Capturing Micro Deformations from Pooling Layers for Offline Signature Verification

Yuchen Zheng†, Wataru Ohyama‡, Brian Kenji Iwana† and Seiichi Uchida†,
† *Department of Advanced Information Technology, Kyushu University, Fukuoka, Japan*
Email: {yuchen, brian}@human.ait.kyushu-u.ac.jp, uchida@ait.kyushu-u.ac.jp
‡ *Saitama Institute of Technology, 1690 Fusaiji, Fukaya-shi, Saitama, Japan*
Email: ohyama@sit.ac.jp

***Abstract*—In this paper, we propose a novel Convolutional Neural Network (CNN) based method that extracts the location information (displacement features) of the maximums in the max-pooling operation and fuses it with the pooling features to capture the micro deformations between the genuine signatures and skilled forgeries as a feature extraction procedure. After the feature extraction procedure, we apply support vector machines (SVMs) as writer-dependent classifiers for each user to build the signature verification system. The extensive experimental results on GPDS-150, GPDS-300, GPDS-1000, GPDS-2000, and GPDS-5000 datasets demonstrate that the proposed method can discriminate the genuine signatures and their corresponding skilled forgeries well and achieve state-of-the-art results on these datasets.**

***Keywords*-Offline Signature Verification; Micro Deformations; Displacement Features; Max-pooling; Feature Extraction**

I. Introduction

Signature verification is a very challenging task that aims to verify the genuine signatures and the corresponding skilled forgeries for a specific user [11]–[13]. Generally, signature verification systems are divided into two categories: online and offline. For online systems, the data is collected as a sequence which includes the positions of the pen, pressure coordinate sequence, pen elevation coordinate sequence, etc [8]. For offline systems, the data is collected from digital images. Since the dynamic information is not available in offline signature verification systems, the task becomes very challenging. In addition, the forgeries can be deliberately imitated by practiced persons, which also increases the difficulty for the verification systems.

To build a robust offline signature verification system, extracting the discriminative features between the genuine signatures and skilled forgeries from original images plays a crucial role. Traditional verification systems apply handcrafted feature extractors to extract the feature representations from the original images, such as geometrical features [2], Local Binary Pattern (LBP) features [4], Scale Invariant Feature Transform (SIFT) features [15], etc. However, designing a good handcrafted feature often need to set different hyperparameters which are just suitable for the specific tasks. Due to the limitation of the handcrafted features, many deep learning based feature extractors are proposed in recent years [5], [7], [16], [22]. However, capturing the micro deformations or distortions between the genuine signatures and skilled forgeries is still difficult for the existing methods.

In this paper, we propose a novel CNN based method that extracts the location information (displacement features) of maximums in max-pooling operations [24], [25] and fuses it with the pooling features to capture the micro deformations or distortions between the genuine signatures and their corresponding skilled forgeries as a feature extraction procedure. The motivation of this paper is to prevent the max-pooling operation from removing the key micro deformations used to discriminate genuine signatures and corresponding skilled forgeries. In the feature extraction process, we train a CNN between the genuine signatures and skilled forgeries on a large scale dataset, named GPDS-10000 [3] to capture the different behaviors of genuine signatures and skilled forgeries. After CNN training process, we take the trained CNN as a feature extractor to obtain the discriminative features from original signature images. Then, we apply linear SVMs as the writer-dependent classifiers for each user to build the signature verification system and evaluate the learned features.

The contributions of this paper are summarized as follows.

- We extract the displacement features and fuse it with pooling features to capture micro deformations between genuine signatures and skilled forgeries for offline signature verification systems.
- We train a proposed CNN based model on a large scale dataset.
- We obtain the state-of-the-art results on GPDS-150, GPDS-300, GPDS-1000, GPDS-2000 and GPDS-5000 datasets.

The rest of the paper is organized as follows: Section II discusses some handcrafted features and deep learning based features related to our own. We introduce the proposed method to capture the micro deformations between the genuine signatures and skilled forgeries in Section III. Finally, we present the experimental results and discussion in Section IV, while Section V concludes this paper with remarks and future work.

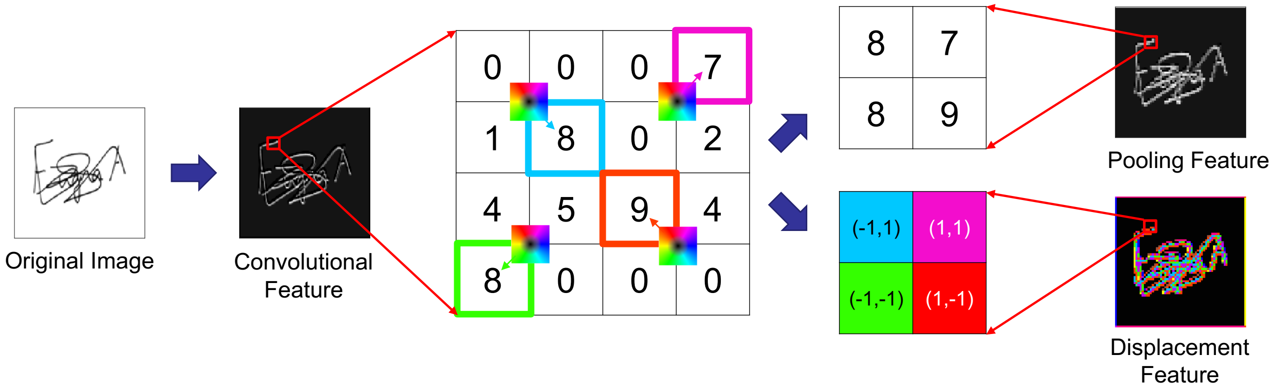

Figure 1: Extracting the pooling features and the displacement features simultaneously. Here, the pooling size is 2×2 with stride 2. The displacement features capture micro-deformation within the pooling window. The displacement vector (-1,1) means that the vertical displacement from the center to the maximum value is -1 and the horizontal displacement is 1.

II. Related Work

A. Handcrafted Features for Offline Signature Verification

Traditional offline signature verification systems often use different handcrafted features to train the writer-dependent or writer-independent classifiers, such as geometrical features [2], LBP features [4], SIFT features [15], and Histogram of Oriented Gradients (HOG) features [20]. More recently, there are many researches that focus on designing robust features [9], [14], [26] to build signature verification systems. In [14], Okawa proposed a feature extraction method based on a Fisher vector (FV) with fused "KAZE" features from both foreground and background signature images. The "KAZE" features consider the structures between strokes and stroke contour information more effectively. In [26], Zois et al. proposed the post-oriented grid features which encode the geometric structure of the signatures by grid templates. However, using the handcrafted features are hard to discriminate the genuine signatures and the corresponding skilled forgeries and often need to set different parameters for specific tasks, which is hard to apply to other verification systems and large scale applications.

B. Deep Learning Based Features for Offline Signature Verification

In recent years, the deep learning based models have widely applied in many fields, such as image classification and detection [18], [21], [27], text recognition and detection [17], [19], [23], online and offline signature verification [1], [6], [7], [16]. In the field of offline signature verification, some deep learning based features are proposed to capture the behaviors of different signatures [5], [7], [10], [22]. In [22], Zhang et al. proposed an unsupervised feature for offline signature verification based on Deep Convolutional Generative Adversarial Networks (DCGANs), which has a robust generalization ability compared to handcrafted features. In [7] and [5], Hafemann et al. proposed a CNN based feature extraction approach, named "Signet" to obtain the discriminative features not only between the genuine signatures and skilled forgeries but also between the different users. However, the "Signet" cannot capture the micro deformations between the genuine signatures and corresponding skilled forgeries and only trained on 531 different users cannot apply to large scale verification tasks.

Compared to the previous methods, the proposed method can train with a huge number of users to capture the micro deformations or distortions between the genuine signatures and skilled forgeries, which is very useful for signature verification systems.

III. Capturing Micro Deformations by Displacement Features

In this section, we introduce the proposed method to fuse the displacement features and pooling features as discriminative features to capture the micro deformations between the genuine signatures and skilled forgeries for signature verification systems. First, we introduce a normal CNN based architecture trained between the genuine signatures and forgeries. Then, based on the pre-trained CNN, we introduce how to extract the displacement features and fuse it with the pooling features in a combined architecture. Finally, we introduce how to train the writer-dependent classifiers based on the fused features.

A. Training a Normal CNN Between the Genuine Signatures and Forgeries

To distinguish genuine signatures and skilled forgeries, we design a CNN architecture with 3 convolutional and pooling layers, 2 fully-connected layers and a softmax layer. In the convolutional layers, the kernel size is 3×3 with stride 1, and the number of the filters is 32, 64 and 128, respectively. The pooling size is 2×2 with stride 2. In the fully-connected layers, the first fully-connected layer

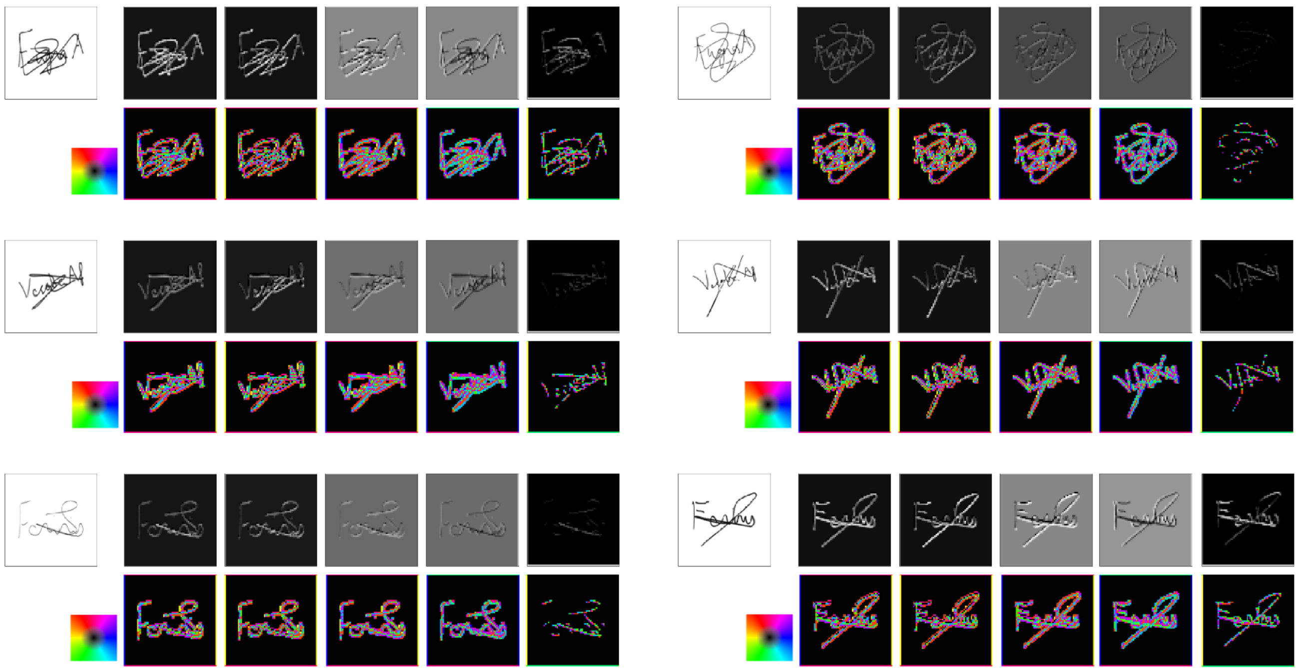

Figure 2: Visualization of the pooling features and the displacement features of some samples on GPDS dataset. The first row represents the original image and corresponding pooling features, the second row represents the displacement features. Each column represents one convolutional filter. The visualization of displacement features is based on an HSV color model whose color and intensity denote the direction and average length of the displacement features.

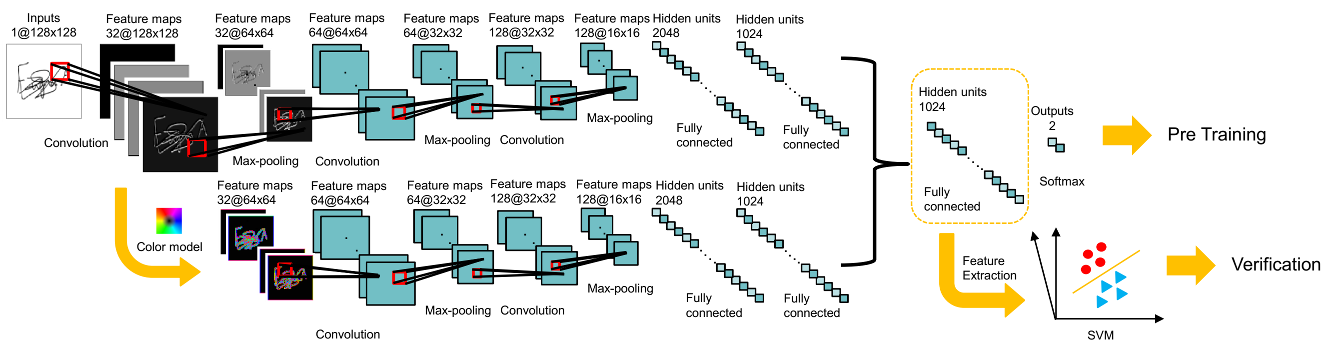

Figure 3: The procedure of feature extraction and verification.

has 2048 nodes and reduces to 1024 in the second fully-connected layer. Rectified Linear Unit (ReLU) is used as the activation function for the network. The final softmax layer contains 2 nodes, which is designed to judge whether the input is a genuine signature or forgery. The cross-entropy is used as the loss function to train the network.

B. Extracting and Fusing the Displacement Features with Pooling Features

To further capture the micro deformations between the genuine signatures and forgeries, we extract displacement features [24], [25] from the first pooling layer in previous pre-trained CNN and fuse it with the pooling features. Fig. 1 shows the procedure of the feature extraction. Here, the pooling size 2×2 with stride 2, the value of the displacement features both in horizontal and vertical directions belong to $[-1, 1]$. Fig. 2 presents the pooling features and displacement features of samples from the GPDS-10000 dataset based on a Hue-Saturation-Value (HSV) color model whose color and intensity denote the direction and average length of the displacement features. The displacement features describe the location information of maximums in max-pooling operation, which might capture some micro deformations of forgery signatures when some skilled writers imitated the genuine signatures.

The architecture that we used for fusing the pooling features and displacement features is shown in Fig. 3. We can see that the architecture for processing the displacement features is the same as the pre-trained CNN without the first convolutional and pooling layers. Here, we divide the displacement into the horizontal and vertical directions and apply the same architecture. We then fuse the pooling features and displacement features in the last fully-connected layer.

C. Training the Writer-dependent Classifiers

After the CNN training procedure, we extract the fused features of each user from the CNN based feature extractor and train the writer-dependent classifiers. For each user (not included in CNN training procedure), we use the genuine signatures as the positive samples and genuine signatures from other users as the negative samples to build the training set. Then, we choose the linear SVM as the writer-dependent classifier to build the verification system. This procedure is shown in Fig. 3.

To overcome the imbalanced problem that the negative samples are much more than the positive samples, we use different weights for the positive and negative class [7]. The SVM objective function becomes,

$$\begin{aligned} &\min \ \frac{1}{2}\mathbf{w}^T\mathbf{w} + C_p \sum_{\substack{i=1 \\ y_i=+1}}^{M} \xi_i + C_n \sum_{\substack{i=1 \\ y_i=-1}}^{N} \xi_i \\ &s.t. \quad y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i \ \text{ and } \ \xi_i \geq 0 \end{aligned} \tag{1}$$

where $\mathbf{x}_i$ is a training sample with target value y_i, ξ_i is the slack variables, M and N are the numbers of the positive and negative samples, C_p and C_n are the weights for the positive and negative class,

$$C_p = \frac{N}{M} C_n \tag{2}$$

For the testing procedure, we used the remaining genuine signatures of the target user and the skilled forgeries to build the test set to evaluate the performance.

IV. Experiment

In this section, we first introduce the experimental protocol of training CNN and the writer-dependent classifiers. Then, we present the experimental results in detail.

A. Experimental Protocol

We conducted the experiments on GPDS-10000 dataset [1] to evaluate the proposed method. The GPDS-10000 is a large scale dataset that contains 24 genuine signatures and 30 skilled forgeries for each user. The number of users is 10,000, so the GPDS-10000 dataset contains 240,000 genuine signatures and 300,000 skilled forgeries and it is very suitable for deep learning based methods. In our experiment, we divided the GPDS-10000 dataset into two parts, the first 5,000 users, and the final 5,000 users.

For the procedure of training CNN, we used the signatures from the final 5,000 users. To normalize the input images, we first resize all images to 128×128. Then, we used 90% data for training and 10% data for validation to train the network. We use Adam as an optimizer to minimize the loss function with mini-batch size 32. The model is trained with 40 epochs. The initial learning rate is set to 1e-4 and reduced by a factor of 0.95 after each epoch. After the CNN training process, we take the trained model as a feature extractor to extract the fused features.

To build the signature verification systems, we trained the linear SVMs as the writer-dependent classifiers for each user. Compared to other state-of-the-art methods, we used 5 sub-datasets, GPDS-150, GPDS-300, GPDS-1000, GPDS-2000, GPDS-5000 (the first 100, 150, 1,000, 2,000, 5,000 users of GPDS-10000 dataset) for final evaluation. For a specific user, we randomly selected 5 genuine signatures as the positive samples and 5 genuine signatures from each of the final 5,000 users as the negative samples as the training set. For the training process, the weights C_n are found by grid search with 5-fold cross-validation, and the C_p is calculated by Equation 2.

For the evaluation of the test set, the remaining genuine signatures from the target user are used for calculating the False Rejection Rate (FRR). The False Acceptance Rate for the skilled forgeries (FAR$_{skilled}$) experiment has been obtained with forgery samples of the target user. The False Acceptance Rate for the random impostor (FAR$_{random}$) experiment has been obtained with the genuine signatures from all the remaining users. The Equal Error Rate for skilled forgeries experiment (EER$_{skilled}$) is calculated by FAR$_{skilled}$ = FRR, and the EER for the random impostor experiment (EER$_{random}$) is calculated by FAR$_{random}$ = FRR.

B. Experimental Results and Discussion

To evaluate the performance of the proposed method, we test the proposed method on GPDS150, GPDS300, GPDS1000, GPDS2000, and GPDS5000 datasets and compared it with the state-of-the-art models [2], [4] and the traditional CNN based features with SVMs. In [2], the authors applied a Hidden Markov Model (HMM) on the geometrical features (GF) for verification systems. In [4], the authors extracted the LBP features to train the SVMs as the writer-dependent classifiers. For the traditional CNN model, it just extracts the features from the last fully-connected layer. Then, using the SVMs for each user to build the verification system. The experimental results are the averages of all users with 10 trials.

Table I shows the results on skilled forgeries. This experiment is to verify whether the query samples are genuine signatures or skilled forgeries. We can see that only using the SVMs with features extracted from a normal CNN can achieve the desired results. The proposed method obtained the best results than other state-of-the-art models on all datasets. When the number of the randomly selected sample

[1] http://www.gpds.ulpgc.es/

Table I: The skilled forgeries experiment (EER$_{skilled}$ in %). The 5 samples and 10 samples represent the randomly selecting 5 or 10 samples of each user for training the SVMs.

Dataset	HMM+GF [2] (5 samples)	SVM+LBP [4] (5 samples)	Traditional CNN (5 samples)	Proposed (5 samples)	Proposed (10 samples)
GPDS-150	11.48	16.45	8.34±0.52	**7.45±0.41**	**6.32±0.35**
GPDS-300	12.11	16.50	8.41±0.61	**7.48±0.52**	**6.25±0.41**
GPDS-1000	11.07	17.01	8.31±0.47	**7.12±0.45**	**6.43±0.38**
GPDS-2000	11.34	16.63	8.20±0.42	**7.23±0.53**	**5.92±0.41**
GPDS-5000	11.10	16.93	8.08±0.53	**7.15±0.48**	**6.11±0.47**

Table II: The random impostor experiment (EER$_{random}$ in %), where micro-deformations is not important.

Dataset	HMM+GF [2] (5 samples)	SVM+LBP [4] (5 samples)	Traditional CNN (5 samples)	Proposed (5 samples)	Proposed (10 samples)
GPDS-150	4.17	**1.31**	4.89±0.48	4.64±0.38	**3.08±0.33**
GPDS-300	4.32	**1.45**	4.77±0.52	4.72±0.54	**3.23±0.42**
GPDS-1000	4.37	**1.63**	4.82±0.58	4.91±0.43	**3.17±0.35**
GPDS-2000	4.44	**1.73**	4.94±0.42	4.95±0.55	**3.22±0.41**
GPDS-5000	4.53	**1.63**	4.56±0.43	4.58±0.41	**2.84±0.43**

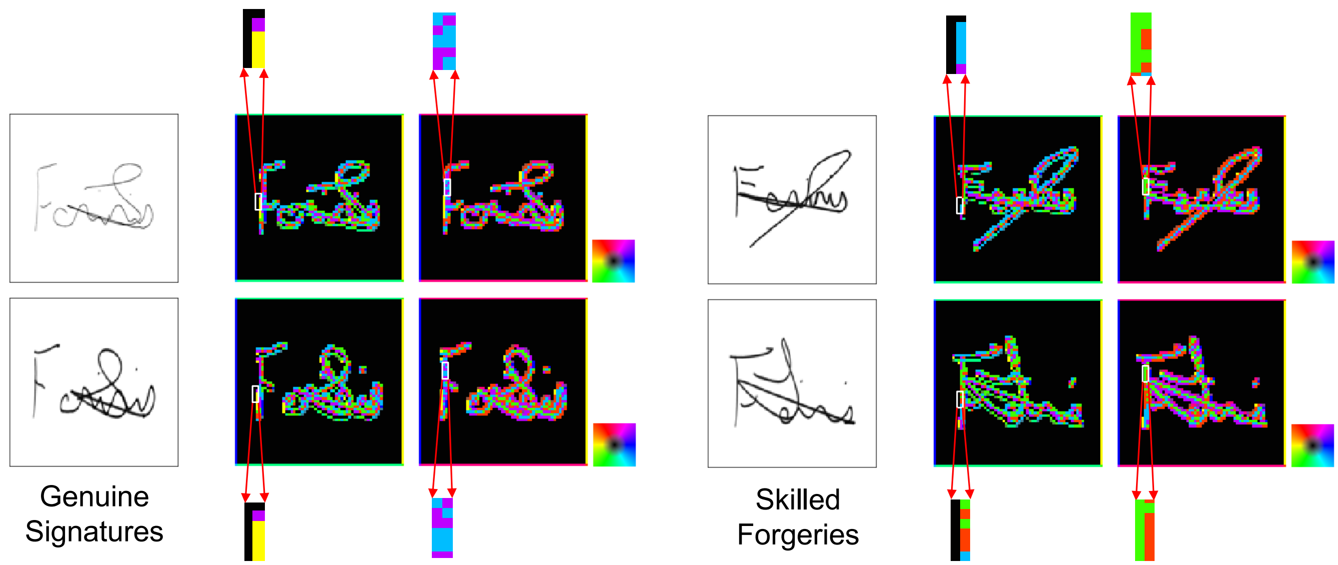

Figure 4: An example improvement of the displacement features. Here, all signatures are from the same user. The first column is the original signature images, the second and the third columns are the displacement features extracted from two filters.

increased to 10, the EERs are smaller than before, which means that using more samples to train the SVMs can improve the performance of the verification systems. In addition, as the datasets become larger (the number of users is increasing), the EER becomes higher when using the traditional models. It means that the traditional models are not general for all the users. The proposed method is more stable than the traditional models when the dataset size becomes huger.

Table II shows the results on random impostors. This experiment is mainly to discriminate the different users. Since the purpose of the feature extractor is to capture different behaviors between the genuine signatures and forgeries, the proposed method did not obtain the best results on this experiment. However, we achieved a competitive performance compared to the state-of-the-art results. Even if the model in [4] can classify the different users well, the ability to distinguish the genuine signatures and skilled forgeries of this model is far worse than the proposed method.

Fig. 4 presents an improved example by using the proposed method. We visualized the displacement features to observe the micro different behaviors between the genuine signatures and skilled forgeries. If we use traditional CNN based features, they all belong to the same class (positive class). But using the fused features by the proposed method, they can be classified correctly. From Fig. 4, we can see that the genuine signatures have similar behaviors on the displacement features and the skilled forgeries are different from the genuine samples in some places. For example, in the bottom left corner of the first filter, the genuine samples have some features in yellow and purple directions, but it is rare in the skilled forgery samples. And in the second filter, the displacement features can capture some blue and purple directions in the genuine signatures, but in the skilled forgeries, the corresponding position is green and red. It means that the proposed method can capture some micro deformations or distortions between the genuine signatures and skilled forgeries.

V. Conclusion

In this paper, we proposed a novel approach that fuses the proposed displacement features and pooling features as a new feature to capture the micro deformations between the genuine signatures and skilled forgeries. Then, we applied linear SVMs as the writer-dependent classifiers to build a verification system. Extensive experimental results demonstrate that the fused features can capture the micro deformations or distortions between the genuine signatures and skilled forgeries well, which is helpful for the offline signature verification systems. For future work, we plan to apply the proposed method to build an end-to-end offline signature verification system and introduce the user information to further improve the performance.

Acknowledgment

This work was supported by JSPS KAKENHI (Grant Number JP18K11373 and JP17H06100) and China Scholarship Council (Grant Number 201706330078).

References

[1] K. Ahrabian and B. Babaali, "Usage of Autoencoders and Siamese Networks for Online Handwritten Signature Verification," *Neural Computing and Applications*, 2017.

[2] M. A. Ferrer, J. B. Alonso, and C. M. Travieso, "Offline Geometric Parameters for Automatic Signature Verification Using Fixed-point Arithmetic," *TPAMI*, 2005.

[3] M. A. Ferrer, M. Diaz, C. Carmona-Duarte, and A. Morales, "A Behavioral Handwriting Model for Static and Dynamic Signature Synthesis," *TPAMI*, 2017.

[4] M. A. Ferrer, J. F. Vargas, A. Morales, and A. Ordonez, "Robustness of Offline Signature Verification Based on Gray Level Features," *TIFS*, 2012.

[5] L. G. Hafemann, L. S. Oliveira, and R. Sabourin, "Fixed-sized Representation Learning from Offline Handwritten Signatures of Different Sizes," *IJDAR*, 2018.

[6] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Analyzing Features Learned for Offline Signature Verification Using Deep CNNs," in *ICPR*, 2016.

[7] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Learning Features for Offline Handwritten Signature Verification Using Deep Convolutional Neural Networks," *Pattern Recognition*, 2017.

[8] L. G. Hafemann, R. Sabourin, and L. S. Oliveira, "Offline Handwritten Signature Verification-Literature Review," in *IPTA*, 2017.

[9] J. Hu and Y. Chen, "Offline Signature Verification Using Real Adaboost Classifier Combination of Pseudo-dynamic Features," in *ICDAR*, 2013.

[10] S. Lai and L. Jin, "Learning Discriminative Feature Hierarchies for Off-Line Signature Verification," in *ICFHR*, 2018.

[11] M. Liwicki, M. I. Malik, C. E. Van Den Heuvel, X. Chen, C. Berger, R. Stoel, M. Blumenstein, and B. Found, "Signature Verification Competition for Online and Offline Skilled Forgeries (Sigcomp2011)," in *ICDAR*, 2011.

[12] M. I. Malik, S. Ahmed, A. Marcelli, U. Pal, M. Blumenstein, L. Alewijns, and M. Liwicki, "ICDAR2015 Competition on Signature Verification and Writer Identification for On-and Off-line Skilled Forgeries (SigWIcomp2015)," in *ICDAR*, 2015.

[13] M. I. Malik, M. Liwicki, L. Alewijnse, W. Ohyama, M. Blumenstein, and B. Found, "ICDAR 2013 Competitions on Signature Verification and Writer Identification for On-and Offline Skilled Forgeries (SigWiComp 2013)," in *ICDAR*, 2013.

[14] M. Okawa, "Synergy of Foreground-background Images for Feature Extraction: Offline Signature Verification Using Fisher Vector With Fused KAZE Features," *Pattern Recognition*, 2018.

[15] J. Ruiz-del Solar, C. Devia, P. Loncomilla, and F. Concha, "Offline Signature Verification Using Local Interest Points and Descriptors," in *CIAPR*, 2008.

[16] A. Soleimani, B. N. Araabi, and K. Fouladi, "Deep Multitask Metric Learning for Offline Signature Verification," *Pattern Recognition Letters*, 2016.

[17] F. Wang, L. Zhao, X. Li, X. Wang, and D. Tao, "Geometry-aware Scene Text Detection with Instance Transformation Network," in *CVPR*, 2018.

[18] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual Attention Network for Image Classification," in *CVPR*, 2017.

[19] Q. Ye and D. Doermann, "Text Detection and Recognition in Imagery: A Survey," *TPAMI*, 2015.

[20] M. B. Yilmaz, B. Yanikoglu, C. Tirkaz, and A. Kholmatov, "Offline Signature Verification Using Classifier Combination of HOG and LBP Features," in *IJCB*, 2011.

[21] X. Zhang, J. Zou, K. He, and J. Sun, "Accelerating Very Deep Convolutional Networks for Classification and Detection," *TPAMI*, 2016.

[22] Z. Zhang, X. Liu, and Y. Cui, "Multi-phase Offline Signature Verification System Using Deep Convolutional Generative Adversarial Networks," in *ISCID*, 2016.

[23] Y. Zheng, Y. Cai, G. Zhong, Y. Chherawala, Y. Shi, and J. Dong, "Stretching Deep Architectures for Text Recognition," in *ICDAR*, 2015.

[24] Y. Zheng, B. K. Iwana, and S. Uchida, "Discovering Class-Wise Trends of Max-Pooling in Subspace," in *ICFHR*, 2018.

[25] Y. Zheng, B. K. Iwana, and S. Uchida, "Mining the Displacement of Max-pooling for Text Recognition," *Pattern Recognition*, 2019.

[26] E. N. Zois, L. Alewijnse, and G. Economou, "Offline Signature Verification and Quality Characterization Using Poset-oriented Grid Features," *Pattern Recognition*, 2016.

[27] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in *CVPR*, 2018.