Offline Signature Verification Using a 2D Attention Encoder-Decoder Network

Zhuohui Chen
Macau University of Science and Technology, Macau, China

Zhicong Lin
Zhuhai Kingsware Information Technology Co. Ltd., Zhuhai, China

Chunlian Lin
Zhuhai Kingsware Information Technology Co. Ltd., Zhuhai, China

Mingchen Wang
Macau University of Science and Technology, Macau, China

Ling Chen*
Beijing Institute of Technology - Zhuhai, Zhuhai, China

ABSTRACT

Verifying an individual's Chinese handwritten signature is a vital biometric technology that is widely used in banking, finance, and legal business. Forged signatures for the purpose of deception endanger these industries' interests. As a result, this paper proposes a network based on 2D Attention to verify the authenticity of the signature. In this paper, we have developed a Chinese handwritten signature dataset (CNSig) and proposed an offline signature verification network (att-OfSVNet) based on 2D Attention. The att-OfSVNet model includes two weight-sharing Encoders and Decoders. The two weight-sharing Encoders receive the inverted genuine and the inverted test signature image, and the Decoder reduces the dimension and concatenates the two extracted feature images. We use 2D Attention to fuse the features extracted by the Encoder and Decoder, which minimizes the information loss in the convolutional layers during the extraction process and enhances the effect of feature extraction by the Encoder. The experimental results show that our att-OfSVNet achieves satisfactory performance on other handwritten signature datasets in three different languages: CEDAR, BHSig-B, and BHSig-H, and it also demonstrates good generalization ability in cross-lingual tests.

CCS CONCEPTS

• **Network Architectures**; • **Network Algorithms**; • **Artificial Intelligence**;

KEYWORDS

Offline signature verification, 2D Attention, Chinese handwritten signature dataset

ACM Reference Format:
Zhuohui Chen, Zhicong Lin, Chunlian Lin, Mingchen Wang, and Ling Chen*. 2023. Offline Signature Verification Using a 2D Attention Encoder-Decoder Network. In *2023 the 12th International Conference on Networks, Communication and Computing (ICNCC) (ICNCC 2023), December 15–17, 2023, Osaka, Japan.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3638837.3638880

ICNCC 2023, December 15–17, 2023, Osaka, Japan

ACM ISBN 979-8-4007-0926-5/23/12
https://doi.org/10.1145/3638837.3638880

1 INTRODUCTION

Verifying the authenticity of signatures is a crucial challenge when a huge number of important financial, business, and legal documents are signed every day around the world today. Fraudulent use of signatures for profit-driven purposes occurs from time to time, posing a serious risk of violations for financial institutions, banks, and government agencies. At the same time, with the increasing number of commercial documents, it is difficult to guarantee the speed and accuracy of manual identification. Therefore, it is increasingly important to develop an automatic, accurate, and efficient signature verification technology.

Although the past few decades have witnessed the development and progress of signature verification, there are still several challenges in the Chinese domain. Firstly, the lack of suitable and public Chinese handwritten signature datasets hinders the research and application of Chinese offline signature verification. Secondly, Chinese strokes are more complex than Western languages, which makes higher requirements for feature extraction of the network. Thirdly, most people's signature style is casual, which makes the same person's signature style distinctly different on different occasions. Moreover, some intentionally forged signatures are very similar to genuine signatures.

In recent years, the field of computer vision has witnessed a growing significance of offline signature verification, thanks to the continuous advancements in deep learning. While Wei et al. [1] and Lu et al. [2] have made notable contributions by collecting large-scale Chinese handwritten signature datasets, unfortunately, neither of these datasets has been made publicly available. Consequently, due to the absence of an accessible Chinese signature dataset, we have taken the initiative to develop a standardized Chinese handwritten signature dataset, termed CNSig, utilizing OpenCV technology and the random elastic distortion method. The CNSig dataset contained 16,720 signature images from 836 writers, including 8,360 genuine signature images and 8,360 forged signature images.

In addition, 2D Attention was proposed by Chattopadhyay et al. [3], and they used 2D Attention to focus on signature features to realize offline signature verification. 2D Attention, a key module in the training pipeline of offline signature verification networks, also focuses on the visually important part of the signature in a regional manner. Therefore, we propose an offline signature verification network (att-OfSVNet) based on 2D Attention. Att-OfSVNet consists of two weight-sharing Encoders, two weight-sharing Decoders, and six 2D Attention modules. Two weight-sharing Encoders in att-OfSVNet receive the genuine and test signature images after black-and-white inversion, respectively, and then extract the features of the signature images to obtain the genuine and test signature feature maps. Two weight-sharing Decoders each receive two feature maps and perform dimensionality reduction on the two feature maps to obtain two vectors. Next, the two vectors are concatenated, passed through a fully connected layer, and finally input into a Sigmoid activation function to calculate the probability, to identify whether the test signature is genuine or forged. 2D Attention inputs a 3D tensor and a 1D tensor, and outputs a 1D tensor with the same shape, so att-OfSVNet can connect the output of the Encoder with the output of the Decoder via 2D Attention.

Additionally, the att-OfSVNet employs 2D Attention to effectively integrate the features extracted from both the Encoder and Decoder. This fusion process aims to minimize information loss during the convolution layer extraction, thereby enhancing the feature extraction capability of the Encoder. Consequently, the performance of the offline signature verification network is improved, leading to more accurate verification of offline handwritten signatures. Furthermore, we validate the robustness and generalization ability of our network by utilizing handwritten signature datasets in three different languages: CEDAR, BHSig-B, and BHSig-H. The results demonstrate that our network exhibits strong generalization across various languages. Moreover, a comparative analysis between our att-OfSVNet model and other offline signature verification networks highlights the superior performance of our network.

The rest of the paper is organized as follows: Section II reviews the current situation of offline signature verification in detail. Section III introduces the Chinese handwritten signature dataset we developed. Section IV describes in detail our proposed offline signature verification network (att-OfSVNet) based on 2D Attention. Section V is our experiment on att-OfSVNet, and we use CEDAR, BHSig-B, and BHSig-H to compare and test the performance of the offline signature verification network. Section VI is the conclusion of the whole paper. In summary, our contributions in this study are as follows: 1) A standardized Chinese handwritten signature dataset (CNSig) is developed, which is helpful for research and application in the field of Chinese signature verification and provides a more comprehensive benchmark for offline verification network. 2) We propose an offline signature verification network based on 2D Attention (att-OfSVNet), which improves the accuracy of offline signature verification by fusing the features extracted by the Encoder and Decoder with 2D Attention, losing as little information as possible from the convolutional layer during extraction, and strengthening the effect of the Encoder in extracting features. 3) We evaluate the performance of att-OfSVNet using three different language handwritten signature datasets: CEDAR, BHSig-B, and BHSig-H and other offline signature verification networks, which demonstrate that att-OfSVNet has strong generalization ability, and its performance is better than other offline signature verification networks.

2 RELATED WORKS

The initial researchers in offline signature verification used machine learning to solve this task. They used machine learning algorithm on an English handwritten signature CEDAR dataset [4] proposed in 2004 and achieved good results [4]-[15]. Among them, there are methods of extracting image features based on gradient, structure, and contour [4], and extracting signature image features. This method has been successfully applied to the CEDAR datasets. There are also methods that extract signature image features based on the description of the signature envelope and geometric signature. The features are calculated by using a 16-bit fixed point algorithm and tested by using classifiers such as Hidden Markov Model, Support Vector Machine (SVM), and Euclidean distance, to realize offline automatic signature verification [5]. Moreover, there are methods based on direction contours to extract features [6], [7]. [6] used exterior contours and shape features for offline signature verification. The signature's exterior contours are determined by the small gaps between different characters in the signature, and the signature's shape features are extracted based on Zernike moments. [7] compared offline signatures using the graph matching method, where each compared signature is regarded as a set of points, and the graph matching contains a mapping function between a deformation metric and a set of points. Besides the above methods, [8] proposed a general wide baseline matching method to replace the signature features extracted based on direction contours and geometric signature methods. They detected and extracted signature image features by using local interest points and classified them via Bayesian classifier to realize offline signature verification. The authors in [9] presented a method to extract offline signature image features based on morphological features. They used the feature analysis technique of multilayer perceptron to extract the morphological features of signatures and used Support Vector Machine with 10-fold cross-validation for offline signature verification. [10] put forward the signature-based surroundedness features. Their extracted features contain both shape and texture features, which are better than the offline signature image features extracted by most previous researchers. An offline signature verification method based on chain code histogram features proposed by [11]. They selected four directional chain code histograms of each network on the signature image contours, augmented the extracted features with Laplacian of Gaussian filter, and used Support Vector Machine as the classifier for offline signature verification.

However, in reality, the number of writers is limited, so [12] based on writer-independent SVM for offline signature verification, which enables the model to perform well even with fewer writers. [13] considered the compact correlated features in the signature image and the spatial information of the compact correlated features plays a key role in the geometric shape of the signature, to realize the offline signature verification. [14] based on the texture feature extraction technique, using Local Binary Patterns and Uniform Local Binary Patterns to extract offline signature image features and using the Nearest Neighbor method to calculate the similarity of signature image features for offline signature verification. The authors in [15] suggested a vector classifier, Artificial Immunity Recognition System, based on the quad-tree histogram of templates structure to extract signature image features and combined with

SVM to classify offline signature images. However, due to the small scale of the CEDAR dataset, the network learns fewer features, resulting in low generalization ability of the network, which cannot play a good signature verification effect on other languages' datasets, so Pal et al. [14] in 2016 published a large BHSig260 dataset of handwritten signatures in Bengali and Hindi, and based on texture features to images for feature extraction, which made greater progress in the task of offline signature verification.

With the gradual development of deep learning and the public availability of large datasets, many researchers have started to use Convolutional Neural Networks (CNN) to solve the offline signature verification problem [1]-[3], [16]-[26], [29], [30], [32]-[36], [38]. Soleimani et al. [16] introduced a new classification method for deep multi-task metric learning to implement offline signature verification. Zhang et al. [17] utilized deep convolutional generative adversarial networks to learn image features instead of hand-crafted signature features. In addition, Hafemann et al. [18] found that in offline signature verification, due to the loss of dynamic information during the signature process, it is difficult to design a suitable feature extractor to distinguish whether a test signature is a genuine signature or not. Consequently, Hafemann et al. [19] further improved the structure of the CNN for capturing visual cues to distinguish real signatures from forged ones, improving the accuracy of the offline signature authentication task. Subsequently, due to the proposal of spatial pyramid pooling [20], Hafemann et al. [21] adapted the CNN structure by using spatial pyramid pooling to learn a fixed-size image dimension from a variable-size signature. Huang et al. [22] proposed a method called Snapshot Ensemble of CNNs that converges to several local minima along the network optimization path on a mainstream CNN network, which can be used to complete offline signature verification tasks.

Meanwhile, Dey et al. [23] introduced a Siamese convolutional neural network to compare the distance between signatures for the offline signature verification task. Inspired by Dey et al. [23], Xing et al. [24] regarded feature extraction and metric learning as two separate and distinct methods and proposed a convolutional Siamese neural network for offline signature verification method. In the meantime, Younesian et al. [25] and Mersa et al. [26] trained offline signature verification networks using pre-trained ResNet [31] on ImageNet dataset [27] and handwriting classification task dataset [28]–[30], respectively. Masoudnia et al. [32] investigated loss functions of offline signature verification networks and presented a dynamic multi-loss function to enable multi-representation learning. Wei et al. [1] built an Inverse Discriminative Networks (IDN) as a Writer Independent (WI) system to determine whether a test signature is genuine or forged. While Lu et al. [2] suggested a method to cut and compare signature images and use an adaptive distance fusion module to fuse the cut graphic areas for offline signature verification.

However, most of the above methods have the problem of overfitting, so Lin et al. [33] put forward a two-channel CNN network by splicing the channel dimension method in the genuine signature and the test signature. Furthermore, previous deep learning-based methods, either paired or triplet samples, were unaware of the internal changes in signatures, and had low training efficiency since they only considered the distance between two points. Therefore, Zhu et al. [34] suggested a new point-to-set metric for offline signature verification. Liu et al. [35] posed a metric learning method to verify offline signatures, which is applicable for both writer-dependent and writer-independent scenarios. Parcham et al. [36], on the other hand, designed a training mechanism, which enables an offline signature verification network to be trained with two signature images at the same level simultaneously.

Although these methods achieve better performances on offline signature verification datasets, they omit the loss of feature information by the convolutional layer, which may hinder the network from learning effective trajectory features and lead to diminished network performance. Transformer [37] discarded convolutional neural networks and raised a novel structure of Encoder-Decoder, using the Attention module instead of the recurring convolutional module. Li et al. [38] proposed a holistic-part unified model based on the Transformer framework for offline signature verification in 2022, following IDN. Chattopadhyay et al. [3] also presented a 2D Attention focused signature feature based on Attention, which improves the accuracy of offline signature verification. However, for the complicated Chinese signature verification problem, there is still no researcher who has used an efficient means to solve it.

Up to now, handwritten signature datasets in many languages are publicly available, such as CEDAR [4], MCYT-75 [28], BHSig260 [14], GPDS Synthetic [29], and UTSig [30]. Nevertheless, there is still no large-scale Chinese signature dataset, which leads to a serious hindrance in the research and application of Chinese offline signature verification. Until 2019, Wei et al. [1] collected the first large-scale Chinese handwritten signature dataset, but unfortunately, they did not release this dataset to the public. Lu et al. [2] collected a large-scale Chinese handwritten signature dataset, ChnSig, but also did not make this dataset publicly available. Therefore, we developed a standardized large-scale Chinese handwritten signature dataset (CNSig) to promote research and application in the domain of Chinese signature verification.

In addition, we propose an offline Chinese signature verification network (att-OfSVNet) based on 2D Attention, which consists of Encoder and Decoder, retaining the method of using convolutional layers to extract features and losing as little information as possible from the convolutional layers during the extraction process. In contrast to the use of 2D Attention to connect the Decoder and Encoder in Chattopadhyay et al. [3], our method also solves the problem that the convolution layer loses information during feature extraction, but since 2D Attention is set before the Decoder, it does not work to fuse the features obtained by the Encoder and Decoder. This paper uses 2D Attention to fuse the features extracted by the Encoder and Decoder, which not only can solve the information lost in the process of feature extraction by the convolutional layer, but also can combine with the Decoder to strengthen the features extracted by the Encoder, thus minimizing the information lost in the extraction process of the convolutional layer, improving the performance of the network, and verifying the offline handwritten signature more accurately.

3 CHINESE SIGNATURE DATASET

In this section, since there is no publicly available Chinese signature dataset, we create a large-scale Chinese Signature (CNSig) dataset

Figure 1: Handwritten signature form: (a) handwritten genuine signature, (b) first imitation signature, (c) second imitation signature.

and distort the forged signatures with random elasticity. The CNSig dataset has a total of 25,080 signature images from 836 writers.

3.1 Tables

We first designed a genuine signature form for 836 writers and two forged signature forms, each with 10 signatures. To ensure that offline signature datasets are challenging, for genuine signatures, we distributed 80 to calligraphers and 756 to volunteers for handwritten signatures. For the forged signatures, we distributed 80 signatures to the calligrapher for imitation and the other 756 signatures were imitated by two random volunteers. Calligraphers who imitate can provide signature samples with a high level of professional skill and artistic style, while random volunteers who imitate can provide more general signature samples, thus offering coverage across various difficulty levels. Different level of difficulty among signatures and the 80 signatures from the calligraphers are highlighted on the dataset as hard cases. During the imitation process, the volunteers imitated the signatures as close as possible to the genuine ones. The calligraphers, on the other hand, carefully observed and studied the style, handwriting, and other characteristics of the genuine signatures before imitating them. Additionally, all signatures were captured on different dates. By collecting signature samples from different dates, a more extensive dataset can be obtained, covering signature styles and variations from different time periods.

After collection, we received 836 writer's handwritten signature forms in total, and an example of one writer's signature is shown in Figure 1.

3.2 Data Preprocessing

We scanned the handwritten signature form into a PDF and made it into a grayscale image. The image resolution is 1655×2340, and the form size is 1160×1150 where the signatures are written. Then we use OpenCV technology to remove the lines of the form, only leaving the signature. Next, the signature is extracted from each image by OpenCV. For the same writer, we randomly extracted 5 forged signatures from two forged signature forms, respectively. The CNSig dataset has a total of 16,720 images from 836 writers, including 8,360 genuine signature images and 8,360 forged signature images. Each writers have 10 genuine signatures and 10 forged signature images, and the example of the dataset is shown in Figure 2.

Since the diversity of forged signatures is not sufficient, we apply random elastic distortion to forged signatures, thus increasing the diversity of forged signatures, and improving the performance of the network.

Random elastic distortion is often used in non-fixed image scenarios, in short, scenarios that may be influenced by small external factors and some areas of the image have simply changed, such as outdoor flower identification, etc. The use of random elastic distortion is not common in the field of offline signature verification, but it is well suited for offline signature verification to employ a small range of random elastic distortion. Since the signatures of the writers are not identical at different times, there may be slight differences, but this does not affect the fact that the signatures are genuine, so the use of random elastic distortion can play a significant role in changing this situation. The size of the random elastic distortion will control the degree of signature distortion. Smaller sizes will lead to a more obvious distortion of the signature, while larger sizes will result in a finer distortion of the signature. Therefore, for each writer's forged signature, we use a smaller size of height and width 3 for random elastic distortion, and the result of elastic distortion is shown in Figure 3.

The comparative effect between the original dataset and the dataset after elastic distortion is shown in Figure 4. It can be observed that random elastic deformation amplifies the differences between genuine signatures and forged signatures, thereby enhancing the accuracy of offline signature authentication networks. By enlarging the discrepancies between genuine signatures and

Figure 2: Example of CNSig dataset.

Figure 3: Example of CNSig dataset after elastic distortion.

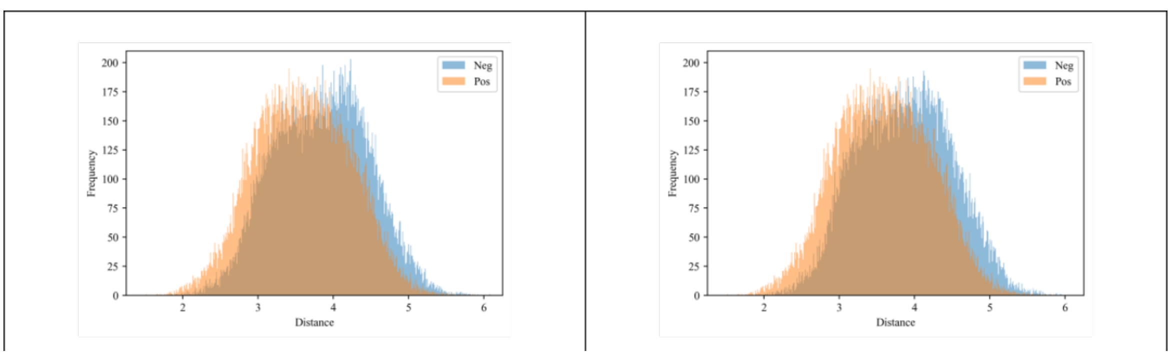


Figure 4: The comparative effect between the original dataset and the dataset after elastic distortion: (a) the original dataset, (b) the dataset after elastic distortion.

forged signatures, the network can more easily distinguish between genuine and forged signatures.

4 2D ATTENTION-BASED OFFLINE SIGNATURE VERIFICATION NETWORK

In this section, we set up an Offline Signature Verification Network (att-OfSVNet) based on 2D Attention. Att-OfSVNet takes the genuine and test signature images as input, and contains two Encoders, two Decoders, and six 2D Attention modules.

4.1 Architecture

The 2D Attention-based offline signature verification network is a novel network defined by us and its structure is shown in Figure 5. The att-OfSVNet first inputs the black-and-white inverted (black background and white foreground) genuine and test signature images into two weight-sharing Encoders, respectively, and obtains two feature maps. The encoding procedure is shown in the following equation:

$$C_r^{(1)}, C_r^{(2)}, C_r^{(3)}, P_r = \text{Encoder}(x_r) \quad (1)$$

$$C_t^{(1)}, C_t^{(2)}, C_t^{(3)}, P_t = \text{Encoder}(x_t) \quad (2)$$

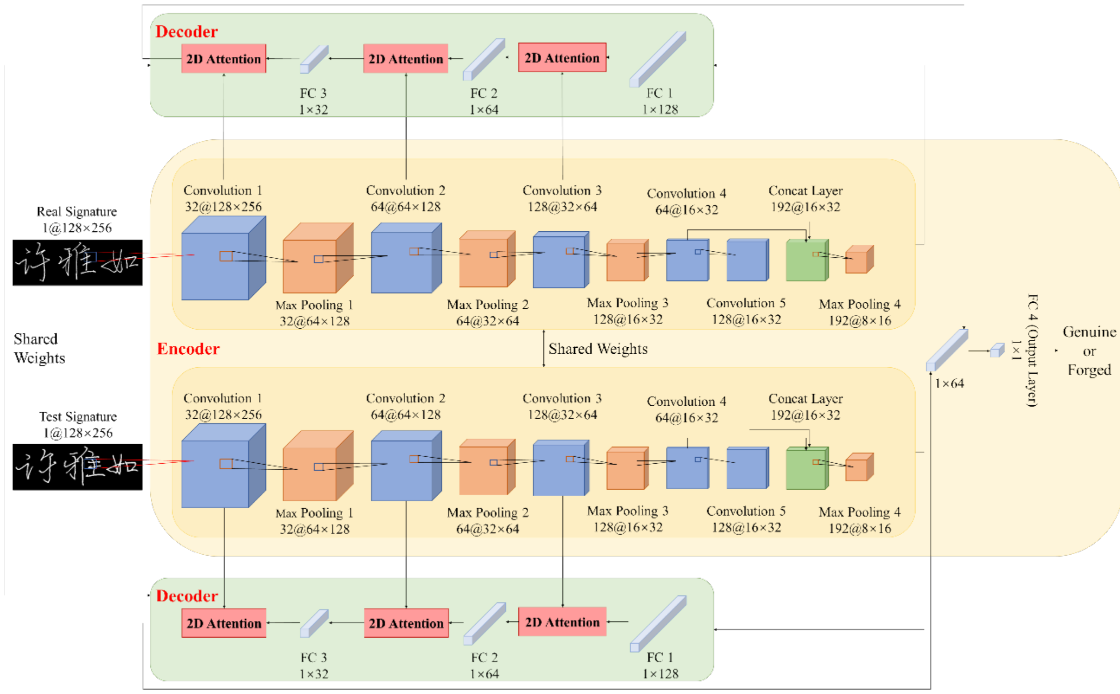


Figure 5: The structure of att-OfSVNet.

where $C^{(i)}$ $(i = 1, 2, 3)$ is the computed result of the i-th convolutional layer in the Encoder, and P is the calculated result of the last max-pooling layer in the Encoder, which is also the signature feature map.

Then the convolution result and the feature map are entered into the two weight-sharing Decoders to obtain two vectors with the length of 32, and the decoding process is shown in the following equation:

$$D_r = \text{Encoder}(C_r^{(1)}, C_r^{(2)}, C_r^{(3)}, P_r) \tag{3}$$

$$D_t = \text{Encoder}(C_t^{(1)}, C_t^{(2)}, C_t^{(3)}, P_t) \tag{4}$$

Finally, two vectors with the length of 32 are concatenated, a full-connected layer with the number of hidden neurons of 1 is taken as the output layer. The concatenated vectors are delivered to the output layer, and the probability that the test signature is genuine or forged is estimated by a Sigmoid activation function with the following equation:

$$D = \text{Concat}(D_r, D_t) \tag{5}$$

$$O = \text{Sigmoid}(WD + b) \tag{6}$$

where Concat is the concatenation operation of two vectors, W is the weight of the output layer, and b is the bias of the output layer.

Att-OfSVNet utilizes (1)-(6) to verify whether a test signature is a genuine signature, thus realizing offline signature verification. Compared with other offline signature verification networks, our network only uses 10 genuine signatures and 10 forged signatures for training, which enables the network to learn the features of the signature images more intensively and to determine more accurately whether the test signature is someone's own signature.

4.2 Dual-stream Encoder

The dual-stream Encoder of att-OfSVNet consists of two weight-sharing encoders, which can serve the purpose of feature extraction, and the structure is shown in Table 1. Encoder consists of five convolutional layers and four pooling layers. The first convolutional layer receives a black-and-white inverted signature image of input size 128 × 256 and has 32 convolutional kernels, thus expanding the feature map from a single channel to 32 channels. The second pooling layer takes the output of the first convolutional layer as input, using a receptive field with the size of 2 × 2 and the step size is 2 pixels, thereby cropping the feature map. One convolutional layer and one max-pooling layer are used together as a module, which is repeated three times until the third max-pooling layer. The number of convolution kernels for the first three convolution layers is 32, 64, and 128, increasing gradually. The output of the third pooling layer is used as input to the fourth and fifth convolutional layers, respectively. The fourth convolutional layer has 64 convolution kernels, which reduces the number of channels to 64. While the fifth convolutional layer has 128 convolution kernels, the number of channels is increased to match the input of the third convolutional layer. Subsequently, the results of the fourth and the fifth convolutional layers are combined in the channel dimension, which means that the number of channels becomes 192. Finally, it passes through a max-pooling layer with a 2 × 2 receptive field and the step size is 2 pixels, then the output of the fourth max-pooling layer becomes the output of the Encoder. All convolutional layers are followed by a BN layer and ReLU activation function, and all convolutional layers have a 3 × 3 kernel size and a 1-pixel step size.

Table 1: The structure of encoder

Layers	Size	Parameters
Convolution 1	$32 \times 3 \times 3$	stride = 1, padding = 1
Max Pooling 1	2×2	stride = 2
Convolution 2	$64 \times 3 \times 3$	stride = 1, padding = 1
Max Pooling 2	2×2	stride = 2
Convolution 3	$128 \times 3 \times 3$	stride = 1, padding = 1
Max Pooling 3	2×2	stride = 2
Convolution 4	$64 \times 3 \times 3$	stride = 1, padding = 1
Convolution 4	$128 \times 3 \times 3$	stride = 1, padding = 1
Concat Layer		
Max Pooling 4	2×2	stride = 2

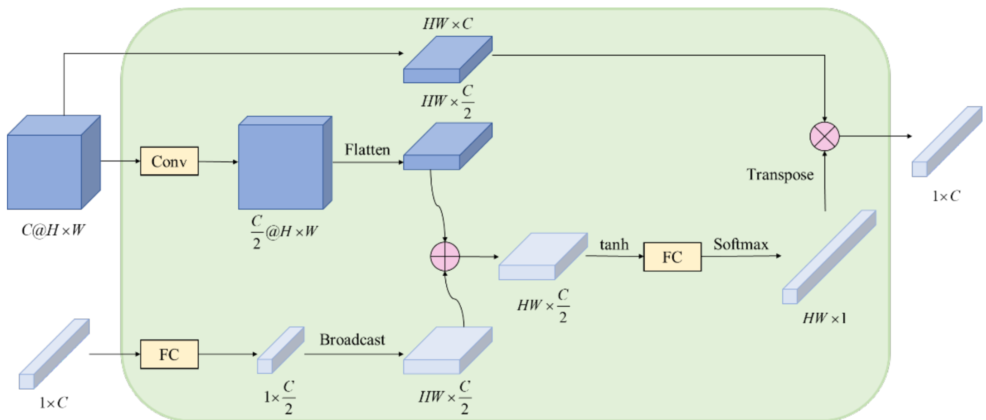


Figure 6: The structure of 2D Attention.

4.3 2D Attention Modules

2D Attention is the key module in the att-OfSVNet network, and we refer to the 2D Attention of Chattopadhyay et al. [3], and the structure is shown in Figure 6. Unlike Chattopadhyay et al. [3], we directly input the output of the convolutional layer in the encoder and the output of the fully connected layer in the decoder to fuse the information of the encoder and decoder. The input of 2D Attention is a 3D tensor and a 1D vector. First, the 3D tensor is input to a convolutional layer with a 3×3 kernel size and the filling is 1, so that the number of channels is halved. Then, the height and width dimensions of the image are combined and downscaled into a 2D tensor. At the same time, the 1D vector is input into a fully connected layer, thus halving the length, and then a broadcast operation is performed on it. The summed tensor is fed into the tanh activation function, the fully connected layer, and the Softmax activation function in turn, and a result is obtained that can be matrix multiplied with the transposed 3D tensor. Finally, a vector with the same shape as the input 1D vector is outputted to complete the 2D Attention, which uses the following formulation given in to fuse the Encoder with the Decoder:

$$M = \tanh(\mathrm{Flatten}(\mathrm{Conv}(C)) + W_1 V_1 + b_1) \tag{7}$$

$$S = \mathrm{Softmax}(W_2 M + b_2) \tag{8}$$

$$V_2 = \mathrm{Flatten}(C) S^{\mathrm{T}} \tag{9}$$

where Flatten flats the image height and width dimensions of the 3D tensor, Conv is the convolution operation, C is the 3D tensor input, V_1 is the 1D tensor input, W_i $(i = 1, 2)$ is the weight of the i-th fully connected layer, and b_i $(i = 1, 2)$ is the bias of the i-th fully connected layer.

This paper fuses the convolutional layer outputs of the first three layers in the Encoder and the three fully connected layers in the Decoder by 2D Attention correspondingly. It feeds the encoder's feature map information in trajectory to the decoder, avoiding the phenomenon of gradient explosion or gradient disappearance compared to placing the decoder directly after the encoder. This aids in extracting richer and more detailed signature features, including stroke shapes, line textures, and spatial structures, thus enhancing the network's understanding and representation capability of signatures. Retaining as much information as possible improves the network's adaptability to variations in input. Offline handwritten signatures may be influenced by factors such as writing speed and stroke intensity, which can result in shape and appearance variations in signatures. By preserving more information, the network can better capture these variations and accurately verify signatures with different variants. By minimizing the loss of information, the network can comprehensively analyze and compare the similarities

Table 2: The structure of decoder

Layers	Size	Parameters
Fully Connected 1	128	
2D Attention 1	2×2	hidden size = 64
Fully Connected 2	64	
2D Attention 2	2×2	hidden size = 32
Fully Connected 3	32	
2D Attention 3	2×2	hidden size = 16

and differences between signature samples. This helps improve verification performance, enabling the network to distinguish genuine signatures more accurately from forged ones.

4.4 Dual-stream Decoder

The dual-stream Decoder of att-OfSVNet consists of two weight-sharing decoders, which are used to extract the image features of the real signature and the signature to be authenticated respectively, the structure of which is shown in Table 2. Decoder consists of three fully connected layers and three 2D Attention layers for dimensionality reduction of feature maps. The Decoder first inputs the output of the Encoder into a fully connected layer with 128 hidden layer neurons, and then puts it into the 2D Attention together with the result of the third layer convolution in the Encoder for fusion. The fused result is then input to a fully connected layer with 64 hidden layer neurons, followed by fusing with the second layer convolution result in 2D Attention. Finally, the fused result is input to a fully connected layer with 32 hidden layer neurons, and then put into 2D Attention together with the third convolutional layer output for fusion, and the fused outcome is the output of the Decoder. The output of the decoder is a more understandable and comparable representation that contains crucial feature information of the signature. It can be used for further signature verification and comparison. By comparing the feature vector of an input signature with the feature vector of a known genuine signature, verification judgments can be made. The role of the decoder is to provide decoding services, transforming abstract feature maps into interpretable and comparable feature vectors, thereby enabling accurate verification of offline handwritten signatures.

4.5 Loss Function

Since the offline signature verification, we study is a binary classification task for test signature images, we adopt the binary cross-entropy function as the loss function of att-OfSVNet, which is formulated as follows:

$$\text{Loss}(y, \hat{y}) = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \tag{10}$$

where N is the sample size, y is the true category, and $\hat{y}$ is the predicted category.

5 EXPERIMENTS

5.1 Experimental Setup

In the experimental process of this paper, the hardware environment uses an Intel Core i9-10900X CPU processor, 3.70 GHz 64 GB RAM, and two NVIDIA GeForce RTX 2080 Ti graphics processing units. The operating system adopts Windows 10 software system, and the model is trained and tested on PyTorch 1.81 framework [39].

The network parameters in Table 3 are used to train att-OfSVNet and other compared offline signature verification networks in all experiments. The parameter quantity of att-OfSVNet is 3,505,268, and the storage capacity is 13.37MB. The model is small enough. Model lightweight can provide higher efficiency, faster inference speed, smaller model size and better adaptability, enabling the model to run effectively in a resource-constrained environment and meet real-time performance, low power consumption and high-performance requirements. We train the network for 100 epochs with and 32 samples per batch and set the image size to 128×256. Using Adam [40] together with the cosine annealing decay method [41] to update the learning rate, and setting the learning rate for stopping the update to 0.1, the equation is as follows:

$$\text{lr} := \frac{1}{2}[1 - \cos(\frac{\text{lr}}{E}\pi)](\text{lr}_f - 1) + 1 \tag{11}$$

where lr is the learning rate, E is the total number of training epochs, and lr_f is the learning rate of stopping updates.

5.2 Evaluation Metrics

This paper uses three metrics, False Rejection Rate (FRR), False Acceptance Rate (FAR), Accuracy (Acc), and Equal Error Rate (EER), to evaluate the performance of att-OfSVNet and uses this criterion to compare other offline signature verification networks. The formula is as follows:

$$\text{FRR} = \frac{\text{FN}}{\text{TP} + \text{FN}} \tag{12}$$

$$\text{FAR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{13}$$

$$\text{Acc} = \frac{\text{TP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{14}$$

where True Positive (TP) means the number of positive samples identified correctly. True Negative (TN) means the number of negative samples identified correctly. False Positive (FP) is the number of positive samples identified incorrectly. False Negative (FN) is the number of negative samples identified incorrectly. EER is the threshold point where the FAR and the FRR are closest to each other.

Table 3: Training hyper-parameters of att-OfSVNet

Epochs (E)	Batch size	Optimizer	Initial learning rate (lr)	Weight Decay
100	32	Adam	0.01	1e-6

5.3 Signature Datasets

5.3.1 CNSig Dataset. Our CNSig dataset is a Chinese Signature dataset, which includes 836 writers. For genuine signatures, each writer has 10 genuine signature images. For forged signatures, each writer has 10 genuine signature images. Totally, this dataset includes 8,360 genuine signature images and 8,360 forged signature images.

Regarding genuine signatures, we two-by-two combine each genuine signature with the remaining genuine signatures as a positive sample, for a total of 836 × $C2$ 10 = 37620 pairs. As for forged signatures, we match each forged signature with all genuine signatures. Then, to avoid the model learning wrong prior information, we balanced the sample size of genuine and forged signature pairs by randomly selecting the same $C2$ 10 pairs as the positive sample out of 100 negative pairs as the negative sample, with a total of 836 × $C2$ 10= 37620 pairs.

5.3.2 CEDAR Dataset. CEDAR dataset [4] is an English signature dataset, which involves 55 writers. For genuine signatures, each writer has 24 genuine signature images. For forged signatures, they collected 24 forged signature images from about 20 skilled imitators respectively, and the remaining signature images were forged by average imitators. This dataset consists of a total of 55 × 24 = 1320 genuine signature images and 55 × 24 = 1320 forged signature images.

Similarly, for genuine signatures, we two-by-two combine each genuine signature with the remaining genuine signatures as a positive sample, totaling 55 × $C2$ 24 = 15180 pairs. For forged signatures, we pair each forged signature with all genuine signatures and randomly select $C2$ 24 pairs out of 576 negative pairs as a negative sample, with a total of 55 × $C2$ 24 = 15180 pairs.

5.3.3 BHSig-B Dataset and BHSig-H Dataset. BHSig260 dataset [14] is a Bengali and Hindi signature dataset that includes 260 writers and is divided into two sub-datasets, BHSig-B, and BHSig-H datasets. The BHSig-B dataset includes 100 writers. For genuine signatures, each writer has 24 genuine signature images. For forged signatures, each writer has 30 images of skillfully imitated forged signatures. This dataset contains a total of 100 × 24 = 2400 genuine signature images and 100 × 30 = 3000 forged signature images. The BHSig-H dataset includes 160 writers. Each writer has 24 genuine signature images. As for forged signatures, each writer has 30 images of skillfully imitated forged signatures. This dataset contains a total of 160 × 24 = 3840 genuine signature images and 160 × 30 = 4800 forged signature images.

Likewise, in the BHSig-B dataset, for genuine signatures, we two-by-two combine each genuine signature with the remaining genuine signatures as positive samples, making a total of 100 × $C2$ 24 = 27600 pairs. For forged signatures, we pair each forged signature with all genuine signatures and randomly select $C2$ 24 pairs out of 720 pairs of negative samples as negative samples, totaling 100 × $C2$ 24 = 27600 pairs.

Regarding the genuine signatures in the BHSig-H dataset, we two-by-two combine each genuine signature with the remaining genuine signatures as positive samples, with a total of 160 × $C2$ 24 = 44160 pairs. For forged signatures, we pair each forged signature with all genuine signatures and randomly select $C2$ 24 pairs out of 720 pairs of negative samples as negative samples, making a total of 160 × $C2$ 24 = 44160 pairs.

5.4 Experimental Results and Analysis

We partition the above four datasets of handwritten signatures in different languages, CNSig dataset, CEDAR dataset, BHSig-B dataset, and BHSig-H dataset according to the number of writers. Then 668 writers from the CNSig dataset were selected as the training set and 168 writers as the test set. A total of 50 writers from the CEDAR dataset were selected for the training set and 5 writers for the test set. 50 writers from the BHSig-B dataset were elected as the training set and 50 writers as the test set. Choose 100 writers from the BHSig-H dataset as the training set and 60 writers as the test set. The detailed segmentation results are shown in Table 4.

Next, performs an ablation study on att-OfSVNet, and the results are shown in Table 5. Table 5 presents the results of ablation studies on different methods, including data augmentation, image flipping, and 2D Attention module. The findings show that the use of data augmentation can significantly improve the performance of the model, especially on the BHSig-B and BHSig-H datasets, where the accuracy can be increased by about 2 percentage points. The use of image flipping can slightly improve the performance of the model, but the improvement is relatively small. The use of 2D Attention module can significantly improve the performance of the model, especially on the CNSig and BHSig-H datasets, where the accuracy can be increased by about 3 percentage points. Overall, the results suggest that the use of data augmentation and 2D Attention module can significantly improve the performance of the model, while the use of image flipping has a relatively weaker effect. And the results of ablation study on CNNs (include VGGNets and ResNets) are shown in Table 6. The findings show that the use of distortion contribute to the performance of the network, and distortion is particularly significant in improving the performance of the network, and distortion can drastically reduce the overfitting of the network. Table 7 displays the feature maps output when the input passes through each of the three 2D Attentions before and after inverting. It can be observed that the input image without inverting finds no features in any of the 2D Attentions, while the input image with inverting extracts many features in the second 2D Attention, and the attention points of signature are reflected in the third 2D Attention. Therefore, inverting the input will make the network retain more signature features.

To estimate the generalization ability of att-OfSVNet, we use three mainstream handwritten signature datasets, CEDAR dataset, BHSig-B dataset, and BHSig-H dataset, for training. Figure 7 shows

Table 4: Sample size of training set and test set in four different language handwritten signature datasets

Datasets	CNSig	CEDAR	BHSig-B	BHSig-H
Training set	60120	27600	27600	55200
Test set	15120	2760	27600	33120

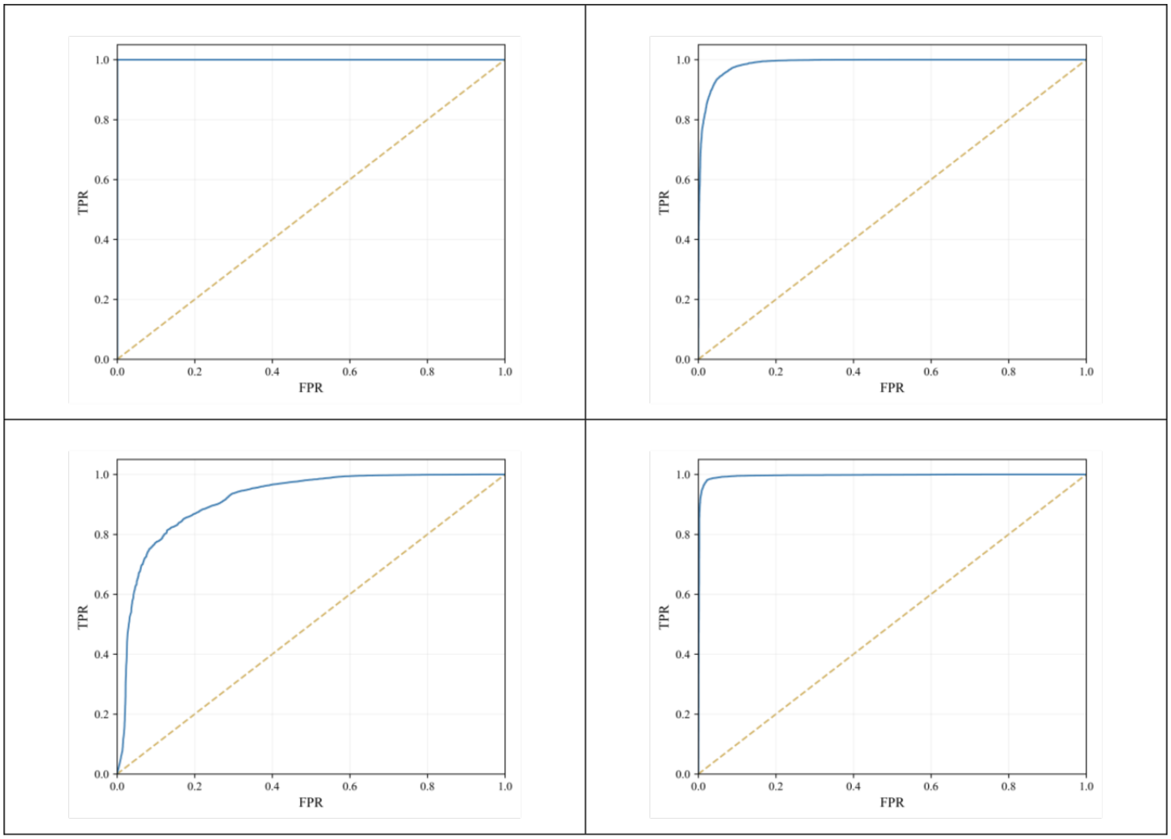


Figure 7: The ROC curves of att-OfSVNet on four datasets: (a) ROC curve on CEDAR, (b) ROC curve on BHSig-B, (c) ROC curve on BHSig-H, (d) ROC curve on CNSig.

Table 5: Ablation study of att-OfSVNet model

Distortion	Inversion	2D Attention	Acc (%)
			59.90
✓			**80.92**
	✓		62.79
		✓	62.71
✓	✓		**86.91**
✓	✓	✓	64.51
✓		✓	**94.42**
✓	✓	✓	**97.81**

the ROC curves of att-OfSVNet on four datasets. And Table 8 demonstrates the results of comparing att-OfSVNet with other offline signature verification networks, to assess its performance on unseen signatures, this approach can validate the network's generalization ability and real-world applicability. It is important to note that our test set is independent of the training set to ensure objectivity and

Table 6: Ablation study of CNNs on CNSig dataset

Methods	Pretrained	Freeze	Distortion	FRR (%)	FAR (%)	ERR (%)
VGG-11	✓			100.00	0.00	0.00
	✓		✓	100.00	0.00	0.00
	✓	✓		**25.70**	47.87	36.60
	✓	✓	✓	45.98	**28.82**	**36.32**
				100.00	0.00	0.00
			✓	100.00	0.00	0.00
VGG-13	✓			100.00	**0.00**	0.00
	✓		✓	**0.00**	100.00	0.00
	✓	✓		**12.59**	70.62	38.76
	✓	✓	✓	37.71	**39.38**	**38.57**
				100.00	0.00	0.00
			✓	100.00	0.00	0.00
VGG-16	✓			**0.00**	100.00	0.00
	✓		✓	100.00	**0.00**	0.00
	✓	✓		**26.39**	46.43	**36.76**
	✓	✓	✓	42.70	**32.39**	37.78
				100.00	0.00	0.00
			✓	100.00	0.00	0.00
VGG-19	✓			**0.00**	100.00	**0.00**
	✓		✓	43.73	**8.47**	8.47
	✓	✓		**47.58**	30.94	**38.31**
	✓	✓	✓	54.19	**29.54**	42.90
				100.00	0.00	0.00
			✓	100.00	0.00	0.00
ResNet-18	✓			55.33	20.17	36.69
	✓		✓	**43.73**	**8.47**	**20.09**
	✓	✓		40.48	**35.36**	37.22
	✓	✓	✓	**33.23**	40.28	**36.77**
				55.33	20.17	36.69
			✓	**43.73**	**8.47**	**20.09**
ResNet-34	✓			52.84	20.70	33.76
	✓		✓	**23.40**	**16.36**	**19.51**
	✓	✓		51.08	**30.66**	39.54
	✓	✓	✓	**44.55**	32.41	**38.15**
				52.84	20.70	33.76
			✓	**23.40**	**16.36**	**19.51**
ResNet-50	✓			45.52	27.39	33.36
	✓		✓	**18.15**	**10.56**	**13.37**
	✓	✓		44.18	38.43	41.03
	✓	✓	✓	**42.98**	**35.34**	**38.89**
				0.00	100.00	**0.00**
			✓	96.80	**6.83**	6.83

Table 7: Input invert and non-invert feature maps output by 2D Attention

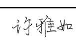
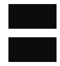
Network layers	Method	Input layer	2D Attention 1	2D Attention 2	2D Attention 3
Non-invert					
Invert					

Table 8: Accuracy comparison of currently popular offline signature verification networks on three different languages handwritten signature test datasets

Datasets	Methods	FRR (%)	FAR (%)	Acc (%)	ERR (%)
CEDAR	Chain Code [11]	9.36	7.84	-	-
	Graph Matching [7]	8.20	7.70	92.10	-
	Surroundedness [10]	8.33	8.33	91.67	-
	SigNet-F [19]	-	-	-	4.63
	Cut and Compare [2]	4.34	4.34	-	4.34
	IDN [1]	2.17	5.87	95.98	3.62
	SigNet-SPP-300dpi [21]	-	-	-	2.33
	Correlated Feature [13]	**0.00**	**0.00**	**100.00**	-
	SigNet [23]	**0.00**	**0.00**	**100.00**	-
	DeepHSV [33]	**0.00**	**0.00**	**100.00**	**0.00**
	Ours	**0.00**	**0.00**	**100.00**	**0.00**
BHSig-B	Texture Feature [14]	33.82	33.82	66.18	33.82
	Correlated Feature [13]	14.43	15.78	84.90	-
	SigNet [23]	13.89	13.89	86.11	-
	DeepHSV [33]	-	-	88.08	11.92
	Cut and Compare [2]	**3.96**	**3.96**	-	**3.96**
	IDN [1]	5.24	4.12	**95.32**	-
	Ours	7.93	4.12	93.97	5.61
BHSig-H	Texture Feature [14]	24.47	24.47	75.53	24.47
	SigNet [23]	15.36	15.36	84.64	-
	Correlated Feature [13]	15.09	13.10	85.90	-
	DeepHSV [33]	-	-	86.66	13.34
	Cut and Compare [2]	5.97	**5.97**	-	**5.97**
	IDN [1]	**4.93**	8.99	**93.04**	-
	Ours	18.36	13.27	84.18	16.04
CNSig	Ours	**1.77**	**2.60**	**97.81**	**2.13**

Table 9: Att-OfSVNet accuracy for cross-language testing (%)

Test set Training set	CNSig	CEDAR	BHSig-B	BHSig-H
CNSig	**97.81**	48.69	51.55	49.09
CEDAR	49.94	**100.00**	50.00	50.00
BHSig-B	52.06	47.98	**93.97**	66.84
BHSig-H	45.60	72.63	64.06	**84.18**

reliability in evaluating the results. Since other methods have not released their source codes, they cannot be compared on the CNSig dataset. It was found that the method of extracting signature features based on machine learning is outdated, and the model of signature feature extraction based on deep learning performs much better than other machine learning-based models due to its scalability. Although CNN achieve better results on handwritten signature datasets, they ignore the loss of feature information in the convolutional layer, which may prevent the network from learning effective handwriting features. Att-OfSVNet loses the information in the extraction process as little as possible in the convolutional layer and strengthens the efficiency of the Encoder in extracting features. Although it does not achieve the best performance on Western language datasets to date, att-OfSVNet is more suitable for verifying complex symbols such as Chinese handwritten signatures, with FRR, FAR and ERR of only 1.77, 2.60 and 2.13.

Table 9 presents the cross-lingual test results of att-OfSVNet, which shows that att-OfSVNet has a strong generalization ability. Indeed, signatures in different languages possess unique writing styles and characteristics. The differences between languages result in a decrease in accuracy when working with cross-lingual datasets. For instance, Chinese and English signatures exhibit distinct stroke structures, glyph features, and writing conventions. Therefore, when the network processes cross-lingual signatures, it may encounter greater challenges, leading to a decline in accuracy. Additionally, datasets from different languages vary in sample quality and diversity. Some language datasets may have higher levels of noise or variations, resulting in more complex signature features. This complexity can increase the learning difficulty for the network model and consequently lower its accuracy. Moreover, after statistics, the verification time of a single Chinese handwritten signature is only 0.02 seconds.

6 CONCLUSION

In this paper, to contribute to the research of Chinese signature verification and other related tasks, we develop a Chinese handwritten signature dataset CNSig and propose a 2D Attention-based offline signature verification network, att-OfSVNet, which enables the input to pass through the weight-sharing Encoder and Decoder to complete offline signature verification. The evaluation of att-OfSVNet with CNSig test set achieved 97.81% accuracy for offline signature verification. Moreover, we conduct experiments on multiple perspectives and demonstrate that att-OfSVNet has strong generalization ability and outperforms other offline signature verification networks. In addition, the verification time of a single Chinese handwritten signature is 0.02 seconds, which can fully meet the engineering requirements.

In the future, we will further develop a larger Chinese signature dataset and release the world's largest Chinese handwritten signature dataset, so that Chinese handwritten signature verification can be applied in various fields. On the other hand, we will continue to study offline signature verification networks and approach the Transformer domain to achieve more accurate and efficient offline signature verification cross-language. Furthermore, we hope to link offline signature verification to robots to automate the offline signature verification process, to identify forged signatures and protect customers' interests.

ACKNOWLEDGMENTS

This work was supported by 2021 Guangdong Province Enterprise Science and Technology Commissioner Special Funding Project (Grant no. GDKTP2021031000).

REFERENCES

[1] Wei, P., Li, H., Hu, P. (2019) Inverse Discriminative Networks for Handwritten Signature Verification. In: 2019 International Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach. pp. 5757-5765. 10.1109/CVPR.2019.00591
[2] Lu, X., Huang, L., Yin, F. (2020) Cut and Compare: End-to-end Offline Signature Verification Network. In: 2020 International Conference on Pattern Recognition (ICPR), Milan. pp. 3589-3596. 10.1109/ICPR48806.2021.9412377
[3] Chattopadhyay, S., Manna, S., Bhattacharya, S., Pal, U. (2022) SURDS: Self-Supervised Attention-guided Reconstruction and Dual Triplet Loss for Writer Independent Offline Signature Verification. In: 2022 International Conference on Pattern Recognition (ICPR), Montreal. pp. 1600-1606. 10.1109/ICPR56361.2022.9956442
[4] Kalera, M.K., Srihari, S., Xu, A. (2004) Offline Signature Verification and Identification Using Distance Statistics. International Journal of Pattern Recognition and Artificial Intelligence, 18 (7): 1339-1360. 10.1142/S0218001404003630
[5] Ferrer, M.A., Alonso, J.B., Travieso, C.M. (2005) Offline Geometric Parameters for Automatic Signature Verification Using Fixed-point Arithmetic. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (6): 993-997. 10.1109/TPAMI.2005.125
[6] Chen, S.Y., Srihari, S. (2005) Use of Exterior Contours and Shape Features in Off-line Signature Verification. In: 2005 International Conference on Document Analysis and Recognition (ICDAR), Seoul. pp. 1280-1284. 10.1109/ICDAR.2005.249
[7] Chen, S.Y., Srihari, S. (2006) A New Off-line Signature Verification Method Based on Graph Matching. In: 2006 International Conference on Pattern Recognition (ICPR), Hong Kong. pp. 869-872. 10.1109/ICPR.2006.125
[8] Ruiz-del-Solar, J., Devia, C., Loncomilla, P., Concha, F. (2008) Offline Signature Verification Using Local Interest Points and Descriptors. In: The Proceedings of Iberoamerican Congress on Pattern Recognition, Berlin. pp. 22-29. 10.1007/978-3-540-85920-8_3
[9] Kumar, R., Kundu, L., Chanda, B., Sharma, J.D. (2010) A Writer-independent Off-line Signature Verification System Based on Signature Morphology. In: The Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia (IITM), New York. pp. 261-265. 10.1145/1963564.1963610
[10] Kumar, R., Sharma, J.D., Chanda, B. (2012) Writer-independent Off-line Signature Verification Using Surroundedness Feature. Pattern Recognition Letters, 33 (3): 301-308. 10.1016/j.patrec.2011.10.009
[11] Bharathi, R.K., Shekar, B.H. (2013) Off-line Signature Verification Based on Chain Code Histogram and Support Vector Machine. In: 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Mysore. pp. 2063-2068. 10.1109/ICACCI.2013.6637499
[12] Guerbai, Y., Chibani, Y., Hadjadji, B. (2015) The Effective Use of the One-class SVM Classifier for Handwritten Signature Verification Based on Writer-independent Parameters. Pattern Recognition, 48 (1): 103-113. 10.1016/j.patcog.2014.07.016
[13] Dutta, A., Pal, U., Lladós, J. (2016) Compact Correlated Features for Writer Independent Signature Verification. In: 2016 International Conference on Pattern Recognition (ICPR), Cancun. pp. 3422-3427. 10.1109/ICPR.2016.7900163
[14] Pal, S., Alaei, A., Pal, U., Blumenstein, M. (2016) Performance of an Off-Line Signature Verification Method Based on Texture Features on a Large Indic-Script Signature Dataset. In: 2016 IAPR Workshop on Document Analysis Systems (DAS), Santorini. pp. 72-77. 10.1109/DAS.2016.48
[15] Serdouk, Y., Nemmour, H., Chibani, Y. (2017) Handwritten Signature Verification Using the Quad-tree Histogram of Templates and a Support Vector-based Artificial Immune Classification. Image and Vision Computing, 66: 26-35. 10.1016/j.imavis.2017.08.004
[16] Soleimani, A., Araabi, B.N., Fouladi, K. (2016) Deep Multitask Metric Learning for Offline Signature Verification. Pattern Recognition Letters, 80: 84-90. 10.1016/j.patrec.2016.05.023
[17] Zhang, Z., Liu, X., Cui, Y. (2016) Multi-phase Offline Signature Verification System Using Deep Convolutional Generative Adversarial Networks. In: 2016 International Symposium on Computational Intelligence and Design (ISCID), Hangzhou. pp.103-107. 10.1109/ISCID.2016.2033
[18] Hafemann, L.G., Sabourin, R., Oliveira, L.S. (2016) Writer-independent Feature Learning for Offline Signature Verification using Deep Convolutional Neural Networks. In: 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver. pp. 2576-2583. 10.1109/IJCNN.2016.7727521
[19] Hafemann, L.G., Sabourin, R., Oliveira, L.S. (2017) Learning Features for Offline Handwritten Signature Verification Using Deep Convolutional Neural Networks. Pattern Recognition, 70: 163-176. 10.1016/j.patcog.2017.05.012
[20] He, K., Zhang, X., Ren, S., Sun, J. (2015) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37 (9): 1904-1916. 10.1109/TPAMI.2015.2389824
[21] Hafemann, L.G., Oliveira, L.S., Sabourin, R. (2018) Fixed-sized Representation Learning from Offline Handwritten Signatures of Different Sizes. International Journal on Document Analysis and Recognition (IJDAR), 21: 219-232. 10.1007/s10032-018-0301-6
[22] Huang, G., Li, Y.X., Pleiss, G., Liu, Z., Hopcroft, J.E., Weinberger, K.Q. (2017) Snapshot Ensembles: Train 1, Get M for Free. Machine Learning. 10.48550/arXiv.1704.00109
[23] Dey, S., Dutta, A., Toledo, J.I., Ghosh, S.K., Lladós, J., Pal, U. (2017) SigNet: Convolutional Siamese Network for Writer Independent Offline Signature Verification. Pattern Recognition Letters. 10.48550/arXiv.1707.02131
[24] Xing, Z.J., Yin, F., Wu, Y.C., Liu, C.L. (2018) Offline Signature Verification Using Convolution Siamese Network. In: 2018 International Conference on Graphic and Image Processing (ICGIP), Qingdao. 10.1117/12.2303380
[25] Younesian, T., Masoudnia, S., Hosseini, R., Araabi, B.N. (2019) Active Transfer Learning for Persian Offline Signature Verification. In: 2019 International Conference on Pattern Recognition and Image Analysis (IPRIA), Tehran. pp. 234-239. 10.1109/PRIA.2019.8786013
[26] Mersa, O., Etaati, F., Masoudnia, S., Araabi, B.N. (2019) Learning Representations from Persian Handwriting for Offline Signature Verification, a Deep Transfer Learning Approach. In: 2019 International Conference on Pattern Recognition and Image Analysis (IPRIA), Tehran. pp. 268-273. 10.1109/PRIA.2019.8785979
[27] Deng, J., Dong, W., Socher, R., Li, L., Li, K., Li, F. (2009) ImageNet: A Large-scale Hierarchical Image Database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami. pp. 248-255. 10.1109/CVPR.2009.5206848
[28] Fierrez-Aguilar, J., Nanni, L., Lopez-Peñalba, J., Ortega-Garcia, J., Maltoni, D. (2005) An On-Line Signature Verification System Based on Fusion of Local and Global Information. In: The Proceedings of International Conference on Audio- and Video-Based Biometric Person Authentication, Berlin. pp. 523-532. 10.1007/11527923_54
[29] Ferrer, M.A., Diaz-Cabrera, M., Morales, A. (2015) Static Signature Synthesis: A Neuromotor Inspired Approach for Biometrics. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37 (3): 667-680. 10.1109/TPAMI.2014.2343981
[30] Soleimani, A., Fouladi, K., Araabi, B. N. (2016) UTSig: A Persian Offline Signature Dataset. In: 2016 International Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas. 10.1049/iet-bmt.2015.0058

[31] He, K., Zhang, X., Ren, S., Sun, J. (2016) Deep Residual Learning for Image Recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas. pp. 770-778. 10.1109/CVPR.2016.90

[32] Masoudnia, S., Mersa, O., Araabi, B.N., Vahabie, A., Sadeghi, M.A., Ahmadabadi, M.N. (2019) Multi-Representational Learning for Offline Signature Verification using Multi-Loss Snapshot Ensemble of CNNs. Expert Systems with Applications, 133: 317-330. 10.1016/j.eswa.2019.03.040

[33] Lin, F., Li, C., Wang, Z.Y., Yu, G., Yuan, L., Wang, H.Q. (2019) DeepHSV: User-Independent Offline Signature Verification Using Two-Channel CNN. In: 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney. pp. 166-171. 10.1109/ICDAR.2019.00035

[34] Zhu, Y., Lai, S., Li, Z., Jin, L. (2020) Point-to-Set Similarity Based Deep Metric Learning for Offline Signature Verification. In: 2020 International Conference on Frontiers in Handwriting Recognition (ICFHR), Dortmund. pp. 282-287. 10.1109/ICFHR2020.2020.00059

[35] Liu, L., Huang, L.L., Yin, F., Chen, Y.B. (2021) Offline Signature Verification Using a Region Based Deep Metric Learning Network. Pattern Recognition, 118. 10.1016/j.patcog.2021.108009

[36] Parcham, E., Ilbeygi, M., Amini, M. (2021) CBCapsNet: A Novel Writer-independent Offline Signature Verification Model Using a CNN-based Architecture and Capsule Neural Networks. Expert Systems with Applications, 185. 10.1016/j.eswa.2021.115649

[37] Vaswani, A., Shazeer, N., Parmar. N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I. (2017) Attention is All You Need. In: The Proceedings of the 31st International Conference on Neural Information Processing Systems. pp. 6000-6010. 10.48550/arXiv.1706.03762

[38] Li, H., Wei, P., Ma, Z.Y., Li, C.K., Zheng, N.N. (2022) Offline Signature Verification with Transformers. In: 2022 IEEE International Conference on Multimedia and Expo (ICME), Taipei. pp. 1-6. doi: 10.1109/ICME52920.2022.9859886

[39] Paszke, A., Chintala, S., Collobert, R., Kavukcuoglu, K., Farabet, C., Bengio, S., Melvin, I., Weston, J., Mariethoz, J. (2017) Pytorch: Tensors and Dynamic Neural Networks in Python with Strong GPU Acceleration.

[40] Kingma, D.P., Ba, J.L. (2015) Adam: a Method for Stochastic Optimization. In: 2015 International Conference on Learning Representations (ICLR), San Diego. 10.48550/arXiv.1412.6980

[41] Loshchilov, I., Hutter, F. (2017) SGDR: Stochastic Gradient Descent with Warm Restarts. In: 2017 International Conference on Learning Representations (ICLR). 10.48550/arXiv.1608.03983