



Robust offline handwritten character recognition through exploring writer-independent features under the guidance of printed data

Yaping Zhang^{a,b}, Shan Liang^a, Shuai Nie^{a,b}, Wenju Liu^{a,*}, Shouye Peng^c

^a National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, No.95, Zhongguancun East Road, Beijing 100190, PR China

^b University of Chinese Academy of Sciences, China

^c Xueersi Online School, China

ARTICLE INFO

Article history:

Received 12 August 2017

Available online 15 February 2018

MSC:

41A05

41A10

65D05

65D17

Keywords:

Handwritten character recognition

Writer-independent features

Adversarial feature learning

Convolutional neural network

ABSTRACT

Deep convolutional neural networks have made great progress in recent handwritten character recognition (HCR) by learning discriminative features from large amounts of labeled data. However, the large variance of handwriting styles across writers is still a big challenge to the robust HCR. To alleviate this issue, an intuitional idea is to extract writer-independent semantic features from handwritten characters, while standard printed characters are writer-independent stencils for handwritten characters. They could be used as prior knowledge to guide models to exploit writer-independent semantic features for HCR. In this paper, we propose a novel adversarial feature learning (AFL) model to incorporate the prior knowledge of printed data and writer-independent semantic features to improve the performance of HCR on limited training data. Different from available handcrafted features methods, the proposed AFL model exploits writer-independent semantic features automatically, and standard printed data as prior knowledge is learnt objectively. Systematic experiments on MNIST and CASIA-HWDB show that the proposed model is competitive with the state-of-the-art methods on the offline HCR task.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

¹ Handwritten character recognition (HCR) has been widely applied in mail sorting [29], historical documents recognition [24], handwritten notes transcription and bank check reading. In addition, it is also an important component of handwritten text recognition [25]. To this end, decades of efforts have been devoted to HCR, but robust HCR is still a challenging task due to the huge variance of handwriting styles. In particular, the recognition task is more difficult when the handwritten characters are offline, while we focus on the offline HCR in this paper.

HCR is a typical classification task, while deep convolutional neural network (DCNN) is one of the most exciting classification models and makes great progress in many fields. At present, DCNN has been widely applied to HCR and achieved significant performance improvements [5,7,8,31,32,35]. However, its success heavily relies on a large amount of labeled data which is high-cost. Moreover, DCNN-based HCR system suffers from the shift between the training and test distributions [33] when the variance of handwriting

styles is large. In the previous works, there are mainly three methods proposed to alleviate this issue, which are summarized as follows: 1) using data augmentation techniques [5,7] to generate more data with different handwriting styles, such as affine transformation [19] and distorted generation [14]; 2) adopting writer-adaptation to match the feature distributions from the source domain to the target domain [30,32,33]; 3) designing writer-independent features manually to reduce within-class variation of character shape [32,35], such as normalization-cooperated gradient features [16]. These methods are well-designed on the basis of domain-specific knowledge to compensate for shape variation caused by various handwriting styles. However, generating more distorted data is insufficient to cover all the variations of handwritten characters, while writer adaptation methods need to match specific writers, and handcrafted writer-independent feature is so subjective that some pretty important information in characters may be lost. Besides, they don't explicitly model the final recognition objective, and they cannot take advantage of extra information.

As known to all, standard printed characters are writer-independent and present more semantic contents of characters. While handwritten characters contain various handwriting styles information of writers, as shown in Fig. 1, which tremendously

* Corresponding author:

E-mail address: lwj@nlpr.ia.ac.cn (W. Liu).

¹ Shuai Nie makes equal contribution to this work with Yaping Zhang.

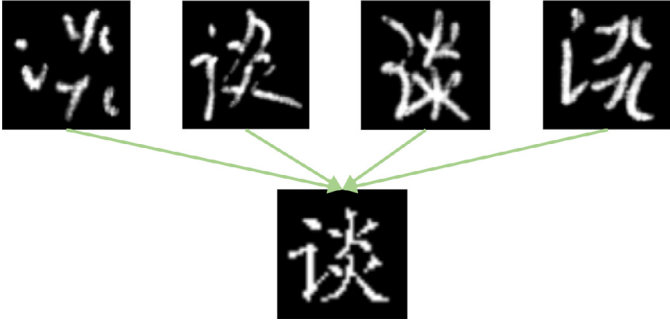


Fig. 1. All different handwritten characters could be representable in standard printed characters which are writer-independent.

interferes with the recognition of handwritten characters. Many experiments show that the printed characters can be more easily recognized without the interference of handwriting styles of writers [1]. In fact, standard printed characters are usually used as stencils to instruct us to recognize new characters, especially when the pupils learn to read from a textbook. Therefore, it is possible to improve the performance of HCR by exploiting writer-independent semantic features, while standard printed characters could be used as prior knowledge to guide models to exploit these features.

Generative adversarial network (GAN) is a well-known adversarial learning model [10]. It is composed of a generator and a discriminator. The discriminator can guide the generator to adjust a complex data distribution to another specific distribution by adversarial learning. When GAN is performed on a handwritten character set, such as MNIST dataset [13], it's interesting to observe that the generator could transfer noise vectors to realistic character images [6,26]. What's more, [28] used GAN to transfer images from street view house number (SVHN) dataset to the domain of MNIST dataset via GAN. Therefore, as shown in Fig. 2(a), it could be expected that the handwritten characters with various handwriting styles could be transferred into standard printed characters by adversarial learning.

Inspired by GAN, we propose a novel adversarial feature learning (AFL) model to exploit writer-independent semantic features for HCR. As shown in Fig. 2(b), AFL is a variant of GAN and composed of a feature extractor, a discriminator and a classifier, which concentrates the strengths of discriminative model and generative model for HCR. The feature extractor is used to extract encoding features of handwritten and printed characters. The discriminator judges whether the extracted features come from handwritten or standard printed characters. With the prior knowledge provided by standard printed characters, it can guide the feature extractor to exploit writer-independent semantic features from handwritten characters automatically. Finally, the extracted features are fed into the classifier to recognize the handwritten characters. The feature extractor, the discriminator and the classifier are jointly optimized by adversarial training. In the process of adversarial training, the prior knowledge from standard printed characters and writer-independent semantic features are incorporated, and hence we could get better performance of HCR.

We summarize our contributions as follows: 1) we introduce the writer-independent standard printed data as prior knowledge, which is learnt by AFL objectively, rather than use handcrafted writer-independent features which need a great amount of domain knowledge; 2) the proposed model could exploit writer-independent semantic features automatically, which in turn alleviates the large variance of handwriting styles for HCR; 3) the proposed AFL model could make better classification, which concentrates the strengths of discriminative model and generative model; 4) we achieve superior performance than the state-of-the-

art model results on offline ICDAR-2013 handwritten Chinese character recognition competition dataset.

The remaining parts of this paper are organized as follows. Section 2 firstly reviews the related works. Then, we describe the proposed AFL method in Section 3. Experimental results and its detailed analysis are presented in Section 4. Finally, we draw concluding remarks in Section 5.

2. Related work

GAN [10] is to learn a generative model synthesizing images similar to real images through a two-player game between a generator and discriminator (in Fig. 2(a)). The key idea of GAN is an adversarial loss that forces the generator could find the mapping between noise vectors and real images. Despite many promising developments [2,11], more recent works focus on image synthesis, such as image generation [4,22] and representation learning [23]. There are a few trials to use GAN to make classification. [27] and [21] generalized GAN to learn a discriminative classifier, which is trained to not only classify images but also identify real and fake images. However, it's incompatible that a single discriminator network plays two competing roles of identifying fake samples and predicting labels, which results in hard achieving equilibrium between discriminator and generator, and in turn prevents the model from predicting labels accurately. Recently, [15] proposed triple-GAN to make two competing roles in discriminator separated to introduce three components for both classification and class-conditional generation. But the structure of triple-GAN is complex for classification. In the proposed AFL model, we concentrate on extracting writer-independent features for classification directly. The domain adversarial training proposed by [9] is closest to our AFL model, while it promotes invariant with respect to the shift between two domains for unsupervised domain adaptation. Our task is quite different from theirs. We introduce the standard printed data to guide the model to extract writer-independent semantic features for character recognition directly.

3. Adversarial feature learning

The proposed AFL model tries to improve the performance of HCR by learning writer-independent semantic features of a handwritten characters, where we provide standard printed characters as prior knowledge. This is different from the conventional feature learning of DCNN, where the primary goal is to extract discriminative feature matching training set. AFL is composed of three neural network components: a feature extractor (F) that characterizes the features of handwritten and standard printed characters, a classifier (C) that could make correct classification for the extracted features, and a discriminator (D) that acts as a teacher could guide F to learn prior knowledge. F , C , and D are jointly optimized by the adversarial training algorithm. The desired equilibrium is that F will concentrate on writer-independent features and can keep discriminative with the constraints of D and C .

3.1. Formulation

Let's denote a handwritten character image as x_h . The corresponding feature vector is denoted as $F(x_h, \theta_f)$ that is extracted by the feature extractor F . The extracted feature $F(x_h, \theta_f)$ is input into the classifier C to compute its classification probability:

$$P(y_c = y_h | x_h; \theta_f, \theta_c) = \frac{e^{-C(y_c=y_h | F(x_h, \theta_f); \theta_c)}}{\sum_{j=0}^L e^{-C(y_c=j | F(x_h, \theta_f); \theta_c)}} \quad (1)$$

where $y_h \in Y_h$, $Y_h = \{1, 2, \dots, L\}$ is a finite label set, L is the number of character classes, y_c denotes the predicted label, θ_f and θ_c

Download English Version:

<https://daneshyari.com/en/article/6940441>

Download Persian Version:

<https://daneshyari.com/article/6940441>

[Daneshyari.com](https://daneshyari.com)