset

BHSig260 dataset [27] contains two subsets: BHSig-B Dataset and BHSig-H Dataset. BHSig-B Dataset contains signature samples of Bengali names. It contains 100 individuals' signature samples. Each individual has 24 genuine signatures and 30 forged signatures. Following other works, we use 50 individuals' samples for training and the rest individuals' samples for test. For each individual, we have 276 reference-genuine pairs and 276 reference-forgery pairs.

BHSig-H Dataset contains signature samples of Hindi names. It contains 160 individuals' signature samples. Each individual has 24 genuine signatures and 30 forged signatures. Following other works, we use 100 individuals' samples for training and the rest individuals' samples for test. For each individual, we have 276 reference-genuine pairs and 276 reference-forgery pairs.

On both the two datasets, we compare our IDN method with other approaches: SigNet [8], Correlated Feature [9], Texture Feature [27], Single Stream, and Double Stream. The Single Stream and Double Stream are as the same definition in Section 4.2.

Table 3 and Table 4 show the results of different approaches on the two datasets, respectively. The performance of the Correlated Feature method [9] here was reported in the work SigNet [8]. Fig. 7 (c) and Fig. 7 (d) show the ROC curves of the Single Stream, the Double Stream, and our four stream IDN. On the two datasets, our IDN model outperforms other approaches by a large margin, which proves the strength of our method.

Model	Type	FRR	FAR	Acc
SigNet [8]	WI	15.36	15.36	84.64
Correlated Feature [9]	WI	15.09	13.10	85.90
Texture Feature [27]	WD	24.47	24.47	75.53
Single Stream	WI	13.39	11.73	87.44
Double Stream	WI	10.44	**8.32**	90.62
Our IDN	WI	**4.93**	8.99	**93.04**

Table 4. Comparison on BHSig-H Dataset (%).

Train / Test	Ours	CEDAR	BHSig-H	BHSig-B
Ours	**90.17**	50.0	57.96	64.53
CEDAR	50.03	**95.98**	50.36	50.01
BHSig-H	50.0	50.0	**93.04**	74.12
BHSig-B	50.0	50.0	74.30	**95.32**

Table 5. Signature verification accuracy of cross-language test (%).

4.5. Cross-Language Test

The datasets used in this work belong to four different languages. We would like to test if signature verification can be done across different languages. Thus, we carried out a cross-language experiment where a model is trained on one dataset and tested on another dataset of a different language. For example, we train a model on the Chi-

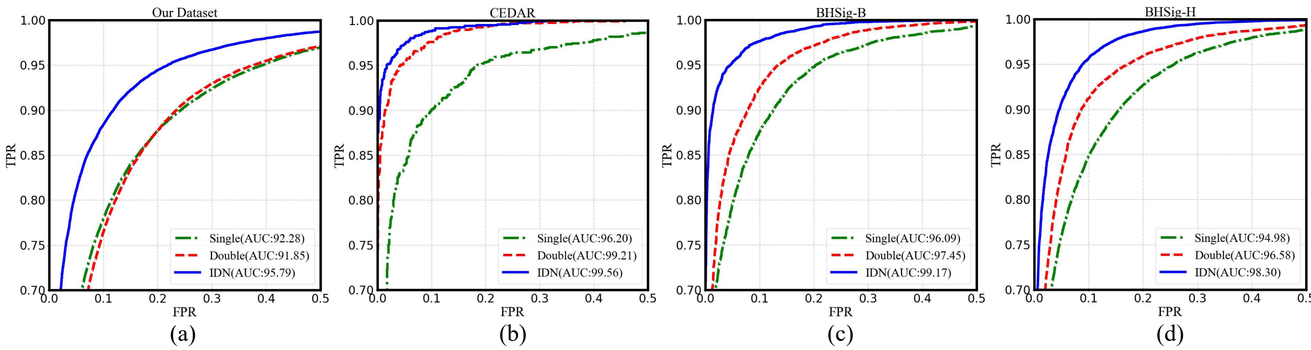

Figure 7. The ROC curve comparison on four datasets. The 'single', and 'double' denote the Single Stream method and the Double Stream method, respectively. (a) Our dataset. (b) CEDAR dataset. (c) BHSig-B dataset. (d) BHSig-H dataset.

nese Signature Dataset and test the model on the BHSig-H Dataset. The training and test data division is the same to the experiments on each independent dataset.

Table 5 shows the accuracy of the cross-language test, where the rows correspond to the training languages and the columns correspond to the testing languages. This table shows that the signature verification performance across languages drops considerably. After all, signatures are closely dependent on the languages and individuals using different languages have different writing habits and styles.

This table also shows that the performance drops of the tests across Bengali and Hindi are not so drastic as other cross-language tests. This can be attributed to the similarity of Bengali and Hindi handwritten signatures in styles and strokes.

5. Conclusion

In this paper, we propose a novel inverse discriminative network (IDN) for writer-independent handwritten signature verification, which contains four weight-shared streams: two discriminative streams that extract the convolutional features of signatures, and two inverse streams that supervise the feature extraction to focus on the signature strokes. An inverse supervision mechanism and a multi-path attention mechanism are used to resolve the sparse information issue in signature verification. In testing, taking the inputs of a reference signature image and a test signature image, our model outputs whether the test signature is genuine or forged. Since there was no proper Chinese signature dataset in the community, we collected a large-scale and challenging Chinese signature dataset. We test our method on the collected Chinese signature dataset and other three signature datasets of different languages. Experiments demonstrate the strength and potential of the proposed method. The future work will focus on the joint system of signature verification and recognition across languages.

Acknowledgement

This research was supported by the grants National Natural Science Foundation of China No. 61876149 and China Postdoctoral Science Foundation 2018M643657.

References

[1] Gabe Alvarez, Blue Sheffer, and Morgan Bryant. Offline signature verification with convolutional neural networks. Technical report, Stanford University, 2016.

[2] H. Baltzakis and N. Papamarkos. A new signature verification technique based on a two-stage neural network classifier. *Engineering Applications of Artificial Intelligence*, 14(1):95–103, 2001.

[3] R. K. Bharathi and B. H. Shekar. Off-line signature verification based on chain code histogram and support vector machine. In *International Conference on Advances in Computing, Communications and Informatics*, 2013.

[4] Long Chen, Hanwang Zhang, Jun Xiao, Liqiang Nie, Jian Shao, Wei Liu, and Tat Seng Chua. Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning. In *CVPR*, 2017.

[5] Siyuan Chen and Sargur Srihari. A new off-line signature verification method based on graph. In *International Conference on Pattern Recognition*, 2006.

[6] Shuhan Chen, Xiuli Tan, Ben Wang, and Xuelong Hu. Reverse attention for salient object detection. In *European Conference on Computer Vision*, 2018.

[7] Sumit Chopra, Raia Hadsell, and Yann Lecun. Learning a similarity metric discriminatively, with application to face verification. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.

[8] Sounak Dey, Anjan Dutta, J. Ignacio Toledo, Suman K Ghosh, Josep Lladós, and Umapada Pal. Signet: Convolutional siamese network for writer independent offline signature verification. *CoRR*, abs/1707.02131, 2017.

[9] Anjan Dutta, Umapada Pal, and Josep Lladós. Compact correlated features for writer independent signature verification. In *International Conference on Pattern Recognition*, 2016.

[10] A. El-Yacoubi, E. J. R. Justino, R. Sabourin, and F. Bortolozzi. Off-line signature verification using hmms and

cross-validation. In *Proceedings of the IEEE Signal Processing Society Workshop*, 2000.

[11] Mounim El Yacoubi, Robert Sabourin, and Flavio Bortolozzi. An off-line signature verification system using hmm and graphometric features. In *IAPR International Workshop on Document Analysis Systems (DAS)*, 2000.

[12] Miguel A. Ferrer, Jesús B. Alonso, and Carlos M. Travieso. Offline geometric parameters for automatic signature verification using fixed-point arithmetic. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(6):993–997, 2005.

[13] Miguel A. Ferrer, Moises Diaz-Cabrera, and Aythami Morales. Synthetic off-line signature image generation. In *International Conference on Biometrics*, 2013.

[14] Julian Fierrez-Aguilar, Loris Nanni, Jaime Lopez-Peñalba, Javier Ortega-Garcia, and Davide Maltoni. An on-line signature verification system based on fusion of local and global information. In *Audio- and Video-Based Biometric Person Authentication*, 2005.

[15] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *NIPS*, 2014.

[16] Yasmine Guerbai, Youcef Chibani, and Bilal Hadjadji. The effective use of the one-class svm classifier for handwritten signature verification based on writer-independent parameters. *Pattern Recognition*, 48:103–113, 2015.

[17] Luiz G. Hafemann, Robert Sabourin, and Luiz S. Oliveira. Writer-independent feature learning for offline signature verification using deep convolutional neural networks. In *International Joint Conference on Neural Networks*, 2016.

[18] Luiz G. Hafemann, Robert Sabourin, and Luiz S. Oliveira. Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognition*, 70:163176, 2017.

[19] Jie Hu, Li Shen, Samuel Albanie, Gang Sun, and Enhua Wu. Squeeze-and-excitation networks. *CoRR*, abs/1709.01507, 2017.

[20] Qin Huang, Chi-Hao Wu, Chunyang Xia, Ye Wang, and C.-C. Jay Kuo. Semantic segmentation with reverse attention. *BMVC*, 2017.

[21] Meenakshi K. Kalera, Sargur Srihari, and Aihua Xu. Offline signature verification and identification using distance statistics. *International Journal of Pattern Recognition and Artificial Intelligence*, 18:1339–1360, 2004.

[22] Hurieh Khalajzadeh, Mohammad Mansouri, and Mohammad Teshnehlab. Persian signature verication using convolutional neural networks. *International Journal of Engineering Research and Technology*, 1, 2012.

[23] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *NIPS*, pages 1097–1105, 2012.

[24] Rajesh Kumar, Lopamudra Kundu, Bhabatosh Chanda, and J. D. Sharma. A writer-independent off-line signature verification system based on signature morphology. In *International Conference on Intelligent Interactive Technologies and Multimedia*, 2010.

[25] Rajesh Kumar, J. D. Sharma, and Bhabatosh Chanda. Writer-independent off-line signature verification using surroundedness feature. *Pattern Recognition Letters*, 33(3):301–308, 2012.

[26] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems Man and Cybernetics*, 9(1):62–66, 1979.

[27] Srikanta Pal, Alireza Alaei, Umapada Pal, and Michael Blumenstein. Performance of an off-line signature verification method based on texture features on a large indic-script signature dataset. In *IAPR Workshop on Document Analysis Systems*, 2016.

[28] Hannes Rantzsch, Haojin Yang, and Christoph Meinel. Signature embedding: Writer independent offline signature verification with deep metric learning. In *Advances in Visual Computing*, 2016.

[29] Javier Ruiz-Del-Solar, Christ Devia, Patricio Loncomilla, and Felipe Concha. Offline signature verification using local interest points and descriptors. In *Iberoamerican Congress on Pattern Recognition: Progress in Pattern Recognition, Image Analysis and Applications*, 2008.

[30] Yasmine Serdouk, Hassiba Nemmour, and Youcef Chibani. Combination of oc-lbp and longest run features for offline signature verification. In *International Conference on Signal-Image Technology and Internet-Based Systems*, 2014.

[31] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.

[32] Fei Wang, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, and Xiaoou Tang. Residual attention network for image classification. In *CVPR*, 2017.

[33] Mustafa Berkay Yilmaz and Kağan Öztürk. Hybrid user-independent and user-dependent offline signature verication with a two-channel cnn. In *CVPR Workshops*, 2018.

[34] Mustafa Berkay Yilmaz, Berrin Yanikoglu, Caglar Tirkaz, and Alisher Kholmatov. Offline signature verification using classifier combination of hog and lbp features. In *International Joint Conference on Biometrics*, 2011.

[35] Mustafa Berkay Yilmaz and Berrin Yankoğlu. Score level fusion of classifiers in off-line signature verification. *Information Fusion*, 32:109–119, 2016.

[36] Sergey Zagoruyko and Nikos Komodakis. Learning to compare image patches via convolutional neural networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[37] Zehua Zhang, Xiangqian Liu, and Yan Cui. Multi-phase offline signature verification system using deep convolutional generative adversarial networks. In *International Symposium on Computational Intelligence and Design*, 2017.

[38] Elias N. Zois, Ilias Theodorakopoulos, and George Economou. Offline handwritten signature modeling and verification based on archetypal analysis. In *The IEEE International Conference on Computer Vision*, 2017.