Text-independent Writer Identification via CNN Features and Joint Bayesian

Youbao Tang, Xiangqian Wu
The School of Computer Science and Technology
Harbin Institute of Technology
Harbin, China
tangyoubao@hit.edu.cn, xqwu@hit.edu.cn

***Abstract*—This paper proposes a novel method for offline text-independent writer identification by using convolutional neural network (CNN) and joint Bayesian, which consists of two stages, i.e. feature extraction and writer identification. In the stage of feature extraction, since a large number of data is essential to train an effective CNN model with high generalizability and the amount of handwriting is limited in writer identification, a data augmentation technique is first developed to generate thousands of handwriting images for each writer. Then a deep CNN network is designed to extract discriminative features to represent the properties of different writing styles, which is trained by using the generated handwriting images. In the stage of writer identification, the training dataset is used to train the CNN model for feature extraction and the joint Bayesian technique is employed to accomplish the task of writer identification based on the extracted CNN features. The proposed method is tested on two standard benchmark datasets, i.e. ICDAR2013 and CVL dataset. Experimental results demonstrate that the proposed method gets the best performance compared to the state-of-the-art approaches.**

Keywords-text-independent writer identification; convolutional neural network; joint Bayesian; data augmentation

I. INTRODUCTION

Offline text-independent writer identification which has been studied for many years, aims to identify the writer of a text from a number of known writers by using their handwriting images. Due to its importance for forensic analysis and documents authorization, offline text-independent writer identification has attracted a number of researchers. To further enlarge its influence, several international contests of writer identification have been successfully held for recent years [1-3]. Extensive good works have been done in this field [4-20]. Generally, the existing approaches of offline text-independent writer identification can be roughly categorized into two classes according to the extracted features, i.e. hand-designed features based approaches [4-18] and deep-learned features based approaches [19, 20].

The hand-designed features based approaches usually extract some hand-designed features (such as SIFT, LBP, and so on) from the local regions of handwriting images, and then get the final feature vectors to represent the properties of handwriting images by using bag-of-word model or other feature representation techniques. For example, Gabor filter [4] and local binary patterns (LBP) [5, 6] are used to extract textural features from handwriting images for writer identification. Some edge-based directional features of handwriting [7, 8], microstructure features [9, 10], allograph based histogram features [7, 10, 11], SIFT based features [12-15], and other heuristic features [16-18] are computed as local structural features for writer identification. These hand-designed features usually capture the appearance information (such as direction, texture, and so on) of handwriting, which cannot well dig the deep and intrinsic writing style information.

Since the convolutional neural network (CNN) has powerful ability to learn deep features and has been successfully applied in computer vision community, several researchers [19, 20] have employed CNN to perform the feature extraction from handwriting images for writer identification in the last year and gotten the state-of-the-art results. Christlein et al [19] first extract a number of local CNN-based features from local image patches centered on the contour of the handwriting image. Then these local features are aggregated to form one global descriptor by using GMM for writer identification. Fiel et al [20] first preprocess the handwriting images by performing text line detection, deskew, and other operations and use sliding windows to extract image patches, and then from which the CNN is used to extract local features. For a given handwriting image, a lot of local CNN-based features will be extracted. They take the mean values of all local feature vectors as the final feature vector for writer identification. These methods extract CNN features from the local structures of handwriting. But they cannot capture the global writing style information from entire handwriting, which is one of most important factors for writer identification.

To overcome the above mentioned problems, this work proposes a novel offline text-independent writer identification method based on CNN. In order to extract global and discriminative features from entire handwriting, a large number of handwriting images for each writer are the first essential for CNN model training. Therefore, a data augmentation technique is proposed to generate the handwriting images, and then based on which a CNN model is trained for global feature extraction. Finally, the joint Bayesian is used for writer identification. Although the CNN features and joint Bayesian have been used in other

2167-6445/16 $31.00 © 2016 IEEE
DOI 10.1109/ICFHR.2016.102

applications, such as face recognition, to the best of our knowledge, it is the first time to directly extract the global CNN features from handwriting images instead of local CNN features from local image patches and utilize the joint Bayesian algorithm to compute the similarity of features for writer identification.

As discussed above, there are three differences between the proposed method and the CNN based approaches [19, 20], which are also the main contributions of this work. The first one is that the global features are directly extracted from entire handwriting images to represent the handwriting styles by using CNN in this work, while other approaches first extract local features from small local image patches with CNN and then aggregate them to form the global features. The second one is that a new data augmentation technique is developed to generate a large number of handwriting images, which overcomes the limitation of handwriting amount during CNN model training. The third one is that the joint Bayesian technique is utilized to measure the similarity between two extracted features for writer identification instead of using cosine and χ^2 distance metrics. With these differences, the proposed method can extract discriminative features and effectively recognize the writers' identities. And the experimental results on two standard benchmark datasets demonstrate that the proposed method gets the state-of-the-art results.

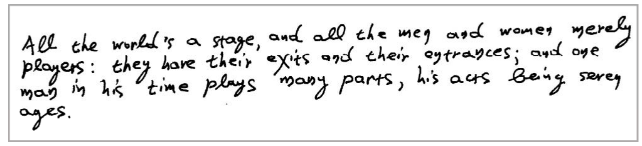
The rest of this paper is organized as follows. Section II gives the process of feature extraction based on CNN. Section III describes joint Bayesian for writer identification. The experimental results are given in Section IV. Finally, the conclusions of this work are presented in Section V.

II. Feature Extraction

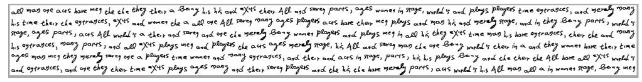
The convolutional neural network (CNN) has been successfully applied in large scale image classification [21] due to its powerful ability of feature representation. And in the field of writer identification, some works [19, 20] get the state-of-the-art results by using CNN. Therefore, we also propose a novel writer identification method by employing CNN in this work. In this section, the details of feature extraction based on CNN will described.

A. Data Augmentation

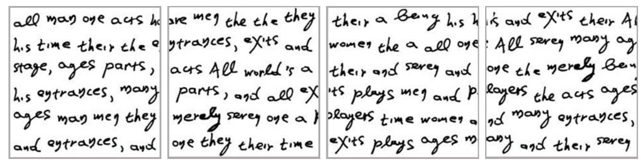
As we know, in order to train an effective CNN model, a large number of training data is essential. However, the amount of handwriting is small for each writer in current existing handwriting dataset for writer identification, such as 4 samples for each writer in ICDAR2013 dataset [3] and 5 samples for each writer in CVL dataset [27]. Obviously, it is serious insufficient for CNN model training. Moreover, it is almost impossible and unrealistic to collect a large number of samples for different writers. To solve this problem, enormous local image patches centered on the contour of the handwriting images [19] or inside the sliding windows on text lines [20] are extracted by previous approaches, which only consider the local structural information of handwriting.



(a)



(b)



(c)

Figure 1. Example of data augmentation for handwriting image. (a) the input handwriting image. (b) the constructed page of (a). (c) four examples of constructed images by using the proposed data augmentation technique.

Compared with the local structural information, the global information is more important for representing the properties of different writing styles. To achieve this goal, we use CNN to directly extract the global information from entire handwriting images instead of local image patches. Therefore, a new problem is introduced, i.e. how to generate a large number of images from the limited handwriting images for CNN model training. To solve this problem, a data augmentation technique of handwriting image is developed in this work.

For text-independent writer identification, we care about the differences between writing styles of handwriting images instead of text contents and text location. Hence, given a handwriting image, we can randomly permute the text contents to generate a lot of different images from it. As we can see, both allograph and word can be considered as the element for text permuting. In Ref. [14], Wu et al. have demonstrated that the word-level extracted features are more powerful than allograph-level ones for writer identification by experiments. Therefore, in this work, we use words segmented from handwriting images as elements to permute the texts to generate a large number of images. Given a handwriting image (as shown in Fig. 1(a)), the data augmentation process is described as follows:

1) Segment the input handwriting image into a number of words by using the word segmentation method proposed by Wu et al [14].
2) Randomly permute all words into a text line. Here, the centers of words are horizontally aligned, and there are 30 pixels width between adjacent words.
3) Repeat step 2 with N times to get N text lines. And every L text lines are concatenated to form a page of

handwriting in the vertical direction, as shown in Fig. 1(b). There are 30 pixels height between adjacent text lines. So $\frac{N}{L}$ pages are constructed.

4) Split each page into a number of non-overlapping images, shown in Fig. 1(c). Wu et al [14] demonstrate that the performance of writer identification is change little when the number of words is larger than 15, which means almost all of write style information for writer identification is included in the 15 words. So when the number of text lines in a page is 6, the number of words in the non-overlapping images is more than 15. Therefore, we set $L = 6$ and $N = 120$ in this work. With this setting, we can make sure that there are at least 15 words in each constructed image. After dealing with all pages, the average number of constructed images is about 500 for each handwriting image on the training dataset.
5) Resize all constructed image into the fixed size of 224*224.

Fig. 1 gives an example of data augmentation. Given a handwriting image (as shown in Fig. 1(a)), the constructed page is shown in Fig. 1(b) and the constructed images are given in Fig. 1(c). With the above data augmentation technique, hundreds of constructed images are generated for each handwriting image, which carry enough information to represent their writing styles. In the preliminary experiments, we also use other approaches to do the data augmentation, such as adding noises, randomly dropping some strokes, and so on. And we find that the proposed data augmentation technique gets the best performance and is much better than other approaches. Therefore, the proposed data augmentation technique is adopted in this work.

B. CNN based Feature Extraction

After doing data augmentation, a large number of images are generated from the training dataset, based on which, the next step is to extract some discriminative features to represent their properties. As done by [19, 20], the CNN is employed to compute the features in this work. Since the handwriting images are simpler and easier to be understood than the natural scene images, we design a lightweight CNN framework compared with the winners of LSVRC, such as Alexnet [21], Googlenet [22], and VGGnet [23].

The CNN framework consists of four convolutional layers, two fully connected layers, one output layer, and one softmax layer, as shown in Fig. 2. From Fig. 2, we can see that each convolutional layer is followed by one rectified linear units (ReLU) nonlinear activation layer, one local response normalization (LRN) normalization layer, and one max pooling layer. And to alleviate overfitting, each fully connected layer is followed by one ReLU nonlinear activation layer and one dropout layer with the dropout ratio of 0.5. The output layer actually is a fully connected layer and the number of its nodes is decided by the number of classes of training dataset. The last layer is a softmax layer during test phase or a softmax with loss layer during training phase. The details of network configure are given in TABLE I. We implement the CNN framework with the popular deep learning tool, i.e. Caffe [24].

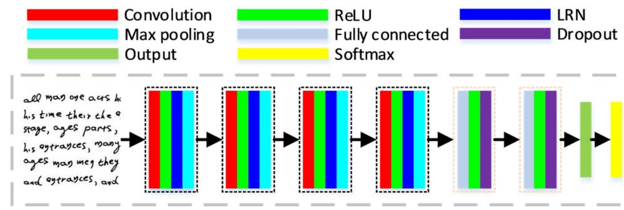


Figure 2. The framework of our CNN model used for feature extraction.

TABLE I. THE NETWORK CONFIGURE OF EACH LAYER IN OUR CNN FRAMEWORK

type	number	size	stride	pad	nodes
conv1	32	7*7	2	3	
pool1		3*3	2		
conv2	64	5*5	2		
pool2		3*3	2		
conv3	96	3*3	1	1	
pool3		2*2	2		
conv4	128	1*1	1		
pool4		2*2	2		
fc5					256
fc6					256

In this work, we use the ICDAR2013 training dataset which consists of 100 writers and each writer with 4 handwriting images to train the CNN model for feature extraction. Therefore, the output layer and softmax layer have 100 nodes. Here, we randomly select one sample from four samples of each writer to form the validation set, and the rest samples are used as training set. For each sample, we can get about 500 constructed handwriting images by using the proposed data augmentation technique. The training process is performed by using the asynchronous stochastic gradient descent with 0.9 momentum, fixed learning rate schedule (the initial learning rate is 0.01, and we decrease the learning rate to 10% until the validation loss is stable). The training process is stopped after decreasing the learning rate twice.

Given an input image, we can use the trained CNN model to extract features for writer identification with the following process. Firstly, do the data augmentation and generate about 20 constructed images. Then for each constructed image, perform the forward behavior with the trained CNN model and use the outputs of the second fully connected layer as feature vector. Therefore, we finally extract a 256-D feature vector by using CNN. It needs to be noticed that the extracted features are directly used for writer identification in this work while the precious work [19, 20] uses extra feature representation techniques to aggregate the CNN based features to form a final feature vector for writer identification.

III. WRITER IDENTIFICATION

With the CNN based features, the next step is to use them for writer identification. Since the samples don't belong to the writers of training samples during test, the outputs of softmax cannot be considered as the final results of writer identification. In this paper, the joint Bayesian [25] technique which has been successfully applied in face verification [26] is used for writer identification based on the extracted features. Here, we simply introduce the technique of joint Bayesian. For more details, please refer to [25]. In joint Bayesian, an extracted CNN based feature x is represented by the sum of two independent Gaussian variables as follows

$$x = \mu + \epsilon \tag{1}$$

where x is the observed feature with the mean of all features subtracted, $\mu \sim N(0, S_\mu)$ represents its identity, and $\epsilon \sim N(0, S_\epsilon)$ the intra-personal variations (e.g. languages, pens, and so on). Joint Bayesian models the joint probability of two handwriting images given the intra- or extra-personal variation hypothesis, $P(x_1, x_2|H_I)$ and $P(x_1, x_2|H_E)$. It is readily shown from Equation (1) that these two probabilities are also Gaussian with variations

$$\Sigma_I = \begin{bmatrix} S_\mu + S_\epsilon & S_\mu \\ S_\mu & S_\mu + S_\epsilon \end{bmatrix} \tag{2}$$

and

$$\Sigma_E = \begin{bmatrix} S_\mu + S_\epsilon & 0 \\ 0 & S_\mu + S_\epsilon \end{bmatrix} \tag{3}$$

respectively. S_μ and S_ϵ can be learned from data with EM algorithm. In test, it calculates the log likelihood ratio

$$r(x_1, x_2) = \log \frac{P(x_1, x_2|H_I)}{P(x_1, x_2|H_E)} \tag{4}$$

which has closed-form solutions and is efficient. Therefore, for two feature vectors x_1 and x_2, if they are from the same person, $r(x_1, x_2)$ will be much larger than the one when they come from different persons. So given a x_i, we compute $r(x_i, x_j)$ between x_i and all other x_j and set the identity of x_i as the same with the identity of x_j by finding the largest $r(x_i, x_j)$. Since for each handwriting image, 20 constructed images are generated for writer identification during test, we use the voting result of 20 feature vectors to make the final decision.

IV. EXPERIMENTS

A. Datasets and Evaluation Criterions

The proposed method is evaluated on two standard benchmark datasets: ICDAR 2013 writer identification contest dataset (called ICDAR2013 dataset) [3] and CVL dataset [27].

The ICDAR2013 dataset [3] consists of a training dataset and a test dataset, containing 100 writers and 250 writers, respectively. Each writer is asked to copy four pages of text in two languages (two in English and two in Greek), which can be used for cross-language writer identification. And the number of text lines that are produced by the writers ranges between two and six.

The CVL dataset [27] is built for writer identification and retrieval and has 310 writers. There are 27 writers who provide 7 texts (1 German and 6 English) and 283 writers who provide 5 texts (1 German and 4 English). The original images in this dataset are RGB color images, so we convert them into gray ones for writer identification in this work. Word segmentation results have been provided in this dataset, so we don't need to do the word segmentation during data augmentation on this dataset.

In this work, the ICDAR2013 training dataset is used to train the CNN model and the joint Bayesian model. The same experimental setting is utilized during test on ICDAR2013 test dataset and CVL dataset.

As done by other approaches [7, 14], we use the following popular criterions, i.e. the "leave-one-out" strategy, the soft Top-N, and the hard Top-N performance, to evaluate the performance of the proposed method on each dataset. The "leave-one-out" strategy means that for each handwriting document sample, the distances to all other samples of these datasets are computed. The soft TOP-N criterion is defined as that a correct hit is considered when at least one document image of the same writer is included in the N most similar document images with K-nearest-neighbors. The hard TOP-N criterion is defined as that a correct hit is considered when all N most similar document images are written by the same writer. As we can see that the hard TOP-N criterions make sense only when at least N reference documents exist for a given writer, which is much stricter than the soft one.

B. Experimental Results

The proposed method is evaluated on the ICDAR2013 test dataset and CVL dataset. In order to have a fair comparison with other approaches, we only use the ICDAR2013 training dataset to train the CNN model for feature extraction. After training the CNN model, the outputs of the second fully connected layer are considered as features. Therefore, a large number of feature vectors are extracted from training dataset and used to train a joint Bayesian model for writer identification. For the test datasets, given a handwriting image, we can get 20 recognition results by using joint Bayesian and set the class with the largest voting value as the identity of the handwriting image. It is considered as a correct recognition if

TABLE II. THE EXPERIMENTAL RESULTS OF THE PROPOSED METHOD AND OTHER APPROACHES ON ICDAR2013 ENTIRE DATASET

Approach	Soft evaluation			Hard evaluation	
	Top-1	Top-5	Top-10	Top-2	Top-3
CS-UMD-a [3]	95.1	98.6	99.1	19.6	7.1
CS-UMD-b [3]	95.0	98.6	99.2	20.2	8.4
HIT-ICG [3]	94.8	98.0	98.3	63.2	36.5
TEBESSA-a [3]	90.3	96.7	98.3	58.2	33.2
TEBESSA-c [3]	93.4	97.8	98.5	62.6	36.5
CVL-IPK [3]	90.9	97.0	98.0	44.8	24.5
Fiel et al [20]	88.5	96.0	98.3	40.5	15.8
Wu et al [14]	95.6	98.6	99.1	63.8	36.5
Christlein et al [15]	97.1	98.8	99.1	42.8	23.8
Nicolaou et al [6]	97.2	98.9	99.2	52.9	29.2
Christlein et al [19]	98.9	N/A	N/A	83.2	61.3
Ours	**99.0**	**99.2**	**99.6**	**84.4**	**68.1**

TABLE III. THE EXPERIMENTAL RESULTS OF THE PROPOSED METHOD AND OTHER APPROACHES ON ICDAR2013 SUB-LANGUAGE DATASETS

Approach	Greek sub-dataset			English sub-dataset		
	Top-1	Top-5	Top-10	Top-1	Top-5	Top-10
CS-UMD-a [3]	95.6	98.6	99.2	94.6	98.4	98.8
CS-UMD-b [3]	95.2	98.8	99.0	94.4	98.4	99.0
HIT-ICG [3]	93.8	97.2	97.8	92.2	96.4	96.8
TEBESSA-a [3]	91.0	96.8	97.8	86.0	94.4	96.0
TEBESSA-c [3]	92.6	98.0	98.4	91.2	96.2	96.6
CVL-IPK [3]	88.4	96.8	97.8	91.4	95.8	97.2
Wu et al [14]	95.0	98.2	98.8	94.6	97.4	98.4
Christlein et al [15]	97.0	99.0	99.2	95.4	98.0	98.2
Nicolaou et al [6]	96.6	99.6	99.8	95.2	98.0	98.4
Christlein et al [19]	99.6	N/A	N/A	97.6	N/A	N/A
Ours	**99.6**	**99.8**	**100**	**97.9**	**98.6**	**99.1**

the class is the same with the label of the handwriting image. The evaluation criterions described in Section IV-A are used to evaluate the performance of different methods. TABLE II and TABLE III list the experimental results of the proposed method and other approaches on ICDAR2013 entire dataset and sub-language dataset. TABLE IV presents the experimental results of the proposed method and other approaches on CVL dataset. From these tables, we can see that the proposed method obtains better performance than other approaches.

Compared with the hand-designed features based approaches [6, 14, 15], the CNN based approach [19] and ours gets much better performance. One possible reason is that the

TABLE IV. THE EXPERIMENTAL RESULTS OF THE PROPOSED METHOD AND OTHER APPROACHES ON CVL DATASET

Approach	Soft evaluation			Hard evaluation		
	Top-1	Top-5	Top-10	Top-2	Top-3	Top-4
CS-UMD [27]	97.9	99.1	99.4	90.0	71.2	48.3
QUQA A [27]	30.5	57.5	67.1	5.7	0.5	0.1
QUQA B [27]	92.9	97.9	98.3	84.9	71.5	50.6
TEBESSA-a [27]	69.8	89.5	94.4	44.5	27.4	12.3
TEBESSA-b [27]	96.0	97.8	98.1	91.4	83.0	64.6
TEBESSA-c [27]	97.6	98.3	98.5	94.3	88.2	73.9
TSINGHUA [27]	97.7	99.0	99.1	95.3	94.5	73.0
Fiel et al [20]	98.9	99.3	99.5	97.6	93.3	79.9
Wu et al [14]	99.2	99.4	99.6	98.3	96.8	91.3
Christlein et al [15]	99.2	99.5	99.6	98.1	95.8	88.7
Nicolaou et al [6]	99.0	99.4	99.5	97.7	95.2	86.0
Christlein et al [19]	99.4	N/A	N/A	98.8	97.3	92.6
Ours	**99.7**	**99.8**	**100**	**99.0**	**97.9**	**93.0**

CNN based features have strong discriminative ability to distinguish different writing styles. The CNN based approach [20] gets worse performance that the hand-designed features based approaches [6, 14, 15]. That's because the approach [20] uses a simple cosine distance metric to compute the similarity of the CNN based features. Compared with the CNN based approaches [19, 20], our method gets the best performance. We summarize two possible reasons after analyzing. The first one is that the proposed method extract more discriminative features from the pages of handwriting to represent the differences between different writing styles, while the approaches [19, 20] extract local features from small local image patches. The second one is that our extracted features are directly used for writer identification with joint Bayesian, while the approaches [19, 20] need other feature representation techniques to aggregate their extracted local features, resulting in losing some information for writer identification. Therefore, the experimental results demonstrate the effectiveness of the proposed method for writer identification.

V. CONCLUSIONS

This paper proposes a novel offline text-independent writer identification method by using the CNN based features and joint Bayesian. Compared to the traditional local features, such as SIFT, the CNN model can learn more discriminative information from original handwriting images and the extracted CNN based features are more powerful to represent the properties of different writing styles for improving the performance of writer identification. Compared to the other CNN based approaches, the proposed method has three advantages. Firstly, a data augmentation technique is developed to avoid the limitation of small amount of handwriting samples. Secondly, the CNN based features are learned from entire handwriting images instead of the small local image patches and are directly used for writer identification. Thirdly, the proposed method uses the joint Bayesian model for writer identification instead of simple cosine and χ^2 distance metrics. With the three advantages, the proposed method gets the best performance on two standard benchmark datasets, which demonstrates the effectiveness of the proposed method for writer identification.

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of China under Grant 61472102 and 61672194. The authors would like to thank the founders of the publicly available datasets and the support of NVIDIA Corporation with the donation of the Tesla K40 GPU used for this research.

REFERENCES

[1] G. Louloudis, N. Stamatopoulos, and B. Gatos, "ICDAR 2011 Writer Identification Contest," Proc. International Conference on Document Analysis and Recognition, 2011, pp. 1475-1479.

[2] G. Louloudis, B. Gatos, and N. Stamatopoulos, "ICFHR 2012 Competition on Writer Identification Challenge 1: Latin/Greek Documents," Proc. International Conference on Frontiers in Handwriting Recognition, 2012, pp. 829-834.

[3] G. Louloudis, B. Gatos, N. Stamatopoulos, and A. Papandreou, "ICDAR 2013 Competition on Writer Identification," Proc. International Conference on Document Analysis and Recognition, 2013, pp. 1397-1401.

[4] B. Helli and M.E. Moghaddam, "A text-independent Persian writer identification based on feature relation graph (FRG)," Pattern Recognition, vol. 43, 2010, pp. 2199-2209.

[5] D. Bertolini, L.S. Oliveira, E. Justino, and R. Sabourin, "Texture-based descriptors for writer identification and verification," Expert Systems with Applications, vol. 40, 2013, pp. 2069-2080.

[6] A. Nicolaou, A.D. Bagdanov, M. Liwicki, and D. Karatzas, "Sparse Radial Sampling LBP for Writer Identification," Proc. International Conference on Document Analysis and Recognition, 2015, pp. 716-720.

[7] M. Bulacu and L. Schomaker, "Text-independent writer identification and verification using textural and allographic features," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, 2007, pp. 701-717.

[8] L. Van Der Maaten and E.O. Postma, "Improving automatic writer identification," Proc. Benelux Conference on Artificial Intelligence, 2005, pp. 260-266.

[9] X. Li and X. Ding, "Writer identification of chinese handwriting using grid microstructure feature," Proc. International Conference on Biometrics, 2009, pp. 1230-1239.

[10] Y. Tang, X. Wu, and W. Bu, "Offline text-independent writer identification using stroke fragment and contour based features," Proc. International Conference on Biometrics, 2013, pp. 1-6.

[11] I. Siddiqi and N. Vincent, "Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features," Pattern Recognition, vol. 43, 2010, pp. 3853-3865.

[12] S. Fiel and R. Sablatnig, "Writer retrieval and writer identification using local features," Proc. International Workshop on Document Analysis Systems, 2012, pp. 145-149.

[13] Y. Tang, W. Bu, and X. Wu, "Text-independent writer identification using improved structural features," Proc. Chinese Conference on Biometric Recognition, 2014, pp. 404-411.

[14] X. Wu, Y. Tang, and W. Bu, "Offline text-independent writer identification based on scale invariant feature transform," IEEE Transactions on Information Forensics and Security, vol. 9, 2014, pp. 526-536.

[15] V. Christlein, D. Bernecker, F. Honig, and E. Angelopoulou, "Writer identification and verification using GMM supervectors," Proc. IEEE Winter Conference on Applications of Computer Vision, 2014, pp. 998-1005.

[16] L. Schomaker and M. Bulacu, "Automatic writer identification using connected-component contours and edge-based features of uppercase western script," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, 2004, pp. 787-798.

[17] A. Brink, J. Smit, M. Bulacu, and L. Schomaker, "Writer identification using directional ink-trace width measurements," Pattern Recognition, vol. 45, 2012, pp. 162-171.

[18] G. Ghiasi and R. Safabakhsh, "Offline text-independent writer identification using codebook and efficient code extraction methods," Image and Vision Computing, vol. 31, 2013, pp. 379-391.

[19] V. Christlein, D. Bernecker, A. Maier, and E. Angelopoulou, "Offline Writer Identification Using Convolutional Neural Network Activation Features," Proc. German Conference on Pattern Recognition, 2015, pp. 540-552.

[20] S. Fiel and R. Sablatnig, "Writer Identification and Retrieval Using a Convolutional Neural Network," Proc. Computer Analysis of Images and Patterns, 2015, pp. 26-37.

[21] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," Proc. Neural Information Processing Systems, 2012, pp. 1097-1105.

[22] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," arXiv preprint arXiv:1409.4842, 2014.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[24] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," Proc. ACM Multimedia Conference, 2014, pp. 675-678.

[25] D. Chen, X. Cao, L. Wang, F. Wen, and J. Sun, "Bayesian face revisited: A joint formulation," Proc. European Conference on Computer Vision, 2012, pp. 566-579.

[26] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1891-1898.

[27] F. Kleber, S. Fiel, M. Diem, and R. Sablatnig, "Cvl-database: An off-line database for writer retrieval, writer identification and word spotting," Proc. International Conference on Document Analysis and Recognition, 2013, pp. 560-564.