

Article

A Two-Stage Siamese Network Model for Offline Handwritten Signature Verification

Wanghui Xiao ^{1,2} and Yuting Ding ^{1,*}

¹ School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; xiaowanghui007@126.com

² Education Information Technology Center, Southwest University of Political Science and Law, Chongqing 401120, China

* Correspondence: s200231024@stu.cqupt.edu.cn

Abstract: Offline handwritten signature verification is one of the most prevalent and prominent biometric methods in many application fields. Siamese neural network, which can extract and compare the writers' style features, proves to be efficient in verifying the offline signature. However, the traditional Siamese neural network fails to represent the writers' writing style fully and suffers from low performance when the distribution of positive and negative handwritten signature samples is unbalanced. To address this issue, this study proposes a two-stage Siamese neural network model for accurate offline handwritten signature verification with two main ideas: (a) adopting a two-stage Siamese neural network to verify original and enhanced handwritten signatures simultaneously, and (b) utilizing the Focal Loss to deal with the extreme imbalance between positive and negative offline signatures. Experimental results on four challenging handwritten signature datasets with different languages demonstrate that compared with state-of-the-art models, our proposed model achieves better performance. Furthermore, this study tries to extend the proposed model to the Chinese signature dataset in the real environment, which is a significant attempt in the field of Chinese signature identification.



Citation: Xiao, W.; Ding, Y. A Two-Stage Siamese Network Model for Offline Handwritten Signature Verification. *Symmetry* **2022**, *14*, 1216. <https://doi.org/10.3390/sym14061216>

Academic Editor: Chin-Ling Chen

Received: 18 April 2022

Accepted: 10 June 2022

Published: 12 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Fingerprints, irises, faces, voices, and handwritten signatures are five prevalent biometric recognition in many practical fields such as financial payment, attendance, computer vision, and contract signature [1–4]. Biometric recognition was started initially with the body measurements. Later with time and necessity, it involved many biometric properties related to the human body to provide authentication. Among these, the handwritten signature is the most commonly accepted symbol [5]. Verifying a person's identity using one's handwritten signatures is challenging, where a forger can access a person's handwritten signature and deliberately attempts to imitate it [6]. The main difficulties of offline handwritten signature verification are the high internal variability of individuals, scarcity of skilled forgery samples, and a limited number of training samples. Moreover, the discrepancy between genuine signature and skilled forgery is subtle since forgers attempt to imitate genuine signatures. With small inter-class variation between genuine signatures and skilled forgery while the larger intra-class variation of genuine signatures from the same person, it is very hard to differentiate between inter-personal forgery and intra-person handwritten signature genuineness.

Recently, powered by the rapid development of pattern recognition and image processing technology [7–9], it has become possible to verify handwritten signatures automatically. However, in the offline handwritten signature verification process, the dynamic information of the signature writing process is lost, and it is difficult to design a good feature extractor

that can distinguish genuine signatures from skilled forged signatures, which makes the problem more challenging.

Up to now, the Siamese neural network model [10,11] is one of the most popular and powerful approaches to address this issue and has greatly promoted the development of image identification. Siamese neural network transforms the features extracted from handwriting verification from traditional image texture features into convolution features [12,13], greatly improving the performance of handwriting verification. However, most of the existing methods regard offline handwritten signature [14–16] as an image recognition problem and cannot fully represent the writers' writing style. When the distribution of positive and negative handwritten signatures is very unbalanced [17,18], the performance of these models is unsatisfactory.

From this point of view, is it possible to improve the Siamese neural network model by considering the imbalance distribution of positive and negative signatures? To address this issue, this study proposes a two-stage Siamese neural network model. First, feature extraction is carried out by a two-stage convolutional neural network, which contains both the verification of the original two handwritten signatures and the verification of the handwritten signature after data enhancement. Second, this study introduces Focal loss as a loss function of the proposed network model, which can fully take into account the extremely unbalanced distribution of positive and negative signatures, making the neural network focus on stroke information rather than background information of the handwritten signatures. Extensive experimental results on four challenging handwritten signature datasets of different languages demonstrate the effectiveness of the proposed model.

The main contributions of this paper are as follows:

1. This is a significant attempt to study Chinese signature identification.
2. A two-stage Siamese network model is proposed to verify the offline handwritten signature.
3. Visualization of the process of feature representation is analyzed.

The remaining of this paper is organized as follows: Section 2 reviews the preliminaries on offline handwriting signature verification. Section 3 presents the proposed model. In Section 4, four experiments have been carried out for evaluating the proposed approach, and the results are explained and discussed. Section 5 concludes the findings and states the future works.

2. Preliminaries

2.1. Related Work

Offline handwritten signature verification can be considered as a two-class classification problem: to decide whether two offline handwritten signatures are signed by the same person or not, and to judge whether a handwritten signature is genuine for a specific user or not. Up to now, many methods have been proposed for offline handwritten signature verification [19–21]. Many studies often use texture features extraction such as gray-level co-occurrence matrix [22] and Local Binary Patterns [23]; directional-based features such as directional-pdf [24] and histogram of oriented gradients [25]; feature extractors specifically designed for offline handwritten signatures, such as the estimation of strokes by fitting Bezier curves [26]. Moreover, an inverse discriminative network [27] is proposed for writer-independent handwritten signature verification. Li, L et al. [28] proposed a region-based deep convolutional Siamese network for feature and metric learning. Wei et al. [5] proposed an inverse discriminative network that is capable of intensifying the effective information of signatures. Mustafa et al. [29] utilized a two-channel CNN as a feature extractor, where the two channels represent the reference and query signatures, respectively. A multi-task architecture based on R-SigNet architecture [30] is proposed, which exploits relaxed loss to learn a reduced feature space for writer-independent signature verification. All these methods have a good effect on signature identification to a certain extent. Despite the remarkable progress, signature verification is still very challenging due to the high intra-class

variety and low inter-class variety among signatures from different writers. Note that these methods have limits. They treat the offline handwritten signature as an image processing problem and fail to represent the writers' writing style fully. When the distribution of positive and negative handwritten signatures is very unbalanced, the performance of these models is unsatisfactory.

Despite great achievements in offline handwritten signature verification, existing models still have some limits as follows:

- Most of them only treat handwriting signature as a picture and do not mine deep signature style.
- They commonly ignore the imbalance distribution of positive and negative signatures that often occurs in real scenarios.
- The signature samples of each writer are usually small and the similarity between real signature and forged signature is high in real scenarios. The existing models usually generate synthetic data that are quite different from the real ones.

In comparison, our proposed model is significantly different from the existing models because:

- It has a two-stage Siamese network module to verify the offline-handwritten signature. This network includes both traditional original handwriting recognition and data-enhanced handwriting recognition to mine the writers' deep signature style.
- It employs the Focal loss to deal with the extreme imbalance between positive and negative offline signatures, which is quite different from previous studies.
- It is the first attempt to study the Chinese signatures with a real Chinese signature dataset.

2.2. CNN and Siamese Neural Network

Convolutional Neural Networks (CNN) are multilayer neural networks consisting of several convolutional layers with different kernel sizes interleaved by pooling layers, which summarize and downsamples the output of its convolutional layers before feeding to the next layers. The structure of the classical Convolutional Neural Network is shown in Figure 1. To obtain nonlinear correction, an activation function is also used. With the gradual increase of the number of convolutional layers, the range of the receptive field is gradually expanding. The closer it is to the subsequent output, the affected range of pixels of the image is wider. The convolutional neural network can learn different features in each volume base by advancing layer by layer and finally realize related recognition and classification functions.

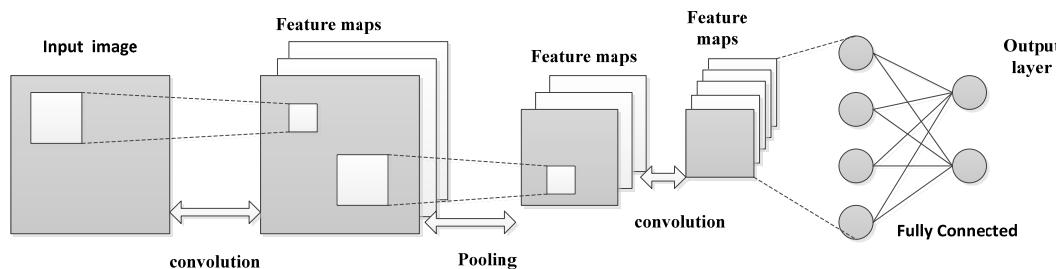


Figure 1. Architecture of Convolutional Neural Network.

Siamese network architecture was first introduced into the field of signature verification by Bromley et al. [31]. Since then, it has been widely used in many different fields such as one-time learning as well as text recognition and face recognition [32]. It consists of two identical subnetworks, maps inputs to higher-dimensional Spaces, and computes distance measures between high-level feature representations. The structure of the Siamese neural network is shown in Figure 2. Two CNNs share the same network and parameters but input different data. Two samples are taken as inputs, and then their representations are embedded into the output in the high-dimensional space to compare the similarity of

the two inputs. Through the forward processing of a convolutional neural network, the data that is difficult to distinguish in the original space can be represented in a specified dimension, making it easy to distinguish. The Siamese neural network is widely used in face verification, signature verification, and other tasks, in which samples are not directly classified, but compared with known patterns to determine whether they belong to the same category. The Siamese neural network model generally adopts the cross-entropy loss function [33] with regularization term. As a typical binary classification problem, it is given by:

$$L(x_1, x_2) = y(x_1, x_2) \log P(x_1, x_2) + (1 - y(x_1, x_2)) \log(1 - p(x_1, x_2)) + \lambda^T |w|^2 \quad (1)$$

where $y(x_1, x_2) = 1$ represents when x_1 and x_2 belong to the same kind of object or $y(x_1, x_2) = 0$ represents when x_1 and x_2 belong to the different kind of object.

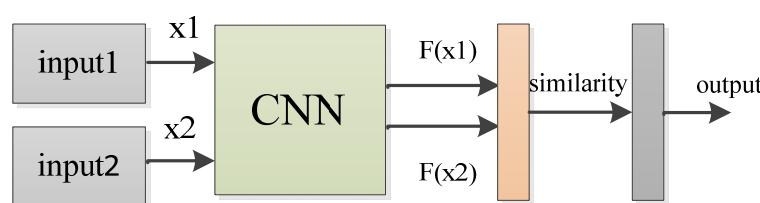


Figure 2. Architecture of Siamese Neural Network.

2.3. Focal Loss

In handwriting signature identification cases, the handwriting signature that can be obtained is usually very limited. The positive signature is usually from a credit card consumption signature, file data signature, or contract signature, while the same document can only provide one positive signature at the same time. A forged signature often comes from someone else's imitation, which can be multiple imitations. In the existing open dataset, the handwriting signature of the positive sample pair is often much fewer than that of the negative sample pair, and the distribution of positive and negative signatures used for comparison in the data set is imbalanced.

Dealing with unbalanced data has always been a challenge in deep learning and machine learning. Based on the classical cross-entropy loss function, Focal Loss [34] was first proposed to handle the object detection scenario where the unbalance exists between foreground and background classes. To handle the imbalanced data of different classes, a weighting factor and modulating factor are used to adjust the loss function. The sharing weight to the total loss can be adjusted by changing the value of the weighting factor α and modulation factor γ . Compared with the classical cross-entropy loss, the Focal Loss focuses more on the difficult and misclassified cases and plays a well-regulating role in the extremely unbalanced distribution of positive and negative samples in the dataset. Formally, the Focal Loss [34] adds the factor $-\alpha(1 - \hat{y})^\gamma$ and $-(1 - \alpha)\hat{y}^\gamma$ to the standard cross entropy criterion. Setting $\gamma > 0$ reduces the relative loss for well-classified examples, putting more focus on hard, misclassified examples, it is given by:

$$FL = \begin{cases} -\alpha(1 - \hat{y})^\gamma \cdot \log(\hat{y}), & y = 1 \\ -(1 - \alpha)\hat{y}^\gamma \cdot \log(1 - \hat{y}), & y = 0 \end{cases} \quad (2)$$

where α is a weighting factor of $[0, 1]$, \hat{y} denotes the predicted value, and γ is an adjustable focus parameter to prevent easy samples from contributing too much.

3. Model

3.1. Problem Formulation

Assume x is a verification sample signature and s is a genuine signature sample. This study aims to distinguish whether x is forged or genuine compared to s . The output

predicted value $y \in [0, 1]$ is the label for validating the decision where $y = 1$ means the verification sample signature is genuine to s signature image and $y = 0$ indicates the verification sample signature is forged to s signature image. Thus, this signature verification problem can be represented as

$$y = \Phi(x, s, \theta) \quad (3)$$

where the decision function $\Phi(\cdot)$ maps the input handwriting signature images x and s to the predicted value y . After training and learning, all parameters are saved in parameter set θ . Note that the problem defined in Equation (3) is similar to but different from the regression problem in the neural network machine learning tradition. First, traditional regression problems in neural network machine learning generally have one input, while handwritten signature verification problems have two original inputs. Second, in most traditional binary classification problems, especially in the field of image processing, the main object of feature extraction is based on color, texture, intensity, etc. In the problem of handwriting signature verification, the key part is to distinguish the difference in handwriting style. Writing style, as an abstract feature, is an indescribable attribute defined by the strokes of a handwritten signature rather than color or background.

3.2. Architecture of the Two-Stage Network

The architecture of the proposed two-stage connected neural network is shown in Figure 3. In the proposed model architecture, the upper and lower layers of the model are completely symmetric, and the left and right sides of the model are relatively symmetric. The model contains a two-stage Siamese neural network and consists of three modules: the convolutional neural feature extractor module, image enhancement module, and the objective function module. The structure and function of each module will be described in detail in later chapters. By adopting a two-stage siamese neural network, the original input handwritten signature verification and the image enhanced signature verification are realized simultaneously, and the output results of the two-stage networks are combined to verify with the label Simultaneously. The two-stage network is beneficial to the extraction and verification of shallow and deep features of handwriting and improves the accuracy of handwriting signature verification.

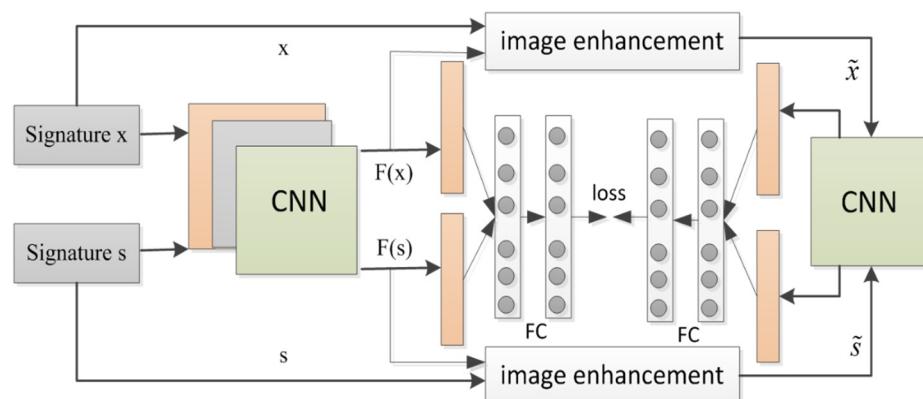


Figure 3. Architecture of the proposed two-stage Siamese Neural Network.

The Focal Loss function was adopted to adjust the extreme imbalance between positive signature and negative signature samples. In this model, \tilde{x} and \tilde{s} are the new images of the original input signature image after a series of image transformations and image enhancement, respectively. This idea is shown in Figure 3. The original input of the check sample signature and the genuine sample signature will obtain the corresponding validation decision tag y . A new inspection sample pair is generated by enhancing the image data of the original inspection sample image and the sample image. These newly generated inspection sample image pairs should have the same validation labels, for the image enhancement processes do not change the signature structure and writing style.

3.3. The Feature Extractor

The architecture of the feature extractor module is illustrated in Figure 4. The proposed networks take two original handwritten signatures as inputs and take the feature of the signatures as outputs. The two input images share the same network parameters. Firstly, the image signature is preprocessed. Since the neural network requires all input sizes to be consistent, all signature sizes are unified as 115×220 dimensions of the gray image in this study. After the signature is input into the network structure, some convolutional neural network layers are used to extract features from the signature image. The feature extraction process of the convolutional neural network includes a convolution layer, nonlinear activation layer, max-pooling layer, batch normalization layer, etc. The convolution layer and full connection layer have learnable parameters, and the parameters are continuously optimized in the training process. After each learnable layer, we apply batch normalization, followed by ReLU nonlinear activation. The last layer adopts Softmax nonlinear activation and interprets the output as a probability.

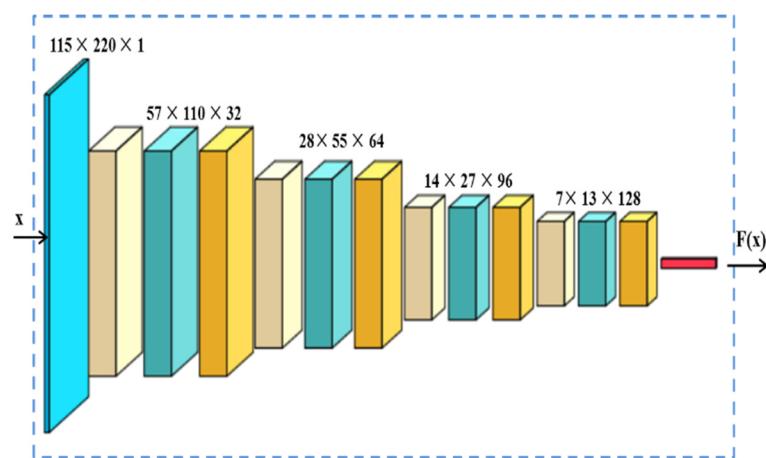


Figure 4. The inner structure of a feature extractor module.

There are four cascaded convolutional operations inspired by the Visual Geometry Group net, and each operator group consists of two convolutional layers followed by Rectified Linear Units function, a normalized layer, and a pooling layer. Generally, through four cascades, the global and local features in the handwritten signature images can be fully represented. The channel number of four cascaded convolutional layers in the corresponding operator group is 32, 64, 96, and 128, respectively. The main function of the batch normalization layer is to standardize the input of each layer and prevent the gradient explosion and gradient disappearance in the subsequent calculation. The batch normalization layer only works for training, not for validation sets. The function of the activation layer is to add nonlinear factors and map features to high-dimensional nonlinear intervals for interpretation. As shown in Figure 4, all the input image processing processes use the same training parameters, that is, all CNN modules share the same network and parameters. By learning these training parameters to extract effective signature author style features, the computational complexity is reduced, and the performance of the model is improved effectively.

3.4. The Signature Image Data Enhancement

The architecture of the image enhancement module is illustrated in Figure 5. In this module, it receives the original input signature image x and the feature map $F(x)$, which comes from the input signature through the output of the feature extraction module as the inputs of this module. The feature $F(x)$ already contains the writing features of part of the signature handwriting after a series of convolution processes. To well represent the writing style features of signature handwriting pictures, this module further refines the features.

After a GAP layer (global average of pooling layer) processing and two FC layers (fully-connected layers), the output feature is reshaped to 57×110 dimensions. The extracted features are restored to the matrix form, and then the nearest neighbor up-sampling and filling layers are carried out. The up-sampling layer can retain the image features extracted from the earlier convolutional layer to the maximum extent, which plays a good role in the subsequent feature extraction.

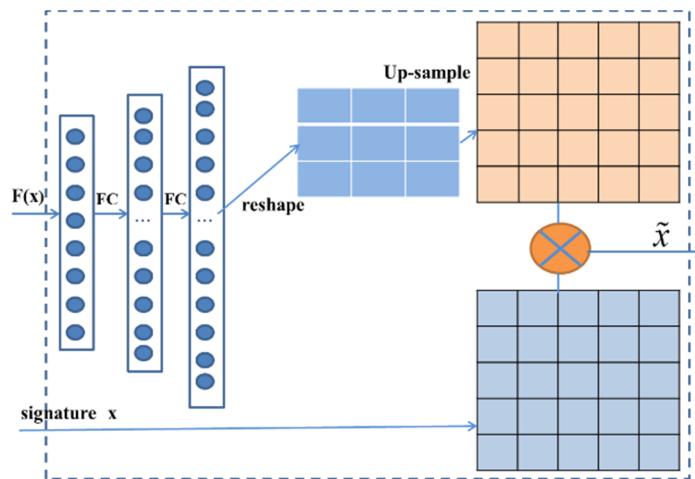


Figure 5. The image enhancement module.

Through the padding layer and the convolutional layer with a nonlinear activation function, the size of the output matrix is completely consistent with the original input signature image, which is enhanced for the subsequent data. The final output result of the data enhancement module is obtained by multiplying the data enhancement weight matrix with the input matrix of the original input signature image. In this way, the feature data on the image can be enlarged and the handwriting features can be further extracted. The symbol \otimes represents element multiplication. Through the processing of the module, the generated image \tilde{x} becomes a new image after data enhancement, which has the same size as the original signature x .

3.5. Loss Function

The loss function can evaluate the performance of the model, which is the most critical part of machine learning, especially in the training phase. Suppose elements combination $\{(x_i, s_i, y_i) | i = 1, \dots, N\}$ is a sample of the training dataset, and x_i and s_i are the i -th verification sample and the genuine verification, respectively. The value $y_i \in \{0, 1\}$ is (x_i, s_i) , where $y = 0$ indicates the verification sample x_i is forged compared to s_i , and where $y = 1$ indicates genuine. With the training dataset, this study aims to optimize the network parameters consisting of two parts. Both the original sample and the image enhanced sample are involved in the training process, the loss function of the model consists of two losses: the loss of the original sample to contrast and the loss of new sample comparison after data enhancement. For the original handwritten signature (x_i, s_i, y_i) , $F(x_i)$ and $F(s_i)$ are the output of the features from the original samples. $P(F(x_i), F(s_i))$ is the signature verification probability predicted. To deal with the extreme imbalance between positive and negative offline signatures, this study adopts the Focal loss function [35] to express the loss of verification results. As a binary classification problem, the loss $FL_O(x_i, s_i)$ result can be calculated:

$$FL_O(x_i, s_i) = \begin{cases} -\alpha(1 - \hat{y})^\gamma \log(\hat{y}), & y = 1 \\ -(1 - \alpha)\hat{y}^\gamma \log(1 - \hat{y}), & y = 0 \\ \hat{y} = P((F(x_i), F(s_i))) \end{cases} \quad (4)$$

Based on the input handwriting signatures x_i and s_i , the image enhanced signatures are defined as \tilde{x}_i and \tilde{s}_i . In handwriting identification studies, $(\tilde{x}_i, \tilde{s}_i)$ should have the same predictive value as the original signature image (x_i, s_i) , for the data enhancement of the signature, the image does not change the writing style of the signature. Therefore, the loss function $FL_E(\tilde{x}_i, \tilde{s}_i)$ for $(\tilde{x}_i, \tilde{s}_i)$ after image enhancement should be defined as follows

$$FL_E(\tilde{x}_i, \tilde{s}_i) = \begin{cases} -\alpha(1 - \hat{y})^\gamma \log(\hat{y}), & y = 1 \\ -(1 - \alpha)\hat{y}^\gamma \log(1 - \hat{y}), & y = 0 \end{cases} \quad (5)$$

$$\hat{y} = P((F(\tilde{x}_i), F(\tilde{s}_i)))$$

Hence, the final loss is the total loss of training samples by combining the losses of the two parts, which can be calculated as follows

$$Loss = \sum_i^N \{FL_O(x_i, s_i) + \lambda \cdot FL_E(\tilde{x}_i, \tilde{s}_i)\} \quad (6)$$

where λ is a hyperparameter whose function is to balance the weight of two parts of the loss and is an empirical value.

3.6. Algorithm Design

Here, we discuss about the training process of the proposed algorithm, which is demonstrated in Algorithm 1.

Algorithm 1: Training Process of the Proposed Algorithm

Require: set up the batch size m , the maximum number of epoch k , the learning rate LR , and the penalty factors λ

Require: Initialize the weights of the networks θ .

for epoch number = 1 : k **do**

 Randomly select m images from the training image dataset: x_i

 Select m corresponding genuine images from the preprocessed dataset: s_i

 Calculate the eigenvector and the Loss according to the network weights θ .

 Update the weights of the networks θ .

$\nabla_\theta \frac{1}{k} \sum_{i=1}^k [FL_O(x_i, s_i) + \lambda \cdot FL_E(\tilde{x}_i, \tilde{s}_i)]$

End for

4. Empirical Studies

4.1. General Settings

This study tested the proposed model on four challenging offline handwritten signature datasets: the Chinese signature dataset, CEDAR signature datasets, BHSig-Hindi signature datasets, and BHSig-Bengali signature datasets, which come from four different languages, respectively: Chinese, English, Hindi, and Bengali. Taking the CEDAR dataset [18] as an example, this dataset contains a total of 55 English signatures, and 24 forged handwritten signatures and 24 genuine signatures were written for each person, so there are $C^2_{24} = 276$ pairs of positive signatures. Combining the genuine and forged signatures of each person, there are $24 \times 24 = 576$ pairs of negative signatures. According to the characteristics of machine learning, 50 signature samples were randomly selected in this study to train model parameters, and the remaining 5 signature samples were used as validation samples. BHSig260 signature Dataset [36] was divided into the BHSIG-Bengali signature set and BHSIG-Hindi signature set, which were trained and verified independently, respectively. The BHSIG-Bengali signature dataset contains a total of 100 handwritten signature images with Bengali signatures. Each person has 30 forged handwritten signatures and 24 genuine signatures, so there are $C^2_{24} = 276$ pairs of positive signatures. Combining the genuine and forged signatures of each person, there are $30 \times 24 = 720$ pairs

of negative signatures. Bhsig—Hindi Signature Dataset is another subset of BHSig260, which contains a total of 160 handwritten signature images of people signing in Hindi. Similar to Bengali signatures, each person has 30 forged handwritten signatures and 24 real signatures, so there are $C^2_{24} = 276$ pairs of genuine signatures. Combining the genuine and forged signatures of each person, there are $30 \times 24 = 720$ pairs of negative signatures. The Chinese signature dataset contains 500 groups of signature data signed in Chinese, and each group of signature data contains one check sample signature and three genuine signatures. The check sample signatures need to be used to verify whether they are genuine handwriting or not, and the other three sample signatures are genuine handwriting that has been confirmed to be written by oneself. The signatures of each dataset come from different scenarios and different sampling times.

Chinese Handwritten Signature Dataset: Since the previous Chinese handwriting signatures were imitated in the laboratory and the data amount was small, there was no suitable Chinese handwriting signature dataset. Therefore, we collected a multi-source Chinese handwriting signature dataset with a large period and strong practical significance. This data set includes both positive and negative signatures sample, which are from the National Forensic Center of Southwest University of Political Science and Law between 2009 and 2020. As it is a real case, these signatures all come from real-life signatures such as credit card consumption signatures, personal file signatures, signatures in document contracts. In a real setting, to ensure the consistency of the data set sample data, the datasets collected an examination of each signature handwriting and three certified true signature handwriting, and formed a set of data signature handwriting. There are altogether 500 sets of such signature handwriting data, including 220 sets of negative signature handwriting data and 280 sets of positive signature handwriting data. The handwriting data of each signature applying for handwriting identification are highly similar. All the signed signatures were scanned into images at 300DPI. The Chinese offline signature dataset consists of 500 names and 2000 signature images. This dataset has the characteristics of multi-source, real, and large scale. First, all the signatures are from real cases, which is often challenging. Secondly, the Chinese data set belongs to a relatively large-scale data set, in which the handwriting period of decades. Third, real signatures are collected at different times and in different scenarios, and the signatures of the same person may be significantly different. All of these characteristics make this data set very valuable and challenging.

To further understand the sample information of signature images in each dataset, some sample signatures are shown in Table 1. In each dataset, the data set details are shown in Table 2.

Evaluation Metrics: In this study, a group of positive samples is composed of two genuine signatures written by the same person, and the corresponding recognition decision label $y = 1$. The evaluation metrics are based on the prediction of the sample pairs in all validation sets and the statistical analysis of the predicted results. Three evaluation indicators were used to evaluate and compare the proposed method with other methods: false acceptance rate (FAR), false rejection rate (FRR), and accuracy (ACC). The false acceptance rate is defined as the ratio of the number of false acceptances divided by the number of negative signature samples. The false rejection rate is defined as the ratio of the number of false rejections divided by the number of positive signature samples. Lower FRR or FAR and higher ACC mean better performance. They are calculated as follows:

$$\begin{aligned} \text{FAR} &= \frac{\text{False positive}}{\text{False positive} + \text{True negative}} \times 100\% \\ \text{FRR} &= \frac{\text{False negative}}{\text{False negative} + \text{True positive}} \times 100\% \\ \text{ACC} &= \frac{\text{True positive} + \text{True negative}}{\text{True positive} + \text{True negative} + \text{False positive} + \text{False negative}} \times 100\% \end{aligned} \quad (7)$$

Baselines: This study compares our proposed model with eight involved state-of-the-art models, including five writer-independent methods models (SigNet [37], Surroundness [38], Chain code [39], Ensemble Learning [40], Morphology [41], and DeepHSV [30])

and three writer-dependent models (Chain code [39], Texture Feature [42] and Fusion of HTF [6]). Table 3 describes the main descriptions of the correlation models.

Table 1. Some genuine samples and skilled forgeries.

	Genuine	Forgery
CEDAR		
BHSig-Bengali		
BHSig-Hindi		
CHINESE		

Table 2. Dataset Details.

Datasets	CEDAR	BHSig-B	BHSig-H	CHINESE
languages	English	Bengali	Hindi	Chinese
People	55	100	160	500
Signatures	2640	5400	8640	2000
Total sample	46,860	99,600	159,360	1500
Positive: negative	276:576	276:720	276:720	840:660

Table 3. Descriptions of All Involved Models.

Model	Description
SigNet	The writer independent Siamese network model proposed in 2017 [37] and is often applied to signature verification.
Surroundness	A signature feature extraction model based on envelopment was proposed in 2012 [38].
Chain code	In 2013 [39], a model based on the histogram features of chain codes was proposed and enhanced by Laplacian Gaussian filter.
Ensemble Learning	Deep learning model proposed in 2019 [40], which improves an integration model for offline writer independent signature verification.
Morphology	Feature analysis technology based on multi-layer perceptron was proposed in 2010 [41].
Texture Feature	a texture-oriented signature verification method was proposed in 2016 [42]. It has good performance for Indian scripts.
Fusion of HTF	A Signature verification model proposed in 2019 [6]. It adopts discrete wavelet and local quantized patterns features
DeepHSV	A neural network model proposed in 2019 [30], which improves the network with a two-channel CNN network

4.2. Comparison with State-of-the-Art Models

This study was independent of the writer, and we labeled writer-independent (WI) in the table of experimental results, such as [37]. It also lists the results of methods that rely on the writer himself and are labeled writer-dependent (WD), such as [39]. The writer-independent approaches train just one model for all test writers, and writer-dependent approaches train a single-minded model for each writer. The writer-dependent model generally has better performance than the writer-dependent model but requires training in everyone's signature sample, which is impractical and cannot be generalized to unobserved people. This study is based on writer-independent and only one model parameter set is trained in the same data set.

The performance of the proposed model was compared with that of the four state-of-the-models, and detailed comparisons are given in Tables 4 and 5 below. In these experiments, To make the networks learn as many characteristic attributes of signatures as possible, we attempt to remove noise and keep foreground information about the signature itself. Firstly, the OTSU algorithm [34] was used to separate foreground and background regions, and the batch normalization was utilized to normalize signature images. Second, the background pixel value is converted to 255 and the original pixel value of signature strokes is retained. The running environment of the model is based on Pytorch 1.3.1 framework, using the NVIDIA 2080Ti GPU graphics card. In this study, the stochastic gradient descent optimization method was adopted, the basic learning rate was 1×10^{-5} , and the batch size was set as 32. The λ in Equation (5) is set as the empirical value 2.5. The proposed model is compared with the baseline model, the traditional Siamese neural network method, and the classical cross-entropy loss function method. As shown in Tables 4 and 5, our model achieves better performance than state-of-the-art models.

Table 4. COMPARISON ON CEDAR DATASET (%).

Method	Type	FRR	FAR	ACC
Morphology	WI	12.39	11.23	88.19
Surroundness	WI	8.33	8.33	91.67
Chain code	WD	9.36	7.84	92.16
Ensemble Learning	WI	8.48	7.88	92.00
ISNN + CrossEntropy	WI	9.38	7.68	92.55
SNN + Focal Loss	WI	8.92	6.94	93.47
Our method	WI	6.78	4.20	95.66

Table 5. COMPARISON OF BHSIG-BENGALI AND BHSIG-HINDI DATASET (%).

Method	Type	BHSig-Bengali			BHSig-Hindi		
		FRR	FAR	ACC	FRR	FAR	ACC
SigNet	WI	13.89	13.89	86.11	15.36	15.36	84.64
Texture Feature	WD	33.82	33.82	66.18	24.47	24.47	75.53
Fusion of HTF	WD	18.42	23.10	79.24	11.46	10.36	79.89
DeepHSV	WI	11.92	11.92	88.08	13.34	13.34	86.66
ISNN + CrossEntropy	WI	18.64	12.86	86.66	15.63	15.49	84.54
SNN + Focal Loss	WI	16.87	9.43	87.69	13.38	10.91	84.79
Our method	WI	14.25	6.41	90.64	12.29	9.6	88.98

4.3. Chinese Signature Dataset

This study is the first to identify handwriting in a real case. The research results of this study will be helpful to judicial identification and have important research value. In the Chinese offline signature handwriting data set, the handwriting information collected is from real handwriting identification cases from 2008 to 2020. In addition, due to a large number of Chinese fonts, similar appearance, ease of confusion, and the randomness of Chinese signatures, the results of handwriting identification will be affected to a certain extent. In addition, different from the multi-arc features of Latin letters, the focus of Chinese handwriting identification is also different. The main characteristics of Chinese signatures are the special and stable parts such as stroke crossing, connection, and collocation. We fully believe that it will contribute to the field of Chinese offline handwritten signature verification and related research. As shown in Table 6, compared with other current methods, this study plays a certain role in handwriting prediction, which is a significant attempt in the field of Chinese signature identification.

Table 6. COMPARISON ON Chinese DATASET (%).

Method	Type	FRR	FAR	Acc
SigNet	WI	42.36	42.36	57.64
DeepHSV	WI	41.87	41.87	58.13
SNN + CrossEntropy	WI	38.98	35.77	64.79
ISNN + CrossEntropy	WI	33.66	31.24	68.88
SNN + Focal Loss	WI	36.74	30.92	65.88
ISNN + Focal Loss	WI	32.18	30.59	70.31

4.4. Process Visualization

Figure 6 shows the feature extraction process of signature image features. It can be seen from the figure that at the beginning of training, neural network learning features

mainly focus on texture features of handwritten pictures, as shown in Figure 6b–d. With the deepening of training, the features learned are gradually abstract, which can be understood as handwriting style features.

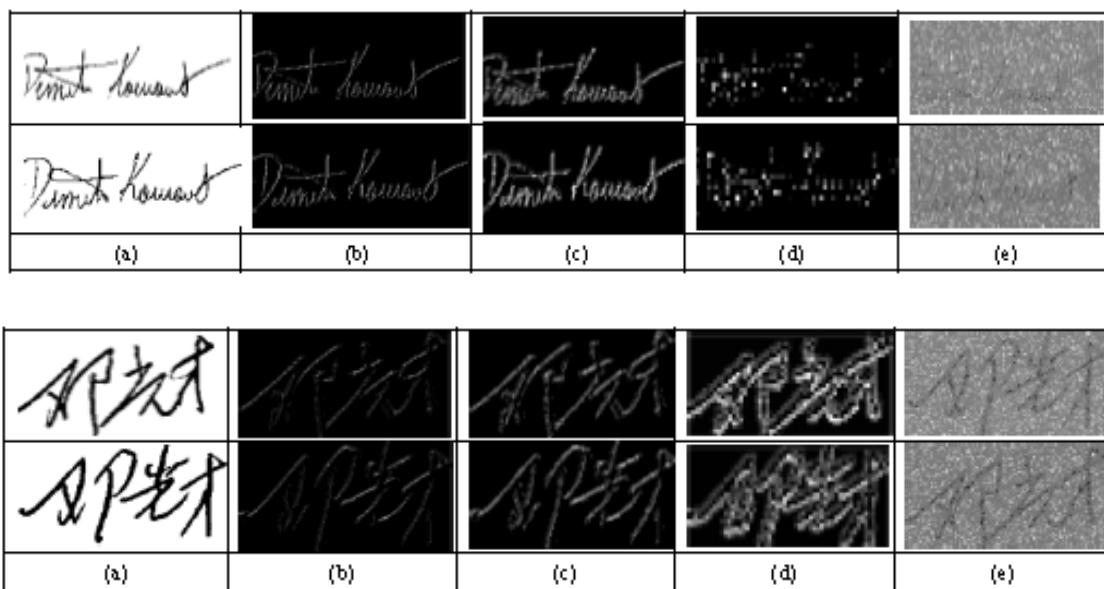


Figure 6. (a) Original Signature images, (b–d), feature extraction in CNN process, (e) Signatures after data enhancement.

From the above conclusions, we can draw the following important implications:

1. Compared with previous methods, this model has better prediction performance. On the CEDAR signature dataset, the FRR, FAR, and ACC of the proposed method reach 6.78%, 4.20%, and 95.66%, respectively, which are superior to the existing comparison methods under all evaluation indicators. On the BHSIG-Bengali and BHSIG-Hindi signature datasets, our model achieves ACC of 90.64% and 88.98%, respectively, which is superior to other models. These results show that our method is superior to other comparison methods. In addition, our writer-independent approach still performs better than the writer-dependent approach.
2. The data enhancement method adopted in this study is only related to the original input signature image. The original input signature image is processed by a series of neural networks to generate a data enhancement weight matrix. Finally, the degree of image data enhancement is adjusted by adjusting the proportion of the weight matrix, which improves the accuracy of experimental results, and the proposed model has strong robustness.
3. The focal Loss function is very effective for solving the problem of unbalanced positive and negative data.
4. The proposed model also has good performance in Chinese signature datasets, and this conclusion will be helpful for further research on offline Chinese signature verification.

5. Conclusions

Aiming at solving the problem of offline handwritten signature verification, this study proposes a two-stage Siamese neural network model to extract the writers' writing style. Based on the end-to-end image enhancement learning method and Focal Loss function, the proposed model can effectively solve the problem of imbalance of positive and negative samples, achieve good performance on challenging datasets with three different languages, and also work well in Chinese offline handwritten signature dataset. To evaluate the proposed model, we conduct extensive experiments on four challenging handwritten signature datasets with different languages. The results demonstrate that the proposed

model achieves better performance than the state-of-the-art models. Future work will focus on the study of Chinese handwriting signatures and improving the accuracy of Chinese handwriting identification.

Author Contributions: Conceptualization, W.X.; methodology, W.X.; software, W.X.; validation, W.X.; formal analysis, Y.D.; investigation, W.X.; resources, Y.D.; data curation, W.X.; writing—original draft preparation, W.X.; writing—review and editing, Y.D.; visualization, W.X.; supervision, Y.D.; project administration, Y.D.; funding acquisition, W.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Science and Technology Project of Chongqing Education Commission, No. KJQN202100304 and the Key Cooperation Project of Chongqing Municipal Education Commission, No. HZ2021008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: For the collection of Chinese handwriting datasets, we thank the National Forensic Center of Southwest University of Political Science and Law.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Vincent, C.; David, B.; Florian, H.; Andreas, M.; Elli, A. Writer Identification Using GMM Supervectors and ExemplarSVMs. *Pattern Recognit.* **2017**, *63*, 258–267.
2. Luo, X.; Liu, Z.; Li, S.; Shang, M.; Wang, Z. A Fast Non-Negative Latent Factor Model Based on Generalized Momentum Method. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *51*, 610–620. [\[CrossRef\]](#)
3. Khan, F.; Tahir, M.; Khelifi, F. Robust Offline Text in Dependent Writer Identification Using Bagged Discrete Cosine Transform Features. *Expert Syst. Appl.* **2017**, *71*, 404–415. [\[CrossRef\]](#)
4. Wu, H.; Luo, X.; Luo, X. Advancing non-negative latent factorization of tensors with diversified regularizations. *IEEE Trans. Serv. Comput.* **2020**, *99*, 1. [\[CrossRef\]](#)
5. Wei, P.; Li, H.; Hu, P. Inverse Discriminative Networks for Handwritten Signature Verification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 5757–5765.
6. Maergner, P.; Pondenkandath, V.; Alberti, M.; Liwicki, M.; Riesen, K.; Ingold, R.; Fischer, A. Combining Graph Edit Distance and Triplet Networks for Offline Signature Verification. *Pattern Recognit. Lett.* **2019**, *125*, 527–533. [\[CrossRef\]](#)
7. Jain, A.; Singh, S.K.; Singh, K.P. Handwritten signature verification using shallow convolutional neural network. *Multimed. Tools Appl.* **2020**, *79*, 19993–20018. [\[CrossRef\]](#)
8. Wu, D.; Luo, X.; Shang, M.; He, Y.; Wang, Y.; Wu, X. A Data-Characteristic-Aware Latent Factor Model for Web Service QoS Prediction. *IEEE Trans. Knowl. Data Eng.* **2022**, *34*, 2525–2538. [\[CrossRef\]](#)
9. Wu, D.; He, Q.; Luo, X.; Shang, M.; He, Y.; Wang, Y. A posterior-neighborhood-regularized latent factor model for highly accurate web service QoS prediction. *IEEE Trans. Serv. Comput.* **2022**, *15*, 793–805. [\[CrossRef\]](#)
10. Zois, E.N.; Alewijnse, L.; Economou, G. Offline signature verification and quality characterization using poset-oriented grid features. *Pattern Recognit.* **2016**, *54*, 162–177. [\[CrossRef\]](#)
11. Luo, X.; Qin, W.; Dong, A.; Sedraoui, K.; Luo, X. Efficient and High-quality Recommendations via Momentum-incorporated Parallel Stochastic Gradient Descent-based Learning. *IEEE/CAA J. Autom. Sin.* **2021**, *8*, 402–411. [\[CrossRef\]](#)
12. Li, H.; Wei, P.; Hu, P. AVN: An Adversarial Variation Network Model for Handwritten Signature Verification. *IEEE Trans. Multimed.* **2021**, *24*, 594–608. [\[CrossRef\]](#)
13. Wu, D.; He, Q.; Luo, X.; Zhou, M. A Latent Factor Analysis-Based Approach to Online Sparse Streaming Feature Selection. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *1*, 1–15. [\[CrossRef\]](#)
14. Alaei, A.; Pal, S.; Pal, U.; Blumenstein, M. An efficient signature verification method based on an interval symbolic representation and a fuzzy similarity measure. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 2360–2372. [\[CrossRef\]](#)
15. Liu, Z.; Luo, X.; Wang, Z. Convergence Analysis of Single Latent Factor-dependent, Non-negative and Multiplicative Update-based Non-negative Latent Factor Models. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 1737–1749. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Jain, A.; Singh, S.K.; Singh, K.P. Multi-task learning using GNet features and SVM classifier for signature identification. *IET Biom.* **2021**, *2*, 117–126. [\[CrossRef\]](#)
17. Luo, X.; Yuan, Y.; Chen, S.; Zeng, N.; Wang, Z. Position-transitional particle swarm optimization-incorporated latent factor analysis. *IEEE Trans. Knowl. Data Eng.* **2020**, *99*, 1. [\[CrossRef\]](#)

18. Sharif, M.; Khan, M.; Faisal, M.; Yasmin, M.; Fernandes, S.L. A framework for offline signature verification system: Best features selection approach. *Pattern Recognit.* **2020**, *139*, 50–59. [[CrossRef](#)]
19. Khan, F.; Tahir, M.; Khelifi, F. Novel Geometric Features for Offline Writer Identification. *Pattern Anal. Appl.* **2016**, *19*, 699–708.
20. Bhunia, A.K.; Alaei, A.; Roy, P.P. Signature verification approach using fusion of hybrid texture features. *Neural Comput. Appl.* **2019**, *31*, 8737–8748. [[CrossRef](#)]
21. Luo, X.; Zhou, M.; Li, S.; Wu, D.; Liu, Z.; Shang, M. Algorithms of Unconstrained Non-Negative Latent Factor Analysis for Recommender Systems. *IEEE Trans. Big Data* **2021**, *7*, 227–240. [[CrossRef](#)]
22. Ruiz, V.; Linares, I.; Sanchez, A.; Velez, J.F. Offline handwritten signature verification using compositional synthetic generation of signatures and Siamese Neural Networks. *Neurocomputing* **2020**, *374*, 30–41. [[CrossRef](#)]
23. Hu, J.; Chen, Y. Offline Signature Verification Using Real Adaboost Classifier Combination of Pseudo-dynamic Features. In Proceedings of the International Conference on Document Analysis & Recognition, Washington, DC, USA, 25–28 August 2013; pp. 1345–1349.
24. Hafemann, L.; Sabourin, R.; Oliveira, L. Learning Features for Offline Handwritten Signature Verification using Deep Convolutional neural networks. *Pattern Recognit.* **2017**, *70*, 163–176. [[CrossRef](#)]
25. Jain, A.; Singh, S.K.; Singh, K.P. Signature verification using geometrical features and artificial neural network classifier. *Neural Comput. Appl.* **2020**, *12*, 6999–7010. [[CrossRef](#)]
26. Dutta, A.; Pal, U.; Lladós, J. Compact Correlated Features for Writer Independent Signature Verification. In Proceedings of the International Conference on Pattern Recognition (ICPR), Cancún, Mexico, 4–8 December 2016; pp. 3422–3427.
27. He, S.; Schomaker, L. Writer Identification Using Curvature Free Features. *Pattern Recognit.* **2017**, *63*, 451–464. [[CrossRef](#)]
28. Li, L.; Huang, L.; Yin, F.; Chen, Y. Offline signature verification using a region based deep metric learning network. *Pattern Recognit.* **2021**, *118*, 108009. [[CrossRef](#)]
29. Li, C.; Lin, F.; Wang, Z.; Yu, G.; Yuan, L.; Wang, H. DeepHSV: User-Independent Offline Signature Verification Using Two-Channel CNN. In Proceedings of the International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 166–171.
30. Danilo, A.; Manoochehr, J.; Luigi, C.; Alessio, F.; Marco, R. R-SigNet: Reduced space writer-independent feature learning for offline writer-dependent signature verification. *Pattern Recognit. Lett.* **2021**, *150*, 189–196.
31. Neculoiu, P.; Versteegh, M.; Rotaru, M. Learning Text Similarity with Siamese Recurrent Networks. In Proceedings of the Repl4NLP Workshop at ACL2016, Berlin, Germany, 11 August 2016; pp. 148–157.
32. Wu, D.; Shang, M.; Luo, X.; Wang, Z. An L1-and-L2-Norm-Oriented Latent Factor Model for Recommender Systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *1*–14. [[CrossRef](#)]
33. Ho, S.L.; Yang, S.; Yao, Y.; Fu, W. Robust optimization using a methodology based on Cross Entropy methods. *IEEE Trans. Magn.* **2011**, *47*, 1286–1289. [[CrossRef](#)]
34. Lin, T.; Goyal, P.; Girshick, R. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *99*, 2999–3007.
35. Zhang, J.; Xing, M.; Sun, G.; Chen, J.; Li, M.; Hu, Y.; Bao, Z. Water Body Detection in High-Resolution SAR Images With Cascaded Fully-Convolutional Network and Variable Focal Loss. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 316–332. [[CrossRef](#)]
36. Xiao, W.; Wu, D. An Improved Siamese Network Model for Handwritten Signature Verification. In Proceedings of the 2021 IEEE International Conference on Networking, Sensing and Control (ICNSC), Xiamen, China, 3–5 December 2021; Volume 1, pp. 1–6.
37. Dey, S.; Dutta, A.; Toledo, J.I.; Ghosh, S.K.; Pal, U. Signet: Convolutional Siamese Network for Writer Independent Offline Signature Verification. *arXiv* **2017**, arXiv:1707.02131.
38. Kumar, R.; Sharma, J.D.; Chanda, B. Writer-independent offline signature verification using surroundedness feature. *Pattern Recognit. Lett.* **2012**, *33*, 301–308. [[CrossRef](#)]
39. Bharathi, R.K.; Shekar, B.H. Offline signature verification based on chain code histogram and Support Vector Machine. In Proceedings of the International Conference on Advances in Computing, Communications, and Informatics (ICACCI), Mysore, India, 22–25 August 2013; pp. 2063–2068.
40. Das, S.D.; Ladia, H.; Kumar, V.; Mishra, S. Writer Independent Offline Signature Recognition Using Ensemble Learning. In *ICDSMLA 2019: Proceedings of the 1st International Conference on Data Science, Machine Learning and Applications (Lecture Notes in Electrical Engineering, 601)*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2020.
41. Kumar, R.; Kundu, L.; Chanda, B.; Sharma, J.D. A Writer-Independent Off-line Signature Verification System based on Signature Morphology. In Proceedings of the 1st International Conference on Intelligent Interactive Technologies and Multimedia, Allahabad, India, 27–30 December 2010.
42. Pal, S.; Alaei, A.; Pal, U.; Blumenstein, M. Performance of an Offline Signature Verification Method Based on Texture Features on a Large Indic-Script Signature Dataset. In Proceedings of the 12th IAPR Workshop on Document Analysis Systems (DAS), Santorini, Greece, 11–14 April 2016.