The Restoration of Style Chinese Characters Based on Deep Learning

Da Lv [a), Yijun Liu

School of Computer, Guangdong University of Technology Guangzhou 510006, China.

[a) Corresponding author: 502472551@qq.com

Abstract. For the image inpainting using incomplete Chinese characters proposed a new method for Chinese characters repair, the first use of U-Net type network structure combining the training method of generating against network design a style Chinese characters converter, and then repair the missing content by Chinese characters style converter. The experimental results show that the structure of the Chinese characters for the repair of the clear, less noise, edge connection also smooth, look real. This paper provides a new idea for image inpainting, it can be converted from other images to the same style of image for image recovery.

Key words: Deep learning; generative adversarial network; structure generated clear.

INTRODUCTION

With the development of new technology, deep learning has launched a new wave in academia. Machine learning, including computer vision, has been rapidly occupied by deep learning, and deep learning is also showing great splendor in every application. Recently, deep learning has also made an important breakthrough. The generation of network Generative Adversarial Networks [1] (GAN) and later people put forward a variety of generation countermeasures network [2-3]. So, we propose an image inpainting method base on deep learning, which allows the computer to learn some characters such as font weight and style, so that the computer has the ability to transform other fonts into another style font. Then a priori information and Chinese characters style converter generates incomplete existing font Chinese characters corresponding to the content information to complete the close, to predict the unknown information defect, then the defect information for prediction of the defect image in Chinese characters to repair style Chinese characters effect. The experiment shows that the reconstructed image is smooth and true to the human vision.

STYLE CHINESE CHARACTER RESTORATION

In this paper, we use the generative adversarial network to repair the missing information of Chinese characters, and accurately predict the missing part's style information, which is credible. First, we train a Chinese character style converter to convert one Chinese character to another style. Then the Chinese character image is repaired by the style converter.

Style Converter

The Chinese character style converter is a style character that converts one style of Chinese characters into another. The generative adversarial network design a Chinese characters style converter as shown in Figure 1, the font A into B font style and font, the font in the A input through a series of convolutional neural network we call generator output font C, through a neural network called D discriminator to generate the font C and to distinguish between target B

font style is the same, in order to deceive the D generator G discriminator will learn to adjust B parameters to the font style, the generated font C as much as possible and B font style similar to D cannot distinguish between two kinds of the discriminator font, D will continue to enhance the ability of adjusting parameters the difference between C and B to the font, it can effectively remove the blur noise, can also carry out some strokes to generate fine adjustment. The final font C is difficult to distinguish between the style and content of the font and the content of the font B.

2.1.1 loss function

The font is the font style conversion process by neural network style transfer [4-5], our goal is to train a G generator, making a style font into another style, in order to generate the corresponding font style, we designed the transform loss containing two parts, the content loss and the style loss.

$$context\_loss = \frac{1}{2n}\sum_{i=1}^{n}(A_i - C_i)^2 \tag{1}$$

$$style\_loss = \sum_{i=1}^{n}\log\left(1 - D\big(G(A_i)\big)\right) \tag{2}$$

$$transform\_loss = \alpha*mse\_loss + \beta*style\_loss \tag{3}$$

Network Structure

The generator network structure is adopted in this paper, pix2pix[6] and U-Net[7] from the structure consists of two parts encode and decode, are using the 5×5 convolution kernel, using convolution extraction of the input image until the 2×2 ×1024 space, and then the recovery of 64 × 64 ×1 using a series of deconvolution, convolution and skip connection, and on each layer with Batch Normalization[8] structure is shown in Figure 1 (a). The generator has reached the purpose of network transformation to a certain extent, but the background is often fuzzy noise multiple strokes is not complete and other issues, so we use generative against network thought using a discriminator to generate the results of optimization, make the image clearer, more realistic and less noise. A network structure is: input a 64 x 64 x 1 pictures, are using the 5×5 size into several layers of laminated volume convolution kernel, after each layer convolution image size decrease and increase in the number of half channel is two times the previous layer, the output layer is a two classification, as shown in Figure 1 (b) shows.
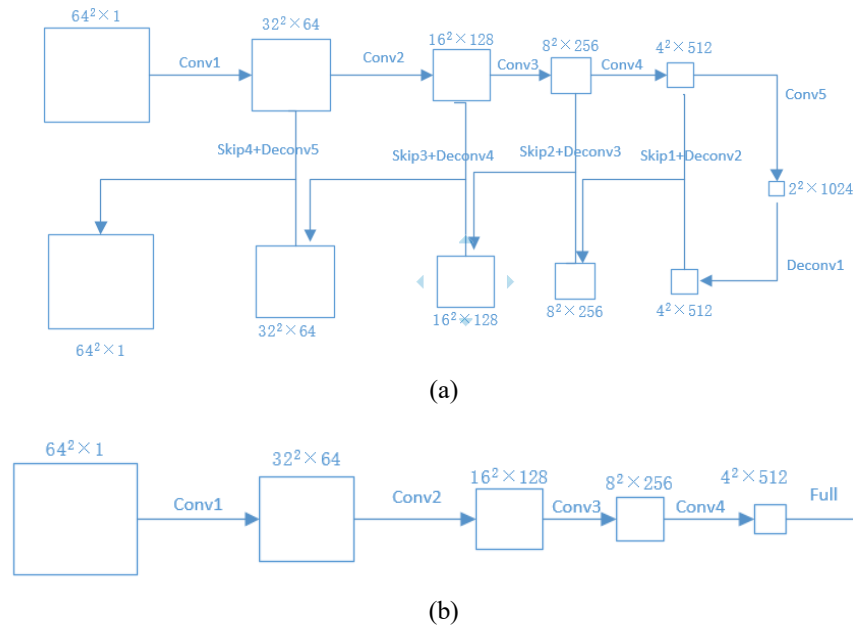

(a)

(b)

FIGURE 1. (a) Generator network structure (b) Discriminator network structure

Repair the Incomplete Chinese Characters

We need to pay attention to two types of information: context information and perceptual information. Contextual information is the pixel information existing in the image. Perceptual information is consistent with human visual perception and can be filled in the incomplete part of credible information.

Context information and perceptual information are very important. Without context information, we don't know that we should fill in the content and style of data. Without perceptual information, we cannot generate information that is consistent with human perception. Here we use the context loss and the perceptual loss

Context Loss

In order to keep the image and the input converter generates the context information of the same, we need ensure consistent image pixel position corresponding to known style conversion and the original incomplete Chinese characters image, we need to reduce the distance between the known information of the defective Chinese character and the corresponding pixel distance on the generated pseudo - style Chinese character.

$$\text{context\_loss} = ||M \odot C - M \odot y||1 \tag{4}$$

The M represents the binary mask that is not incomplete, and the C represents the Chinese character generated by the style converter, and the Y is used to represent the incomplete Chinese character.

Perceptual Loss

In order to make the content of the filling conforms to our visual perception, the perceptual loss function uses the conversion loss function transform loss of the Chinese character style converter, that is:

$$\text{Perceptual\_loss} = \sum_{i=1}^{n} \log\left(1 - D\big(G(A_i)\big)\right) \tag{5}$$

Total Loss

So, the total loss:

$$\text{total\_loss} = \lambda 1 * \text{context\_loss} + \lambda 2 * \text{perceptual\_loss} \tag{6}$$

Where $\lambda 1$ and $\lambda 2$ are super parameters.

If we want to complete the incomplete information, we need to repair Chinese characters, smooth information and forecast the Chinese characters in one piece, the repair incomplete images are Chinese characters.

$$\text{image}_{\text{reconstructed}} = M \odot y + (1 - M) \odot C \tag{7}$$

The M represents the binary mask that is not incomplete, and the C represents the Chinese character generated by the style converter, and the Y is used to represent the incomplete Chinese character.

EXPERIMENTS

In order to validate the experiment, using Simsun, Fangsong,Xinhua clerical scripts, Zhun yuan, official script and so on ,for font conversion experiment. we use the style converter convert fang song to clerical script, and then repair the clerical script incomplete Chinese characters.

First, A powerful Chinese character style converter is needed to repair Chinese characters, we use the previous method to train a style converter. This paper uses fang song into clerical script. The experimental results are as shown in Figure 2. As is shown in Figure 2, we use this algorithm convert fang song to clerical script. Watch the upper row of pictures from source font and generated fonts to real target, we find that the generated fonts's structure is clear, at first glance, not observed differences, the lower the generated font magnifier can send images generated with the real image of the target structure is basically the same, the details generated is clear and consistent, less noise, heavy style also expressed words very clearly, more real, some of the tiny details such as point ,pie and na, also can generate delicate strokes.
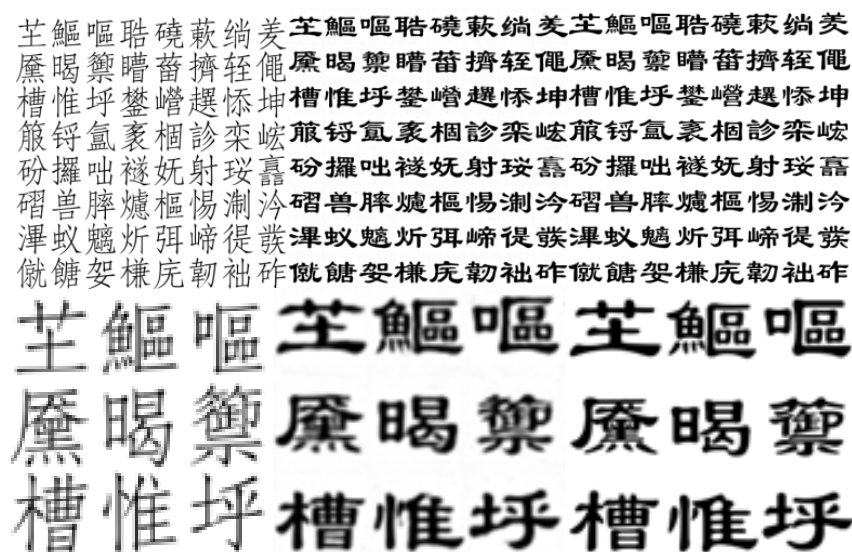
FIGURE 2. style conversion results. left up: source font; middle and upper: generated fonts; right up: real target font; Lower left: magnified source font; middle and down: magnified generation font; right down: magnified real target font.

Second, we use the style converter convert fang song to clerical script and repair the clerical script incomplete Chinese characters. It can be observed from Figure 3 that the repaired Chinese characters image look real, and the font structure is relatively complete, and the repaired content is consistent with the Chinese character without missing information style, it is also smooth at the junction.
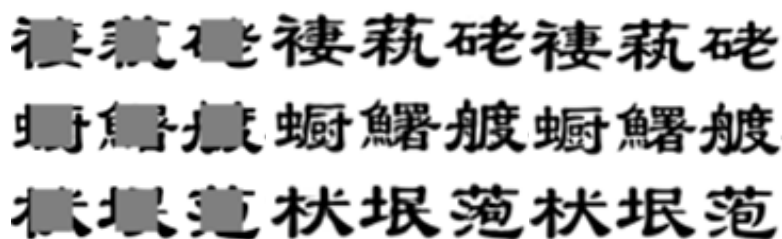

FIGURE 3. Repair results. Left: incomplete Chinese characters; Middle: the repaired font; Right: the real target font.

CONCLUSION

There is a possibility that deep learning tends to be perfect, and it is no doubt that it has a wider future than the artificial algorithm. In this paper, we use Chinese character to put forward a way of inpainting. We use the style conversion of font to transform a style font into the style font of incomplete Chinese characters, and then repair the incomplete Chinese characters. The whole process is generated using U-Net structure and generative adversarial network to design a powerful style converter, then use the Chinese character style converter to repair Chinese characters, the experimental results show that this method can automatically transform and generate contents that need to be repaired and achieve perfect cohesion. The repaired Chinese characters are very consistent with human vision sense.

ACKNOWLEDGEMENTS

REFERENCES

1. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in neural information processing systems. 2014: 2672-2680.
2. Radford A, Metz L, Chintala S. Unsuper-vised representation learning with dee-p convolutional generative adversarial networks[J]. arXiv preprint arXiv:1511.06434, 2015.
3. Denton E L, Chintala S, Fergus R. Deep generative image models using a laplac-ian pyramid of adversarial networks[C]//Advances in neural information proces-sing systems. 2015: 1486-1494.
4. 4Gatys L A, Ecker A S, Bethge M. A neur-al algorithm of artistic style[J]. arXiv preprint arXiv:1508.06576, 2015.
5. Johnson J, Alahi A, Fei-Fei L. Percept-ual losses for real-time style transferand super-resolution[C]//European Conf-erence on Computer Vision. Springer, C-ham, 2016: 694-711.
6. Isola P, Zhu J Y, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks[J]. arXiv preprint arXiv:1611.07004, 2016.
7. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
8. Ioffe S. Batch renormalization: Towards reducing minibatch dependence in batch-normalized models[C]//Advances in Neu-ral Information Processing Systems. 2017: 1942-1950.
9. Kingma D P, Ba J. Adam: A method for s-tochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.