applied sciences

MDPI

Article

FC-ResNet: A Multilingual Handwritten Signature Verification Model Using an Improved ResNet with CBAM

Yusnur Muhtar [1], Mahpirat Muhammat [2], Nurbiya Yadikar [1], Alimjan Aysa [3,*] and Kurban Ubul [1,3,*]

[1] School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China; ysnr@stu.xju.edu.cn (Y.M.); nurbiya611@163.com (N.Y.)
[2] Institue of International Culture and Exchange, Xinjiang University, Urumqi 830046, China; xmahpu@xju.edu.cn
[3] Key Laboratory of Xinjiang Multilingual Information Processing, Urumqi 830046, China
* Correspondence: alim@xju.edu.cn (A.A.); kurbanu@xju.edu.cn (K.U.)

Abstract: Offline signature verification is a widely used biometric method in finance, law, and administrative procedures. However, existing deep convolutional neural network models perform poorly on signature datasets that span different regions and ethnic people, while also suffering from problems such as large parameter counts and slow inference speeds. To address these issues, we propose an improved residual network model (FC-ResNet). This model introduces a convolutional block attention module into the classical residual network to adapt to the diversity and variability of signatures, while also compressing the model for lightweight deployment. Due to the lack of public, offline handwritten signature datasets for ethnic people, we collected a large-scale offline handwritten signature dataset, including genuine signatures and forged signatures in Chinese, Uyghur, Kazakh, and Kirgiz, totaling 38,400 images. Our FC-ResNet model achieved an accuracy of over 96% for each language in our self-built dataset, as well as accuracy rates of 96.21%, 98.42%, and 97.28% on the public datasets CEDAR, BHSig-B, and BHSig-H, respectively. Based on the above experimental results, our proposed model demonstrates great potential for both public and self-built signature datasets, while also exhibiting significant advantages in lightweight model deployment. We believe that this work can provide a feasible solution for ethnic people signature verification.

Keywords: offline signature verification (OSV); residual network (ResNet); model compression; convolutional block attention module (CBAM)

check for updates

Citation: Muhtar, Y.; Muhammat, M.; Yadikar, N.; Aysa, A.; Ubul, K. FC-ResNet: A Multilingual Handwritten Signature Verification Model Using an Improved ResNet with CBAM. *Appl. Sci.* **2023**, *13*, 8022. https://doi.org/10.3390/app13148022

Academic Editor: Andrea Prati

Received: 20 May 2023
Revised: 2 July 2023
Accepted: 6 July 2023
Published: 9 July 2023

Copyright: © 2023 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

Handwritten signature verification is the process of verifying the authenticity, legitimacy, and validity of handwritten signatures. All over the world, a large number of important financial, commercial, and judicial documents are signed every day, so verifying signatures has become a critical issue that requires attention. Considering the widespread use and the huge number of handwritten signatures, it has become particularly urgent and necessary to develop an automatic, accurate, and efficient signature verification technology. Offline Signature Verification (OSV) requires the signature images to be recorded in advance. Then, pattern recognition, computer vision, image processing, and other technologies are used to extract the signature's characteristic information for analysis and comparison in order to determine the authenticity of the signature.

Although OSV has made remarkable progress over the past few decades, there are still some challenges that need to be solved. First, there is a lack of public, offline handwritten signature datasets for ethnic people such as Uyghur, Kazakh, and Kirgiz research and applications for these ethnic peoples' signatures. Second, the diversity of signatures requires OSV algorithms to have the ability to generalize across different regions and nationalities so as to adapt to the application needs of different scenarios and user groups. Most people have a lot of variability in their signature styles, making it difficult for algorithms to

correctly verify different signatures from the same person. Third, modern deep learning models usually have a large number of parameters, which require a lot of memory and computing resources to store and process. This will not only slow down the calculation speed of the model but also increase the storage and calculation burden of the model in resource-constrained environments such as mobile devices.

The primary achievements of our research can be outlined as follows:

1. First, we constructed a signature dataset containing Uyghur, Kazakh, Kirgiz, and Chinese signatures. This dataset consists of 38,400 signature images from 200 individuals of each ethnic group;
2. In this study, we conducted experiments on a signature dataset spanning different regions and ethnicities using the ResNet18 network. Furthermore, we introduced the Convolutional Block Attention Module (CBAM) to enhance the model's generalization ability. The CBAM can adaptively adjust the weight and scale of feature maps so that the network can better adapt to different signature styles and features. This experiment proves that the CBAM is robust and adaptable to the diversity and variability of the signature itself;
3. By compressing the model, we have successfully reduced its storage and computational costs while maintaining high accuracy. This achievement provides an excellent solution for the application of the model to resource-constrained devices, such as embedded and mobile devices. As a result, our model can be widely used in various scenarios where resources are limited.

In the field of signature verification, due to the lack of public ethnic people datasets for research, we collected a large-scale offline handwritten signature dataset that includes Chinese, Uyghur, Kazakh, and Kirgiz. We tested our method on both our self-built dataset and public datasets. Through experiments, it is proved that our method has strong effectiveness and generality in the verification of cross-language signature datasets.

2. Related Works

OSV has been widely researched and applied in the past decades, and the review articles [1–5] provide insights into the advantages and limitations of existing methods, emphasizing opportunities for further research and development. With the continuous development of technology, more and more signature datasets are publicly released, such as CEDAR [6], MCYT-75 [7], BHSig260 [8], GPDS [9,10], and so on. These datasets not only promote the research and development in the field of OSV but also improve the application level of OSV technology. However, there is no large-scale ethnic people signature dataset, and there are certain difficulties in the research and application of OSV for ethnic people. Therefore, we need a new dataset to satisfy this demand.

Classic manual feature extraction methods, such as LBP, HOG, and GLCM, are usually used for OSV [11–14]. However, these handcrafted features are prone to problems in the face of noise and complex background interference and require a lot of time and effort to adjust. Therefore, they perform poorly when applied to some complex data.

To address the limitations of traditional handcrafted feature extraction methods in signature verification, researchers commonly employ deep learning techniques. For instance, Dey S et al. [15] proposed a Siamese CNN network based on contrastive loss for OSV in their study. Shariatmadari S et al. [16] proposed a layered CNN for feature learning from signature patches. Wei P et al. [17] proposed an Inverse Identification Network (IDN), which focuses on the characteristics of signature strokes. Recent work includes topological graphs [18], Siamese networks [19], deep metric learning for regions [20], RNN networks [21], graph neural network methods [22], models based on Capsule and CNN networks [23], and the encoder–decoder architecture approach [24]. By adopting deep learning methods, it can be widely applied in fields such as forensic identification and finance to improve the progress of issuing certifications and reduce work intensity. However, most of the current OSV methods only use a single or a small number of language signature datasets for model training and do not fully model multi-language signatures, which may

lead to poor performance in multi-language signature verification. Therefore, the further research and development of deep learning methods capable of modeling multilingual signatures are required to improve the accuracy and reliability of signature verification.

We propose a verification method for multilingual signature datasets spanning regions and ethnic groups, which fuses the ResNet18 [25] and the CBAM [26]. By adopting the ResNet18 network, we can enhance the model's generalization ability to better handle the differences between different languages. In addition, combined with the CBAM, the method is more robust and adaptable to the variability of the signature itself.

In modern computing applications, deep learning models have emerged as a critical technology. However, due to the constraints of computational resources, many devices are unable to support large models, thereby hindering the deployment of deep learning models on these devices. To address this issue, researchers have proposed model compression techniques, which employ a variety of methods to reduce the size and computational complexity of the models, enabling deep learning models to be applied on resource-constrained devices. These methods include weight pruning, low-rank decomposition, quantization, and knowledge distillation, which can effectively reduce the model size, accelerate computation, and have a minimal impact on model performance. Consequently, model compression techniques have become a viable means of enabling the application of deep learning models on resource-constrained devices. Researchers have made significant efforts to accelerate convolutional neural networks, which can be categorized into three types: optimization implementation [27], quantization [28], and structural simplification [29]. He Y et al. [30] proposed a new channel pruning method to accelerate ultra-deep convolutional neural networks. Inspired by the aforementioned works, the objective of this research is to transform the model into a compact ResNet18 and validate its performance in comparison to the original model through a series of experiments. This study aims to provide new insights and inspirations for the more efficient design of deep learning models.

3. Self-Built Signature Dataset

To bridge the gap in the field of multilingual signature verification, we established a multilingual offline handwritten signature dataset consisting of Chinese and ethnic people. This dataset includes signatures from ethnic people such as the Uyghur, Kazakh, and Kirgiz people, covering diverse cultural backgrounds in Central Asia. With this dataset, we are able to provide more comprehensive data support for multilingual OSV, as well as new resources for language learning and cross-cultural research. Figure 1 shows examples of the multilingual signature dataset.



Figure 1. Some samples of our self-built signature dataset. Each row represents a language, and from top to bottom, there are samples of signatures in Chinese, Uyghur, Kazakh, and Kirgiz. The two samples on the left are genuine signatures, while the two samples on the right are forged signatures.

To build a multilingual offline handwritten signature dataset, we followed a three-step process: collection, scanning, and segmentation. During the collection phase, volunteers were asked to write their names in their native language in 24 boxes on a sheet of A4 paper. Each individual user was given 24 corresponding forged signatures, which were imitated by volunteers of the same ethnicity according to their own writing style and habits. In the scanning stage, we used a Lenovo M7400 scanner to scan the collected signature images at 300 dpi and used Adobe Photoshop CC2022 software to crop and number sequentially and save them as BMP images. Finally, we used these images to create a multilingual dataset of offline handwritten signatures by removing signatures unsuitable for human experimentation, which we use for research alongside publicly available datasets.

To establish a comprehensive basis for research such as multilingual signature verification in Central Asia, we present a dataset that possesses several unique characteristics and challenges. First, the dataset is a large-scale multilingual signature dataset that includes four languages: Chinese, Uyghur, Kazakh, and Kirgiz. These languages represent multiple ethnicities and cover people from various regions and cultural backgrounds, thus exhibiting a high degree of diversity. Second, the handwritten styles and features of these signatures are also distinct, rendering the dataset more challenging. Third, the dataset includes a significant number of real, personal handwritten signatures that were collected manually at different times and in natural environments.

4. Methods

4.1. Overview

This article proposed a novel offline signature verification method (FC-ResNet) that is applicable to multi-lingual signatures spanning different regions and ethnic groups. The FC-ResNet model structure is shown in Figure 2. First, on the basis of the ResNet-18 structure, we use methods such as minimizing the size of the cropped convolution kernel and reducing the number of convolution kernel channels as much as possible to reduce the number of model parameters, maintain effectiveness, and convert it into a compact ResNet18. Second, we introduce the CBAM into the residual block, which helps the model to make better use of multi-scale feature information by combining the two attention modules of spatial attention (SA) and channel attention (CA) and improves the accuracy of the model for signature images.

4.2. ResNet18 Model Compression

When training deep neural networks, the backpropagation of gradients often results in weight updates that are either too small or too large, rendering the network untrainable. ResNet [25] introduced residual connections, which allow gradients to flow freely throughout the network, effectively addressing the issues of vanishing and exploding gradients in deep neural networks and making network training easier.

OSV is a typical verification problem based on limited features and a relatively small dataset. In this scenario, the use of deep neural network structures can easily lead to overfitting, thereby affecting the accuracy and robustness of the validation algorithm. To overcome this problem, it may be considered to use a shallow network structure instead of a deep neural network. Compared with ResNet34, ResNet50, and ResNet101, ResNet18 has a lighter network structure, fewer parameters, and computational complexity. Therefore, this model has a smaller size and memory footprint, making it easier to deploy on resource-limited devices. In addition, compared to deeper network structures, the training time and hyperparameter settings of ResNet18 are easier to adjust and handle.

This study aims to address the issues of the excessive parameters and slow inference speed of ResNet18 in offline handwritten signature verification by optimizing it through network compression techniques. Our main improvement strategy is to prune the weights with smaller contributions in the network to reduce the total number of parameters and optimize the model structure. The corresponding improvements are shown in Table 1. The

pruned network can achieve similar or even better performance compared to the original network while being more memory-efficient and computationally efficient.

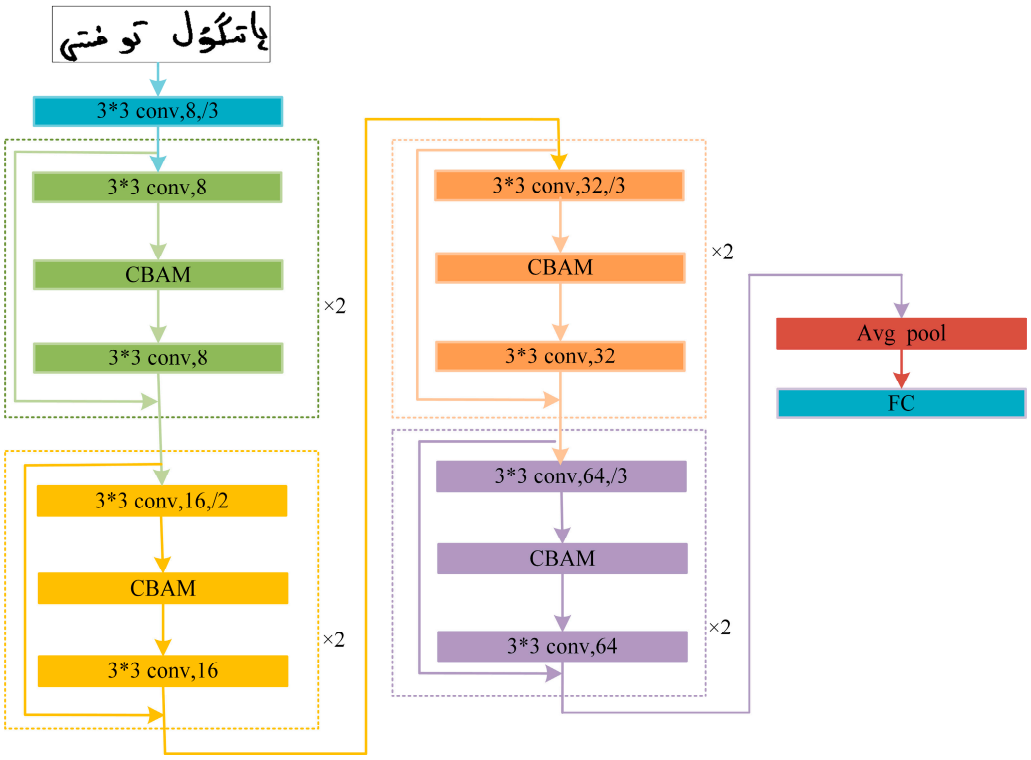


Figure 2. FC-ResNet model structure.

Table 1. Comparison of detailed structures between the ResNet18 and modified ResNet18 models.

Layer Name	ResNet18	Modified ResNet18
Conv1	7×7, 64, stride 2	3×3, 8, stride 3
Conv2_x	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 8 \\ 3 \times 3, & 8 \end{bmatrix} \times 2$
Conv3_x	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 16 \\ 3 \times 3, & 16 \end{bmatrix} \times 2$, stride 2
Conv4_x	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 32 \\ 3 \times 3, & 32 \end{bmatrix} \times 2$, stride 3
Conv5_x	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 2$, stride 3
FC	Average pool, 1000-d fc, softmax	Average pool, 2-d fc, softmax

As shown in the above table, the ResNet18 network consists of a 7×7 convolutional layer, two pooling layers, eight residual blocks, and a fully connected layer, where each residual unit consists of a two 3×3 convolutions-layer composition. For the compression of the ResNet18 network model, we mainly adopted the following two steps: First, reduce the channel number of the first convolution layer from 64 to 8, and replace the 7×7 convolution kernel with a step size of 2 in a 3×3 convolution kernel with a stride of 3. Doing so can effectively reduce the number of parameters and calculations in the model and reduce the risk of overfitting. Second, within each residual block, the number of channels of the convolution kernel is reduced to 1/8 of the current number of channels. Specifically, the number of channels of the two convolutional layers in each residual block is gradually reduced until the number of channels in the residual block of the last layer is 64. Both steps aim to reduce the complexity of the model, making it more lightweight and efficient.

The first step reduces the parameter size and computational cost of the model by decreasing the number of channels in the convolutional layers and using smaller convolu-

tion kernels. The second step compresses the network structure effectively by gradually reducing the channel numbers of the convolution kernels. These operations can reduce the model size without compromising its performance. Through the organic combination of these two steps, we successfully achieved the compression of the ResNet18 network model.

4.3. Improved Residual Block

In this study, we observed that ResNet18 employs fixed convolution kernels and channels, which restricts the model's ability to learn the correlations and importance between different features, thus limiting the model's representation capacity. To address this issue, we introduced CBAM into the residual block to better learn the correlations between and importance of different feature channels and spatial positions. The incorporation of an attention mechanism can improve the model's adaptability to different signature styles and features, making the model more robust and with a stronger generalization ability, particularly in verifying multilingual signatures across different regions and ethnic groups. Moreover, the CBAM attention module has a lightweight structure, which can enhance the model's performance without increasing the number of network parameters.

In our study, we utilized a combination of convolutional layers, batch normalization layers, and activation functions in each residual block to extract features and enhance the stability and convergence speed of the network. The convolutional layers can extract spatial features from the images, the batch normalization layers can accelerate the training process and improve the stability of the network, and the activation functions can introduce non-linear factors. To further strengthen the network's attention control over the input feature map, we introduce the CBAM module. This module has a lightweight structure that does not increase the network's complexity. Meanwhile, by adopting residual connections that directly add the input to the output, we achieve faster training and higher accuracy.

The CBAM [26] is a hybrid attention mechanism that aids the neural network in better focusing on target regions containing important information in both channel and spatial dimensions, while suppressing irrelevant information, thereby enhancing model performance and generalization. Specifically, the CBAM module first generates the channel mask using the CA module to strengthen the focus on important channel features and then generates the spatial mask using the SA module to further attend to the regions containing important information in the image. CBAM is an efficient and lightweight attention module that can be seamlessly integrated into various convolutional neural network architectures and can be end-to-end trained with baseline models.

The CBAM module is divided into two components, the CA module and the SA module, as illustrated in Figure 3. Initially, the feature map is fed into the CA module, which produces a corresponding attention map. The input feature map is then multiplied element-wise with the attention map and fed into the SA module, where the same element-wise multiplication is performed. Finally, the output feature map is obtained. The purpose of CA is to focus on which specific features are meaningful for a given task. The specific structure of the CA module is shown in Figure 4.

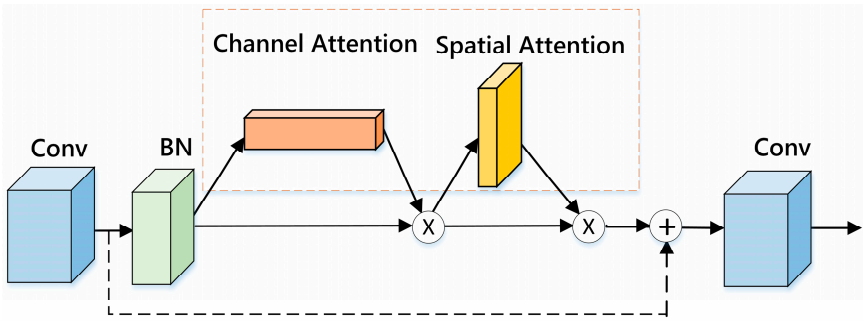


Figure 3. Structure diagram of the improved residual block.

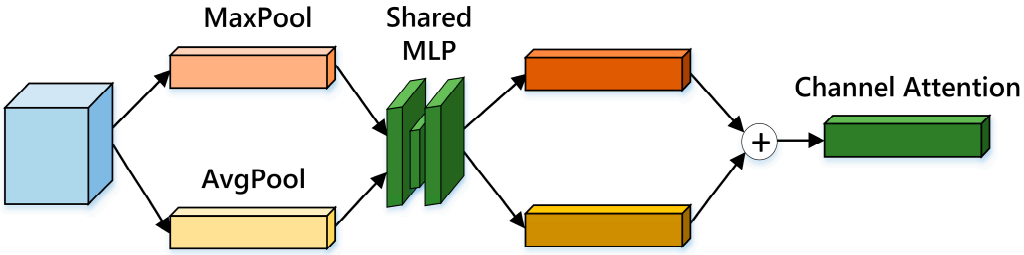


Figure 4. Structure of CA.

The CA module of the CBAM is different from the SE module in that it uses average pooling and max pooling to obtain global receptive field information. Then, these two $1 \times 1 \times c$ feature maps are fed into a shared MLP for feature extraction. Finally, the two feature maps are fused, and the result is converted to the range (0, 1) through a sigmoid function for feature recalibration and combined with the original feature map. The detailed operational steps of CBAM are as follows:

$$\text{Attention}_c(F) = \sigma\left(\text{MLP}\left(G_{AVG(F)}\right) + \text{MLP}\left(G_{MAX(F)}\right)\right) \quad (1)$$

In the expression, F represents the input 2D feature map, and G_AVG and G_MAX are used to obtain feature information under average pooling and max pooling, respectively. Then, these two feature maps are separately fed into a shared, fully connected layer for feature extraction and finally activated. After the CA module, the SA module is used to focus on which positions in the feature map are meaningful for a particular task. The specific structure of the SA module is shown in Figure 5.

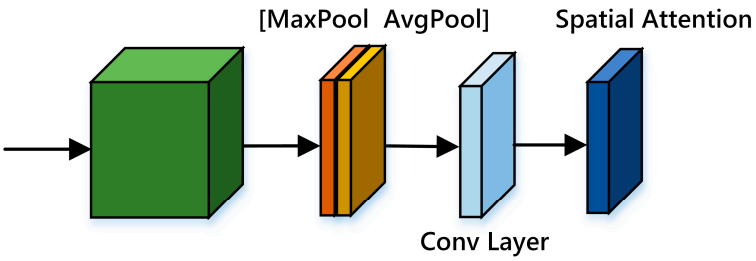


Figure 5. Structure of SA.

In CBAM, the CA and SA modules have different objectives. The CA module aims to evaluate the effectiveness of each channel, thus transforming the $h \times w \times c$ feature map into a $1 \times 1 \times c$ map. On the other hand, the SA module aims to evaluate the effectiveness of the spatial positions of the feature map, which requires it to be transformed into an $h \times w \times 1$ map. To achieve this, CBAM first employs max pooling and average pooling operations to transform the feature map into the $1 \times 1 \times c$ form and then performs feature extraction using a 7×7 convolutional network. Finally, they are combined with the original inputs for feature recalibration. The detailed implementation is as follows:

$$\text{Attention}_s(F) = \sigma(\text{Conv}_{7\times7}([\text{AVG}(F); \text{MAX}(F)])) \quad (2)$$

5. Experiments

The primary focus of this study lies in writer-dependent offline handwritten signature verification, employing signature datasets from seven different languages. Specifically, Chinese, Uyghur, Kazakh, and Kirgiz scripts are included, each consisting of 200 individuals. For each individual, there are 24 genuine samples and 24 forged samples. Additionally, to assess the effectiveness of the experiments, two publicly available datasets, namely, CEDAR and BHSig260, were incorporated. Based on the foundations of other relevant works, we allocated 80% of the samples for training, reserving the remaining 20% for testing. Conse-

quently, our self-constructed dataset encompasses a total of 38,400 handwritten signature samples, far exceeding the combined scale of the three publicly available datasets, which comprise a mere 16,640 signature samples. This demonstrates the substantial applicability of our established dataset in offline handwritten signature verification research, spanning multiple languages.

In the experiment, we employed a single NVIDIA GeForce RTX 3090 graphics card with 24GB of memory, an Intel(R) Xeon(R) Platinum 8338C CPU operating at 2.60GHz, and 80GB of RAM. The system configuration consisted of Ubuntu 20.04, Python 3.8, PyTorch 1.11.0, Cuda 11.3 for code compilation, and PyCharm 2021.1 Professional as the integrated development environment. Prior to training, the images were resized to 224 $\times$ 224 pixels, and a batch size of 32 was set. We utilized the Adam optimizer with a learning rate of 5×10^{-4} and the weight decay set to 5×10^{-2}.

5.1. Preprocessing

Image preprocessing technology can effectively remove irrelevant information while amplifying or enhancing useful information. For offline handwritten signature images, different preprocessing operations are usually performed based on specific feature selection to obtain the maximum amount of information. This study designs a preprocessing technique aimed at ensuring that the signature image is cropped to a tight region with minimal background pixels. The preprocessing steps encompass image grayscale conversion, normalization, binarization, and smoothing for artifact removal. Finally, all signature samples are resized to a standard 224 $\times$ 224 pixels to eliminate height and width variation between different signature samples. The use of this preprocessing technique is of great significance in OSV, which can improve the robustness of the model.

5.2. Evaluation Metrics

In the study of OSV, three common evaluation metrics were employed to assess the performance of the proposed method: False Rejection Rate (FRR), False Acceptance Rate (FAR), and Accuracy (ACC). The use of these evaluation indicators can help researchers carry out reasonable algorithm design and selection. Among them, the FAR refers to the probability of falsely detecting a genuine signature as a forged signature, that is, the ratio of failing to successfully identify a genuine signature. The FRR refers to the probability of mistaking a forged signature as a genuine signature, that is, the rate of failing to successfully identify a forged signature. The accuracy rate is one of the indicators for the comprehensive evaluation of the performance of the multi-language OSV algorithm. It refers to the ratio of the correct verification of genuine signatures and forged signatures after all the test samples are verified. Therefore, a higher accuracy indicates that the proposed algorithm can better verify the signature, demonstrating stronger discriminability and better processing capabilities.

5.3. Results and Discussion

To evaluate the effectiveness of our proposed method and investigate the impact of different modules on the overall performance, we conducted ablation experiments. The CEDAR dataset is classic and has a moderate size; thus, it is more authoritative. The evaluation on the Uyghur dataset can further improve the verification ability and practicability of the method in this paper. We have summarized and organized the results of the ablation experiments and presented them in Table 2 for a more intuitive and accurate representation of the experimental results.

5.3.1. Results of the Different Model

In this section, we finally selected the baseline model of the full text by comparing the parameter quantity and accuracy of several ResNet networks. ResNet18 and ResNet50 are very popular residual networks widely used for computer vision tasks. They are both part of the ResNet, which uses residual blocks in deep neural networks. Compared with

ResNet50, ResNet18 has a lighter network structure, fewer parameters, and less calculation. Therefore, ResNet18 is easier to deploy on devices with limited resources. We compared the performance of the network model regarding signature verification using the CEDAR public dataset and the self-built Uyghur dataset. Through these experiments, we chose the baseline model that is most suitable for our research purposes.

Table 2. Results obtained from the ablation study.

Method		CEDAR			Uyghur			Parameter (M)	Memory (MB)
Compression	CBAM	FAR (%)	FRR (%)	ACC (%)	FAR (%)	FRR (%)	ACC (%)		
×	×	7.58	2.27	95.08	5.83	2.88	95.65	11.18	31.03
✓	×	6.06	4.54	94.70	6.15	2.92	95.47	0.08	1.83
×	✓	6.44	1.52	96.02	3.75	3.23	96.51	11.87	31.14
✓	✓	5.30	2.27	96.21	4.17	3.02	96.41	0.08	2.06

According to the experimental results shown in Table 3, on the CEDAR dataset, the accuracy rate of ResNet18 is 95.08%, and the accuracy rate of Resnet101 is the highest, which is 95.64%. However, Resnet101 has the largest number of parameters and the most memory usage, and its verification effect has no obvious advantage compared with ResNet18.

Table 3. Performance indicators of different models on the CEDAR public dataset.

Method	FAR (%)	FRR (%)	ACC (%)	Parameter (M)	Memory (MB)
ResNet101	6.06	2.65	95.64	44.5	202.3
ResNet50	6.43	3.03	95.27	25.51	134.3
ResNet34	7.19	3.41	94.69	21.28	45.89
ResNet18	7.58	2.27	95.08	11.18	31.03

As shown in Figure 6, on the Uyghur language dataset, Resnet101 still has the largest number of parameters, but it can only achieve 86.61% accuracy, 9.04% lower than that of ResNet18. Compared to the others, ResNet18 achieved the best results of two different signature datasets and had the lowest number of parameters and the lowest memory usage. Considering both computational performance and model effectiveness, ResNet18 was chosen as the baseline model in this paper.

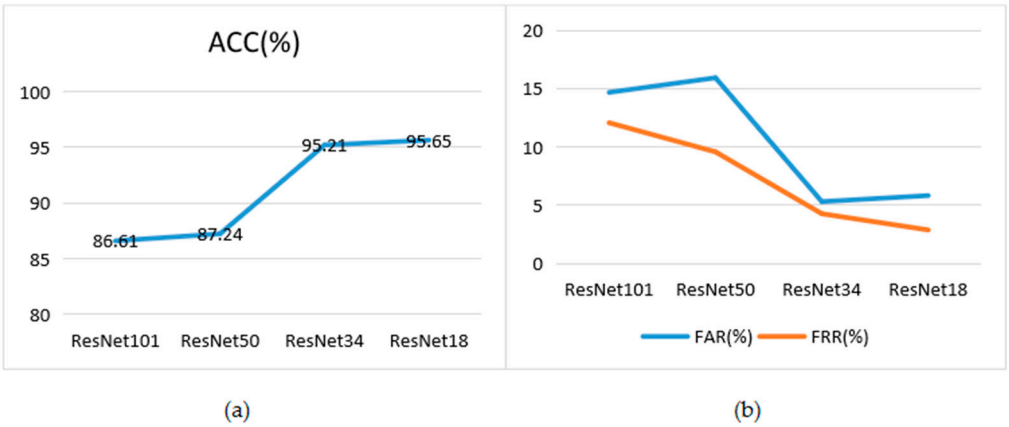


Figure 6. Performance indicators of different models on the self-built Uyghur dataset. (**a**,**b**) show the results of FAR, FRR, and ACC indicators, respectively.

5.3.2. Results of the Model Compression

We use a network pruning technique based on ResNet18 to reduce storage and computation costs while maintaining high accuracy. It has been verified by experiments that this

operation was successful. Specifically, the number of model parameters after compression is reduced from 11.18 M to 0.08 M, and the memory usage is also reduced from 31.03 MB to 2.06 MB, which shows that the compression technology is very effective in model optimization. The use of this technology can not only reduce the storage and computing overhead of the model but also improve the performance and efficiency of the model in practical applications, and it has broad application prospects.

5.3.3. Results of the CBAM Fusion

According to the experimental results in Table 3, we observe that the discrimination accuracy of Resnet101 is slightly higher than that of ResNet18, without adding the attention module. This is notable because the former has a deeper network with 101 convolutional layers, while the latter has only 18, suggesting that the network depth plays a significant role in model performance. In Table 2, we also show the result that the accuracy is significantly improved after adding the CBAM in the residual block of ResNet18. For example, on the CEDAR dataset, after adding the CBAM, the accuracy rate increased from the initial 95.08% to 96.02%. The CBAM module enables the improved ResNet18 to achieve better performance relative to ResNet101 by combining spatial and channel information while maintaining the same structure. In addition, the method improved the accuracy rate by 0.86% on the Uyghur language dataset, while the number of parameters and the memory usage did not increase significantly. These results show that the CBAM effectively enhances the performance and expressiveness of the model by learning different signature styles and features, leading to more accurate signature verification results.

5.3.4. Results of the FC-ResNet

Based on the above experimental results, we yielded comprehensive statistics on the accuracy of the CEDAR dataset and the Uyghur language dataset, and the results are shown in Table 2. On the CEDAR dataset, the accuracy rate is 96.21%, indicating that the method proposed in this paper is still effective for signature data with few samples and has good performance. In addition, on the Uyghur language dataset, the accuracy rate is 96.41%, indicating that the model can accurately identify signature samples across languages.

5.3.5. Results of the Different Parameters

To evaluate the performance of the proposed method and find the best experimental results, we conducted a series of experiments on the CEDAR dataset, testing different parameter conditions. We tested different batch sizes and plotted the experimental results in Figure 7. The experimental results show that the accuracy does not strictly follow a linear growth relationship and may decrease as the batch size increases. Based on our experimental results, we found that using a batch size of 32 achieved the best performance for our proposed method on the CEDAR dataset, with an accuracy rate of 96.21%. This result provides important foundational support for our subsequent experiments.

5.3.6. Results of the Self-Built Datasets

To further illustrate the generality of our proposed method, we conduct experiments on our self-built signature datasets from four different languages. Compared with the commonly used CEDAR and BHSig260 public datasets, our dataset has a larger scale, which is beneficial for the training and convergence of deep learning networks.

Table 4 shows the performance of FC-ResNet on the self-built dataset. The dataset includes signature data in four languages: Chinese, Uyghur, Kazakh, and Kirgiz. The FAR and FRR of this method are kept low, and the accuracy rate has exceeded 96%. Our FC-ResNet exhibits high accuracy and reliability in solving the problem of OSV across different languages and cultural backgrounds.

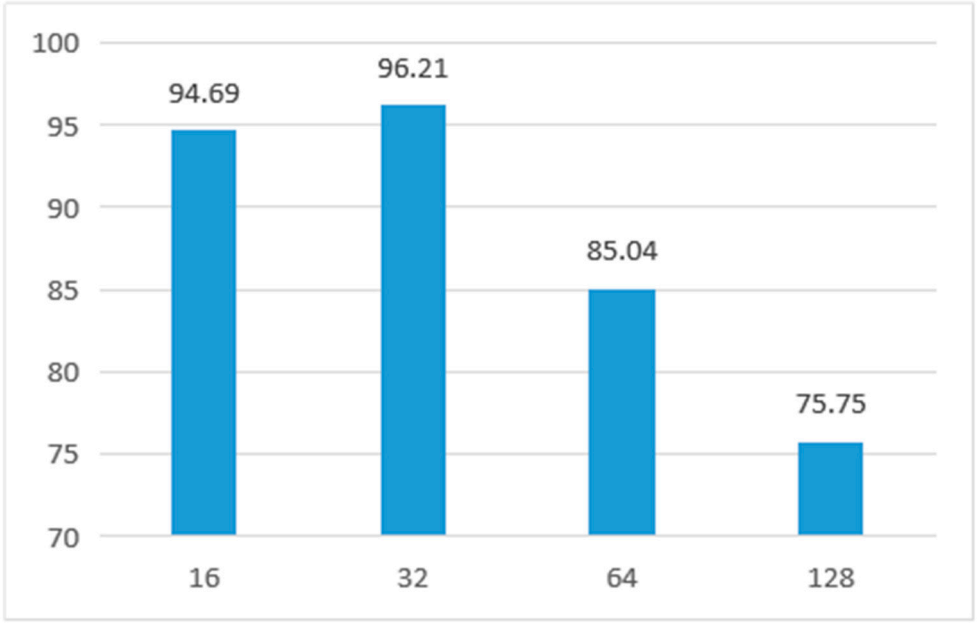


Figure 7. Accuracy rate (%) of our proposed method at different batch sizes.

Table 4. Experimental results of FC-ResNet on self-built datasets.

Dataset	FAR (%)	FRR (%)	ACC (%)
Chinese	3.44	3.13	96.72
Uyghur	4.17	3.02	96.41
Kazakh	4.48	3.23	96.15
Kirgiz	3.48	3.18	96.67

5.4. Comparison with the Classical Method

The method in this paper is compared with the existing state-of-the-art methods on indicators such as FAR, FRR, and ACC on different datasets. Among them, the dataset contains CEDAR, BHsig260, and the Chinese dataset, and there is no order restriction between signature images in the experimental setting. Through such an experimental design, we can more comprehensively evaluate the advantages of each method.

The CEDAR signature dataset [6] contains signature samples of English names. It is composed of 55 individuals' samples, and each individual has 24 genuine and 24 forged signatures. We compare other methods: Grid Feature [31], Surroundedness Feature [32], CBCapsNet [23], DeepHSV [33], HOCCNN [16] and SigNet-F [34]. Table 5 shows the results of the various methods, with CBCapsNet and DeepHSV achieving the best accuracy performance of 100% for the validation of the CEDAR dataset. After an experimental evaluation of this dataset, the fact that our FC-ResNet performs higher in terms of all of the reported evaluation metrics confirms the superior performance of our FC-ResNet.

Table 5. Comparison on the CEDAR Dataset.

Authors	Method	FAR (%)	FRR (%)	ACC (%)
Zois. et al. [31]	Grid Feature	11.52	5.83	91.32
Kumar R. et al. [32]	Surroundedness Feature	8.33	8.33	91.67
Parcham E. et al. [23]	CBCapsNet	0	0	100
Li C. et al. [33]	DeepHSV	0	0	100
Shariatmadari S. et al. [16]	HOCCNN	5.07	4.79	95.06
Hafemann L G. et al. [34]	SigNet-F	-	-	95.37
Ours	FC-ResNet	5.30	2.27	96.21

The BHSig260 dataset [8] contains two subsets: the BHSig-B Dataset and BHSig-H Dataset. The BHSig-B Dataset contains signature samples of Bengali names. It contains 100 individuals' signature samples. Each individual has 24 genuine signatures and 30 forged signatures. We compare other methods, including IDN [17], CBCapsNet [23] and SURDS [24]. Tables 6 and 7 show the results of different methods on the two datasets, respectively, with the CBCapsNet [23] method on the BHSig-H dataset showing the highest accuracy of 100%. On these two datasets, our proposed method has obvious advantages over other methods. In particular, our proposed method has low FAR and FRR and high accuracy, and it performs well on the signature verification problem on three public datasets, verifying its effectiveness and generalization ability. In addition, our model has fewer parameters than other models, has the advantage of being lightweight, and is suitable for resource-constrained practical application scenarios.

Table 6. Comparison on the BHSig-B dataset.

Authors	Method	FAR (%)	FRR (%)	ACC (%)
Li C. et al. [33]	DeepHSV	-	-	88.08
Wei P. et al. [17]	IDN	4.12	5.24	95.32
Lu X. et al. [35]	STN	3.96	3.96	96.04
Parcham E. et al. [23]	CBCapsNet	5.11	6.29	94.3
Li H. et al. [36]	AVN	5.07	7.33	93.8
Jain A. et al. [37]	Geometrical features	-	-	97.79
Chattopadhyay S. et al. [24]	SURDS	19.89	5.42	87.34
Xamxidin N. et al. [38]	IDN + SE + ESA	1.5	3.14	97.17
Zhang X. et al. [39]	MA-SCN	9.96	5.85	92.86
Longjam T. et al. [40]	CNN	-	-	95
Ours	FC-ResNet	0.73	2.42	98.42

Table 7. Comparison on the BHSig-H dataset.

Authors	Method	FAR (%)	FRR (%)	ACC (%)
Li C. et al. [33]	DeepHSV	-	-	86.66
Wei P. et al. [17]	IDN	8.99	4.93	93.04
Lu X. et al. [35]	STN	5.97	5.97	94.03
Parcham E. et al. [23]	CBCapsNet	0	0	100
Li H. et al. [36]	AVN	5.46	5.91	94.32
Jain A. et al. [37]	Geometrical features	-	-	95.29
Chattopadhyay S. et al. [24]	SURDS	12.01	8.98	89.50
Xamxidin N. et al. [38]	IDN + SE + ESA	2.31	6.65	96.86
Zhang X. et al. [39]	MA-SCN	5.73	4.86	94.99
Longjam T. et al. [40]	CNN	-	-	90.00
Ours	FC-ResNet	3.13	2.31	97.28

Our Chinese signature dataset includes the signature samples of 200 individuals. We compared other methods, including IDN [17], STN [35], DenseNet-36 [20], IDN+SE+ESA [38], etc. Table 8 shows the results of various methods. According to the experimental results, our FC-ResNet model shows better results for all the evaluation metrics, indicating that our proposed approach has a higher generality and robustness. In summary, our proposed method shows excellent experimental results for different datasets, which fully proves its effectiveness.

Table 8. Comparison on the Chinese signature dataset.

Authors	Individual	Method	FAR (%)	FRR (%)	ACC (%)
Wei P. et al. [17]	749	IDN	11.52	5.47	90.17
Lu X. et al. [35]	1243	STN	10.21	10.21	89.79
Liu L. et al. [20]	1243	DenseNet-36	-	-	90.09
Xamxidin N. et al. [38]	100	IDN + SE + ESA	4.46	5.2	95.17
Ours	200	FC-ResNet	3.44	3.13	96.72

6. Conclusions

The application of deep learning methods in OSV faces many challenges, among which the diversity and variability of signatures make it difficult for algorithms to correctly verify different signatures of different user groups or the same person. To address these challenges, we propose a new OSV model that aims to improve the accuracy while making the model lightweight. Due to the lack of public, offline handwritten signature datasets for ethnic people, we collected a large-scale signature dataset that includes genuine and forged signature images in Chinese, Uyghur, Kazakh, and Kirgiz. We used ResNet18 as the base model, which is a classic residual network model with a smaller model size and better performance. Although the ResNet18 model has the advantages of ease of implementation and training, it does not perform well for the diversity and variability of signature datasets spanning different languages and ethnicities. The contribution of this paper is that, by introducing the CBAM, the relationship between different feature channels and spatial locations is better captured, thus improving the adaptability of the model to different signature styles and features. Meanwhile, we compressed the model and successfully reduced its storage and computational costs while maintaining high accuracy. The evaluation results show that our FC-ResNet models exhibit significant advantages on both self-built and public datasets, thus demonstrating the strong effectiveness and generality of our method on cross-lingual datasets. In future work, we will aim to further expand our datasets, optimize our model, and apply it to a wider range of domains for more accurate and robust signature verification.

Author Contributions: Conceptualization, Y.M.; methodology, Y.M.; software, Y.M.; investigation, Y.M.; validation, Y.M.; resources, M.M., N.Y. and K.U.; writing—original draft preparation, Y.M. and A.A.; writing—review and editing, Y.M. and A.A.; visualization, Y.M.; project administration, M.M., N.Y. and K.U.; funding acquisition, A.A. and K.U. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 62266044, 62061045, 61163028, 61862061) and the Science and Technology Plan Project of Xinjiang Uyghur Autonomous Region, China (2021D01C119).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The self-built datasets used in this paper, Chinese, Uyghur, Kazakh and Kirgiz datasets, have not been made public. This is the dataset collected by our college, which will be published soon.

Acknowledgments: The authors are very thankful to the editor and the referees for their valuable comments and suggestions for improving the paper.

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

References

1. Hafemann, L.G.; Sabourin, R.; Oliveira, L.S. Offline handwritten signature verification—Literature review. In Proceedings of the Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), Montreal, QU, Canada, 28 November–1 December 2017; pp. 1–8.
2. Diaz, M.; Ferrer, M.A.; Impedovo, D.; Malik, M.I.; Pirlo, G.; Plamondon, R. A perspective analysis of handwritten signature technology. *Acm Comput. Surv.* **2019**, *51*, 1–39. [CrossRef]
3. Hameed, M.M.; Ahmad, R.; Kiah, M.L.M.; Murtaza, G. Machine learning-based offline signature verification systems: A systematic review. *Signal Process. Image Commun.* **2021**, *93*, 116139. [CrossRef]
4. Muhtar, Y.; Kang, W.; Rexit, A.; Ubul, K. A Survey of Offline Handwritten Signature Verification Based on Deep Learning. In Proceedings of the 3rd International Conference on Pattern Recognition and Machine Learning (PRML), Chengdu, China, 15–17 July 2022; pp. 391–397.
5. Kaur, H.; Kumar, M. Signature identification and verification techniques: State-of-the-art work. *J. Ambient. Intell. Humaniz. Comput.* **2023**, *14*, 1027–1045. [CrossRef]
6. Kalera, M.K.; Srihari, S.; Xu, A. Offline signature verification and identification using distance statistics. *Int. J. Pattern Recognit. Artif. Intell.* **2004**, *18*, 1339–1360. [CrossRef]
7. Fierrez-Aguilar, J.; Nanni, L.; Lopez-Penalba, J.; Ortega-Garcia, J.; Maltoni, D. An on-line signature verification system based on fusion of local and global information. In Proceedings of the Audio-and Video-Based Biometric Person Authentication: 5th International Conference, AVBPA, Hilton Rye Town, NY, USA, 20–22 July 2005.
8. Pal, S.; Alaei, A.; Pal, U.; Blumenstein, M. Performance of an off-line signature verification method based on texture features on a large indic-script signature dataset. In Proceedings of the 12th IAPR Workshop on Document Analysis Systems (DAS), Santorini, Greece, 11–14 April 2016; pp. 72–77.
9. Ferrer, M.A.; Alonso, J.B.; Travieso, C.M. Offline geometric parameters for automatic signature verification using fixed-point arithmetic. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 993–997. [CrossRef]
10. Ferrer, M.A.; Diaz-Cabrera, M.; Morales, A. Synthetic off-line signature image generation. In Proceedings of the International Conference on Biometrics (ICB), Madrid, Spain, 4–7 June 2013; pp. 1–7.
11. Yilmaz, M.B.; Yanikoglu, B.; Tirkaz, C.; Kholmatov, A. Offline signature verification using classifier combination of HOG and LBP features. In Proceedings of the International Joint Conference on BIOMETRICS (IJCB), Washington, DC, USA, 11–13 October 2011; pp. 1–7.
12. Serdouk, Y.; Nemmour, H.; Chibani, Y. Combination of OC-LBP and longest run features for off-line signature verification. In Proceedings of the Tenth International Conference on Signal-Image Technology and Internet-Based Systems, Marrakech, Morocco, 23–27 November 2014; pp. 84–88.
13. Serdouk, Y.; Nemmour, H.; Chibani, Y. Handwritten signature verification using the quad-tree histogram of templates and a support vector-based artificial immune classification. *Image Vis. Comput.* **2017**, *66*, 26–35. [CrossRef]
14. Batool, F.E.; Attique, M.; Sharif, M.; Javed, K.; Nazir, M.; Abbasi, A.A.; Iqbal, Z.; Riaz, N. Offline signature verification system: A novel technique of fusion of GLCM and geometric features using SVM. *Multimed. Tools Appl.* **2020**. [CrossRef]
15. Dey, S.; Dutta, A.; Toledo, J.I.; Ghosh, S.K.; Lladós, J.; Pal, U. Signet: Convolutional siamese network for writer independent offline signature verification. *arXiv* **2017**, arXiv:1707.02131.
16. Shariatmadari, S.; Emadi, S.; Akbari, Y. Patch-based offline signature verification using one-class hierarchical deep learning. *Int. J. Doc. Anal. Recognit.* **2019**, *22*, 375–385. [CrossRef]
17. 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, Long Beach, CA, USA, 15–20 June 2019; pp. 5764–5772.
18. Zois, E.N.; Zervas, E.; Tsourounis, D.; Economou, G. Sequential motif profiles and topological plots for offline signature verification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 13248–13258.
19. Souza, V.L.; Oliveira, A.L.; Cruz, R.M.; Sabourin, R. An investigation of feature selection and transfer learning for writer-independent offline handwritten signature verification. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 7478–7485.
20. Liu, L.; Huang, L.; Yin, F.; Chen, Y. Offline signature verification using a region based deep metric learning network. *Pattern Recognit.* **2021**, *118*, 108009. [CrossRef]
21. Ghosh, R. A Recurrent Neural Network based deep learning model for offline signature verification and recognition system. *Expert Syst. Appl.* **2021**, *168*, 114249. [CrossRef]
22. Roy, S.; Sarkar, D.; Malakar, S.; Sarkar, R. Offline signature verification system: A graph neural network based approach. *J. Ambient. Intell. Humaniz. Comput.* **2023**, *14*, 8219–8229. [CrossRef]
23. Parcham, E.; Ilbeygi, M.; Amini, M. Cbcapsnet: A novel writer-independent offline signature verification model using a cnn-based architecture and capsule neural networks. *Expert Syst. Appl.* **2021**, *185*, 115649. [CrossRef]
24. Chattopadhyay, S.; Manna, S.; Bhattacharya, S.; Pal, U. SURDS: Self-Supervised Attention-guided Reconstruction and Dual Triplet Loss for Writer Independent Offline Signature Verification. In Proceedings of the 26th International Conference on Pattern Recognition (ICPR), Montreal, QU, Canada, 21–25 August 2022; pp. 1600–1606.

25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
26. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 14–17 May 2018; pp. 3–19.
27. Bagherinezhad, H.; Rastegari, M.; Farhadi, A. Lcnn: Lookup-based convolutional neural network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7120–7129.
28. Rastegari, M.; Ordonez, V.; Redmon, J.; Farhadi, A. Xnor-net: Imagenet classification using binary convolutional neural networks. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Proceedings, Part IV. Springer International Publishing: Cham, Switzerland, 2016; pp. 525–542.
29. Jaderberg, M.; Vedaldi, A.; Zisserman, A. Speeding up convolutional neural networks with low rank expansions. *arXiv* **2014**, arXiv:1405.3866.
30. He, Y.; Zhang, X.; Sun, J. Channel pruning for accelerating very deep neural networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1389–1397.
31. 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]
32. Kumar, R.; Sharma, J.D.; Chanda, B. Writer-independent off-line signature verification using surroundedness feature. *Pattern Recognit. Lett.* **2012**, *33*, 301–308. [CrossRef]
33. 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 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, Australia, 20–25 September 2019; pp. 166–171.
34. Hafemann, L.G.; Sabourin, R.; Oliveira, L.S. Learning features for offline handwritten signature verification using deep convolutional neural networks. *Pattern Recognit.* **2017**, *70*, 163–176. [CrossRef]
35. Lu, X.; Huang, L.; Yin, F. Cut and compare: End-to-end offline signature verification network. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 3589–3596.
36. Li, H.; Wei, P.; Hu, P. Avn: An adversarial variation network model for handwritten signature verification. *IEEE Trans. Multimed.* **2021**, *24*, 594–608. [CrossRef]
37. Jain, A.; Singh, S.K.; Singh, K.P. Signature verification using geometrical features and artificial neural network classifier. *Neural Comput. Appl.* **2021**, *33*, 6999–7010. [CrossRef]
38. Xamxidin, N.; Mahpirat Yao, Z.; Aysa, A.; Ubul, K. Multilingual Offline Signature Verification Based on Improved Inverse Discriminator Network. *Information* **2022**, *13*, 293. [CrossRef]
39. Zhang, X.; Wu, Z.; Xie, L.; Li, Y.; Li, F.; Zhang, J. Multi-Path Siamese Convolution Network for Offline Handwritten Signature Verification. In Proceedings of the 2022 the 8th International Conference on Computing and Data Engineering, Bangkok, Thailand, 11–13 January 2022; pp. 51–58.
40. Longjam, T.; Kisku, D.R.; Gupta, P. Multi-scripted Writer Independent Off-line Signature Verification using Convolutional Neural Network. *Multimed. Tools Appl.* **2023**, *82*, 5839–5856. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.