

Chinese Characters Style Analysis Using Generative Adversarial Network

Haiyan Deng^{1, a)}, Yijun Liu^{2, b)}

¹*School of Computers. Guangdong University of Technology. Guangdong, Guangzhou, China*

²*School of Information Engineering. Guangdong University of Technology. Guangdong, Guangzhou, China*

^{a)} peptreltongxu@163.com

^{b)} yjliu@gdut.edu.cn

Abstract. Various methods have been proposed in previous works to achieve effective printed Chinese character recognition. Feature extraction and production of large scale multi-font Chinese character remains a major challenge owing to the wide variety in the shape, layout, and grey-level distribution of single Chinese characters across different font styles. Convolutional neural networks (CNNs) have shown outstanding performances in many fields. Convolutional layer is the dominant algorithm used in training neural networks. In this paper, we propose a Generative Adversarial Network (GAN) to analyze Chinese character fonts, extract font features through CNN, and we can output the type of fonts from the learning characters. There are many competitive CNN applications, aiming to achieve Chinese font performance. In order to capture rich and discriminative information of fonts, we combine GAN with CNN to learn good features for the fonts and put out the desired fonts as required. The approach can generate more obvious font features and better display than original structure of GAN

Key words: Chinese font; Character font; Convolutional neural network CNN; GAN.

INTRODUCTION

Chinese characters play an important role in many applications, such as typography culture, advertising decoration and handwriting. Including the recognition of fonts, classification, generation and other applications in many different directions. Font are closely related with font generation and data storage and management. Especially in recognition and new font generation. Many researchers have proposed different methods of extraction and analysis of font features. No matter words or characters, whether English [1], Roman [2], Chinese characters, and others. Moreover, the number of Chinese characters are pretty large and translating into the corresponding fonts are time-consuming and labor-intensive. Chinese characters are increasing applied into Neural networks for recognition. Such as HCL2000 [3], HIT.MWI [4], SCUT-COUCH2009 [5] and CASIA.HWDB [6]. In recent years, ICDAR handwritten Chinese character recognition contest [7,8], researchers advanced different deep learning algorithms to obtain better recognition results. The generation of MNIST data set provides a solution to the problem of the conversion of Chinese character fonts from generation adversarial network. Based on this, this paper proposes a conditional-based generative adversarial network model. Experiment shows it is very effective and the method combines the generation of adversarial network and convolutional neural network to achieve the transformation of the font and generate corresponding Chinese fonts.

RELATED WORK

CNN(Convolutional Neural Network)

Convolutional neural networks are neural networks that use convolution operations instead of general matrix multiplication operations at least one level of the network. Convolutional neural networks provide an end-to-end learning model that can be

trained by traditional gradient descent methods. The training convolutional neural network can extract features from the training data set, and can realize the extraction and classification of training features. In previous work, CNN is employed in many in image classification [9][10], object detection [11][12], many other recognition tasks [13][14] and even non-recognition tasks, especially in text recognition [15,16]. Convolution neural network is characterized by weight-sharing structure and similar to biological neural network structure can reduce the parameters by the weight. CNN for image-related tasks present the implementation details in pre-processing on images, initializations of Networks, selections of activation functions, methods of ensemble multiple deep networks, diverse regularizations etc aspects, thereby reducing the complexity of the network model. we can use cnn these advantages to analyze the font characteristics. Chinese character research involves segmentation, thinning algorithms, feature extraction and neural networks [16]. The experimental results of answer recommendation on the chinese font dataset show that the CNN exhibits better performance.

Erik Bernhardsson [17] proposed Analyzing 50k fonts using deep neural networks. Zhuoyao Zhong, et [18] employed a CNN with new techniques for multi-font printed Chinese character recognition and proposed the addition of a multi-pooling layer to conventional CNN, it was shown to be robust and suitable for PCCR in their experiments. Shumeet Baluja [19] presented a learning method to analyze typographic style. Many of the overall shapes, weights, angles and serifs were successfully modeled. Xiangyu Qiu, et [20] proposed a novel glyph description method to generate the rest with the similar writing style. Xiangyu Qiu, Wenhua Jia, Hongmei Li. Institution of Computer Science & Technology. and so on. including the skeleton lines and bounding boxes of the strokes [21], stroke information [22,23]. Qingsheng Li, et [24] advanced a novel dynamic description and generation method for Chinese character. By using stroke-segments-vector, stroke elements, feature points and their weights, characteristic expression and weight vector of the feature points. Wenfa Qi, et [25] proposed an automatic watermarked font library generation method by modifying vector contour curves of Chinese characters.

GAN (generation adversarial network)

GAN is a new deep learning framework for training Generative Model based on the Adversarial training process, and GAN was proposed by Ian Goodfellow of Google Brain Scientists in 2014. GAN is composed of two models: a generative model G, which is used to obtain the distribution of the input sample x (indicating the learning representation viewpoint that the deep learning is a process of fitting the (x,y) relationship to the input sample, which is actually the learning input. The distribution of sample x), a discriminative model D, which estimates the probability that a sample is a real sample rather than a sample generated by G. Because the degree of freedom to generate adversarial network is very large. In other words, GAN's training procedure will easily lose its direction and become barbaric. In the early days of GAN, the theoretical G and D only needed to be able to fit the functions of the corresponding generation and discrimination, and the training of the generator lacked guidance and often caused people to collapse. We needed to add some extra information such as category information to wait. Constraints to solve the deep neural network itself difficult training and instability, such as the CGAN network.

THE ARCHITECTURE

Generator and Discriminator

The generator in this paper adopt the structure U-Net [26] where the first layer and the last layer are span connection, between the second layer and the penultimate layer, and so on. The first four layers are mutually Pooling operation, the last three layers are convolutions, and the last four layers are U-Net neural network structures used in the upper convolution generator. For discriminate, we Adopt optimized DCGAN structure. Discriminator structure follows discriminator structure in DCGAN.

Loss function

The construction of the objective function defines an a priori input z based on the noisy input variable p(z), then maps the input z to the generator output y' with G(z, Q), where G is everywhere It is represented by a multi-layer neural network and uses its parameter Q to fit the distribution of the input sample x. Then use a multi-layer network to represent D and output a scalar that represents the probability of originating from x instead of p(z). For G, it is equivalent to minimizing log(1-D(G(z))) expectation; for D, it is equivalent to the expectation of maximum logD(x), similar to a minimax game. From this we can get the objective function V(G,D) of GAN:

$$\min_G \max_D V(D, G) = IE_{x \sim p_{data}(x)} [\log D(x)] + IE_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

After add additional information (y) directly to the target function of generator G and discriminator D directly after adding the constraint condition, and form the conditional probability with inputs Z and X:

$$L_{cGAN}(C, D) = IE_{x,y \sim p_{data}(x,y)}[\log D(x, y)] + IE_{x \sim p_{data}(x), z \sim p_z(z)}[\log 1 - D(x, G(x, z))] \quad (2)$$

$$L_{L1}(C) = IE_{x,y \sim p_{data}(x,y), z \sim p_z(z)}[\|y - G(x, z)\|_1] \quad (3)$$

The image generation with L1 distance is more sharp than with L2 distance. Ultimately optimize the objective function:

$$G^* = \arg \min_G \max_D L_{cGAN}(C, D) + \lambda L_{L1}(G) \quad (4)$$

EXPERIMENT RESULT

There are more than 20,000 training images, 1,000 validation images, and 1,000 testing images, including 10 fonts. The style of font includes the BSKaiTi, FangSong, HeiTi, XianHei, SimSun, and so on, the only said that each font corresponding to generate more than 20,000 characters, including a commonly used words 3500. Samples of the dataset are showed as follow(a): This article follows the method in DCGAN with a small batch(128) of stochastic gradient descent (SGD) to train a large number of fonts. The standard renormalization of ownership relies on the zero standard deviation of the normal distribution set to 0.02 and incorporates the Adam optimizer. The slope is set to 0.2 in LeakyReLU and learning rate is 0.0001. In addition, the momentum item is 0.5 in the beginning.



FIG. 1 Dataset



FIG.2 epochs

Fig (1): Some fonts of the dataset, the BSKaiTi, FangSong, HeiTi, XianHei, SimSun in order about the dataset; Fig (2): The first and second lines are the input of the Founder script and the target output. The other output is the number of intermediate iterations of 200, 1000, 5000, and 10000. The more times the number is learned, the clearer it is, until the level is unchanged.



FIG.3 Results



FIG.4 mutual transformation

Fig (3): Different Networks generate corresponding results. The first line is original structure of GAN, the second line is generated by our new network, which is based on DCGAN. Fig (4): The mutual transformation between different fonts has a better effect.

CONCLUSION

In this article, we propose to use the DCGAN-based structure to achieve style transitions between different fonts. Our experiments show that this structure also has a better generation effect than the traditional DCGAN neural structure. It also significantly improves neural network performance. The build structure used in this article is expected to extend to other architectures to achieve similarly generated tasks. Finally, by applying the generated adversarial network to our Chinese fonts, we have achieved excellent results in different styles of Chinese characters. Moreover, on the basis of the transformation of font style realized in this article, can we generate other art forms other than existing fonts? This study provides a powerful and easy-to-

implement benchmark and learning method for the new architecture to compare. We use imagery to generate fonts. Another potential direction is to use non-image-based methods, such as the use of long-term short-term memory (LSTM) to generate characters directly in TrueType language. If the results can be generated directly in the TTF encoding, this will produce another orthogonal result to complement this study. This method is feasible in future research.

ACKNOWLEDGMENTS

The paper is supported by funds from Department of Science and Technology at Guangdong Province and Guangzhou City (No.201604010051, 2015B090901060, 2016B090903001, 2016B090904001, 2016B090918126, 2016KZ010101).

REFERENCES

1. Roger D. Hersch, Claude Bétrisey, and Justin BurEcole Polytechnique Fédérale de Lausanne, School of Design. Perceptually Tuned Generation of Grayscale Fonts. *IEEE Computer Graphics and Applications*. 1995.78-89.
2. JM Coueignoux. Generation of Roman printed fonts., Thesis. 1975. Ph.D.--Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science.; Vita.; Bibliography: leaves 128-130.
3. Zhang H, Guo J, Chen G, et al. HCL2000-A large-scale handwritten Chinese character database for handwritten character recognition[C]//Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on. IEEE, 2009: 286-290.
4. Su T, Zhang T, Guan D. HIT-MW dataset for offline Chinese handwritten text recognition[C]//Tenth International Workshop on Frontiers in Handwriting Recognition. SuvisioR, 2006.
5. Jin L, Gao Y, Liu G, et al. SCUkoUCH2009—a comprehensive online unconstrained Chinese handwriting database and benchmark evaluation[J]. *International Journal on Document Analysis and Recognition (DDAR)*, 2011, 14(1): 53-64.
6. Liu C L, Y'm F, Wang D H, et al. CASIA online and offline Chinese handwriting databases[C]//Document Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE, 2011: 37-41.
7. Liu C L, Y'm F, Wang Q F, et al. ICDAR 2011 Chinese Handwriting Recognition Competition[C]//Proceedings of the 2011 International Conference on Document Analysis and Recognition. IEEE Computer Society, 2011: 1464-1469.
9. Y'm F, Wang Q F, Zhang X Y, et al. ICDAR 2013 Chinese handwriting recognition competition[C]//Document Analysis and Recognition (ICDAR), 2013 12th International Conference on. IEEE, 2013: 1464-1470.
10. A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1106-1114.
11. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional neural networks," *arXiv:1311.2901*, 2013.
12. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580-587.
13. W. Y. Zou, X. Wang, M. Sun, and Y. Lin, "Generic object detection with dense neural patterns and regionlets," in *ArXiv:1404.4316*, 2014.
14. A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. DeepVision Workshop*, 2014, pp. 806-813.
15. Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1701-1708.
16. Xinhao Liu, Takahito Kawanishi, Xiaomeng Wu, Kunio Kashino. Scene Text Recognition with High Performance CNN Classifier and Efficient Word Inference. *ICASSP 2016*. 1322-1326.
17. Shuye Zhang, Lianwen Jin, Liang Lin. Discovering similar Chinese characters in online handwriting with deep convolutional neural networks. *IJDAR 2016 May 12*. 237-252.
18. <https://erikbern.com/2016/01/21/analyzing-50k-fonts-using-deep-neural-networks/>.
19. Zhuoyao Zhong, Lianwen Jin, Ziyong Feng. Multi-font Printed Chinese Character Recognition using Multi-pooling Convolutional Neural Network. 2015 13th International Conference on Document Analysis and Recognition (ICDAR). [19] S Baluja. Learning Typographic Style. Shumeet Baluja shumeet@google.com Google, Inc. *arXiv:1603.04000v1 [cs.CV]* 13 Mar 2016.

20. X Qiu, W Jia, H Li. A font style learning and transferring method based on strokes and structure of Chinese characters. *Computer Science & Service System (CSSS), 2012 International Conference on*.2012:1836-1839.
21. Hobby, John and Guoan Gu, "A Chinese Meta-font," *Dept Of Computer Science, Stanford University, TUGboat, Volume 5 (1984). No. 2. 1983:119-136.*
22. Yu JinHui, Peng QunSheng, "Realistic synthesis of cao shu of Chinese calligraphy," *Computer Graphics, 2005, vol. 29, pp.145-153.*
23. Xu Songhua, Lau, C.M. Francis, Cheung, K. William and Pan, Yunhe. "Automatic generation of artistic Chinese calligraphy," *IEEE intelligent systems, 2005, pp.32-39.*
24. Qingsheng Li, Quan Liu, A novel dynamic description and generation method for Chinese character. 978-1-5090-0154-5/15 \$31.00 © 2015 IEEE DOI 10.1109/CIT/IUCC/DASC/PICOM.2015.105.719-724.
25. Qi, Yuxin Liu and Wei Guo, An automatic Chinese font library generation method by modifying vector contour curves. 978-1-4799-5390-5/14 \$31.00 © 2014 IEEE DOI 10.1109/IIH-MSP.2014.44.150-154.
26. O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *MICCAI*, pages :234-241. Springer, 2015. 2, 3, 4.