ICIC Express Letters
Part B: Applications
Volume **12**, Number **7**, July **2021**
 ISSN 2185-2766
pp. 645–650

IMAGE CORRECTION METHOD TO BEAUTIFY JAPANESE HANDWRITTEN CHARACTERS USING GENERATIVE ADVERSARIAL NETWORKS

Takumi Otsuki[1], Takuya Takahashi[1], Shugo Nakamura[1]
Hideaki Orii[2] and Hideaki Kawano[3]

[1]Department of Electrical and Electronic Engineering
[3]Department of Electrical Engineering and Electronics
Faculty of Engineering
Kyushu Institute of Technology
1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka 804-8550, Japan
o108027t@mail.kyutech.jp; kawano@esc.kyutech.ac.jp

[2]Department of Electrical Engineering
Faculty of Engineering
Fukuoka University
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-0180, Japan

Received November 2020; accepted February 2021

Abstract. *In this paper, we describe a technique to generate beautified Japanese handwritten characters. Although handwritten characters contain more non-text information such as the writer's personality and feelings than digitally printed documents, writing 'good' Japanese characters by hand is not easy. In this paper, we propose an image processing method that corrects 'messy' handwritings that contain imperfections in local shapes of character to well-balanced character form with important features of 'Tome', 'Hane' and 'Harai'. Specifically, using the Deep Convolutional Generative Adversarial Network (DCGAN), simultaneously train a generator that transforms 'messy' handwritings into 'good' handwritings and a discriminator that identifies 'messy' and 'good' handwritings. As a result, handwriting images produced by our method contained the correct features of 'Tome', 'Hane' and 'Harai'.*
Keywords: Image correction, Japanese handwriting, Handwriting beautification, Deep learning, Neural networks, GAN, DCGAN

1. **Introduction.** It is easier for handwritten characters to convey the individuality of the writer than characters printed by a printer or the like. Even when digitization progresses, it is used when you want to communicate important information with your own thought to the other party [1]. There are three kinds of characters in Japanese. One is called "Hiragana", another is called "Katakana" and the last one is called "Kanji". The total number of characters used daily in Japanese is as many as 2,000 [2]. All three types of characters should be used in a single sentence in Japanese at the same time. When describing it by handwriting, it is necessary to write those characters one by one in a well-balanced manner. For this reason, knowledge, experience, and technique are required to write 'good' handwritings. There are some important features for well-balanced characters [3]. For example, 'Tome' means 'a stop', 'Hane' means 'an upward turn at the bottom of a stroke', and 'Harai' means 'a shake off'. As shown in Figure 1, several these important features are contained in only one character. Although a method for balancing handwritten characters by averaging handwritten strokes is already studied, no consideration has been given to 'Tome', 'Hane' and 'Harai' [4]. In this paper, we propose an

DOI: 10.24507/icicelb.12.07.645

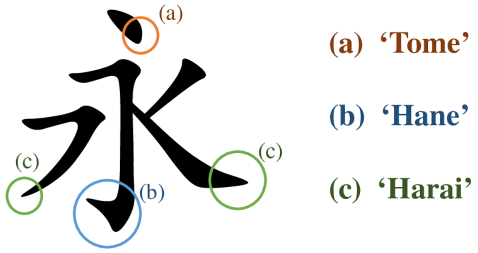

Figure 1. Example of 'Tome', 'Hane' and 'Harai'

image processing method to correct 'messy' handwritten characters into 'good' handwritten characters with 'Tome', 'Hane' and 'Harai'. It is preferable to use a generalizable model for correcting handwritten characters in languages with many types of characters such as Japanese. Recently, Deep Convolutional Generative Adversarial Network (DCGAN) is attracting attention as an image generation method that can handle many kinds of image data [5]. In this paper, we propose a method to beautify Japanese handwritten characters by adding features of 'Tome', 'Hane' and 'Harai' using DCGAN. By using convolutional layers, high-resolution images are supported. The structure of this paper is as follows. Chapter 2 describes related research, Chapter 3 describes the proposed method, and Chapter 4 describes the content of experiments using the proposed method. Chapter 5 summarizes the experimental results. The conclusions are given in Chapter 6.

2. Related Works.

2.1. Handwriting characters. Even in the modern age of digitalization, there is a lot of demand for handwriting beautification. Zitnick [6] proposed a handwriting correction method by averaging of multiple instances of the same written word or shape. Matayoshi et al. [4] extended this handwritings correction method so that it can handle Japanese, and realized real-time processing. However, these methods do not take account of 'Tome', 'Hane' and 'Harai'.

Chang et al. [7] formulated the Chinese handwritten character generation problem as learning a mapping from an existing printed font with 'Tome', 'Hane' and 'Harai' to a personalized handwritten style by using GAN. This study shows the effectiveness of GAN in correcting characters including important features of 'Tome', 'Hane' and 'Harai'.

2.2. Deep Convolutional Generative Adversarial Network. Deep Convolutional Generative Adversarial Network (DCGAN) is a kind of deep neural network used to generate synthetic data [8]. The architecture consists of two deep neural networks, a generator and a discriminator, which learn while 'adversarial' to each other. The generator section generates new data instances, while the discriminator evaluates the data for authenticity and decides whether each instance of data is "real" from the training dataset, or 'fake' from the generator. Together, the generator and discriminator are trained to work against each other until the generator is able to create realistic synthetic data that the discriminator cannot determine fake or real. After successful training, the data produced by the generator can be used to create new synthetic data, for potential use as input to other deep neural networks. This adversarial network can be described as a minimum-maximum optimization problem as follows:

$$\min_G \max_D V(G, D) = \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{x \sim p_z(x)}[\log D(G(x))] \tag{1}$$

where $p_{data}(x)$ is a probability distribution of real data x and $p_z(x)$ is a probability distribution of the latent variable z input to the generator network.

The DCGAN is a notable method for unsupervised learning tasks [5]. The CNNs are considered the state of the art for a range image recognition task. The DCGAN not only incorporates the CNN, there are various ideas. In CNN, downsampling is usually performed by max pooling. However, in the discriminator model, it is replaced by the convolution layer of stride 2. In the discriminator model, the global average pooling layer is used instead of the fully connected layers. This slows down the convergence and has the effect of preventing from overlearning. Radford et al. generated the image of bedrooms using DCGAN [5]. It produces an image that is visually indistinguishable from the visually generated data.

3. **Method.** In this section we describe our system overview. Figure 2 shows our image correction method consisting of the generator model and the discriminator model. First, the generator model corrects 'messy' handwritings. Next, the discriminator cannot distinguish the beautified corrected 'messy' handwritings from 'real good' characters provided as train data. Figure 3 shows the structure of the generator model. This generator model is an autoencoder-based architecture designed to learn features of 'good' handwritings. It is composed by convolutional layers and fully connected layers. As shown in Figure 4, the discriminator is constructed as follows: Five convolutional layers are combined with one global average pooling layer. Finally, fully connected layers classify the handwritings as 'good' or 'messy'. Once a messy handwritten character image is fed into the generator, it is corrected to a 'good' handwriting image with 'Tome', 'Hane' and 'Harai'. The

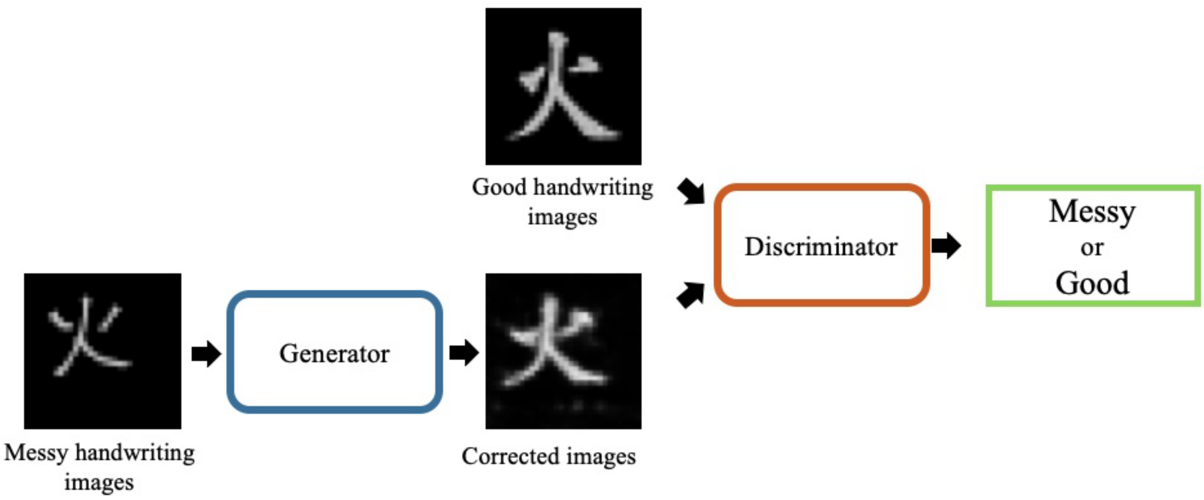

Figure 2. Visual representation of handwritten character correction method using convolutional generative method

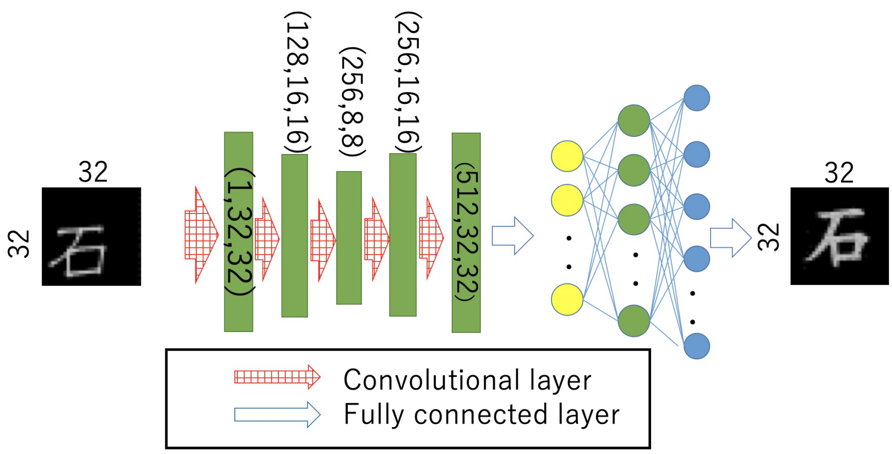

Figure 3. Detailed structure of generator model

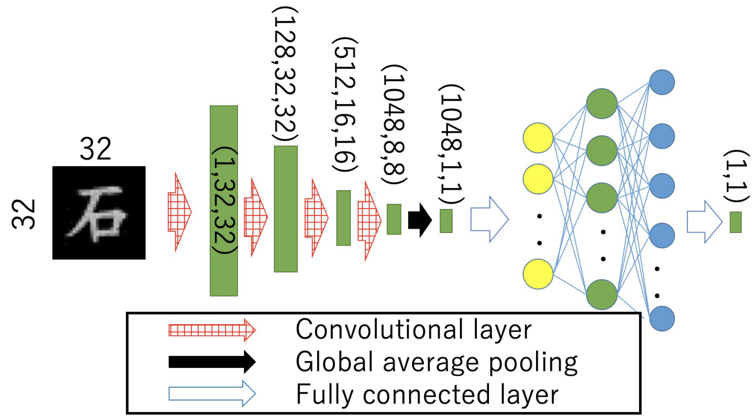

FIGURE 4. Detailed structure of discriminator model

corrected image and a 'real good' handwriting image are sent to the discriminator. The discriminator is being trained in parallel with the generator and able to identify 'good' or 'messy' characters.

4. **Experiments.** To obtain handwritings dataset, we chose 80 types of Kanji characters. 'Good' handwritten character images were prepared by scanning directly from the dictionary. And, for 'messy' handwritten character images, we prepared eight types of actually handwritten characters for each character ('Messy' dataset: 640 images, 'Good' dataset: 80 images). Additionally, for further increase of training material, the images were rotated, rescaled, and misaligned. With this approach, we prepared 20,000 images ('Messy' dataset: 10,000 images, 'Good' dataset: 10,000 images). Figure 5 shows examples of experimental images.

FIGURE 5. Example of training dataset. 'Good' handwriting images are shown on upper row and the corresponding 'messy' handwriting images are in the bottom row.

5. **Results.** Figure 6, Figure 7 and Figure 8 show the results of the trained generator model. The first rows are the input 'messy' handwriting images, and the bottom rows are the corrected handwriting images generated by the trained generator, in each figure. Figure 6 shows the results focusing on 'Tome', 'Hane' and 'Harai'. Although, there is no features like 'Tome', 'Hane' and 'Harai' in the input 'messy' handwriting characters, the corrected handwritings have them.

Figure 7 shows the results of different kinds of characters, and Figure 8 shows the results of the same characters. As shown in these figures, the length of the line and the position of the characters have not changed, but the features of 'Tome', 'Hane', and 'Harai', which are important for 'good' handwritten characters, are firmly added. In particular, in Figure

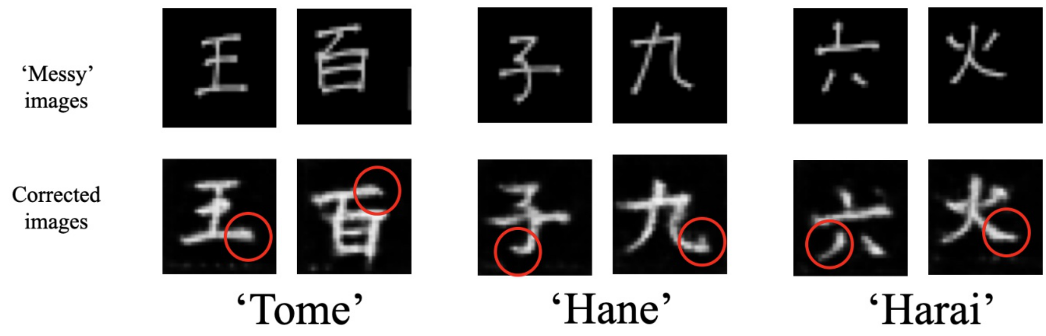

Figure 6. Results focusing on important features of Kanji, 'Tome', 'Hane' and 'Harai'. As shown by the red circle, the features appear in the corrected images on the bottom row.

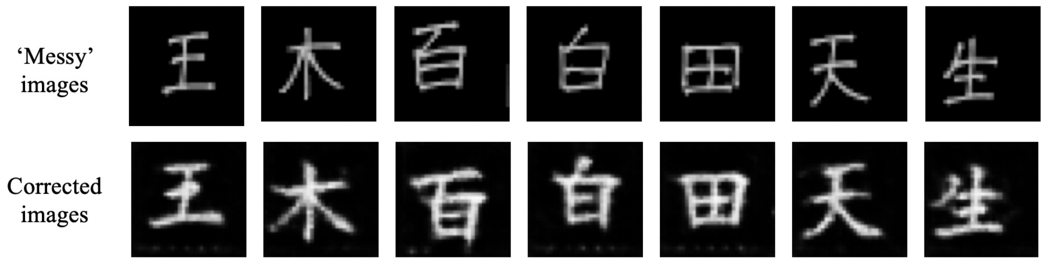

Figure 7. We show several results for different characters of handwritings correction. Upper row shows input messy handwritings. Bottom row shows corrected characters.

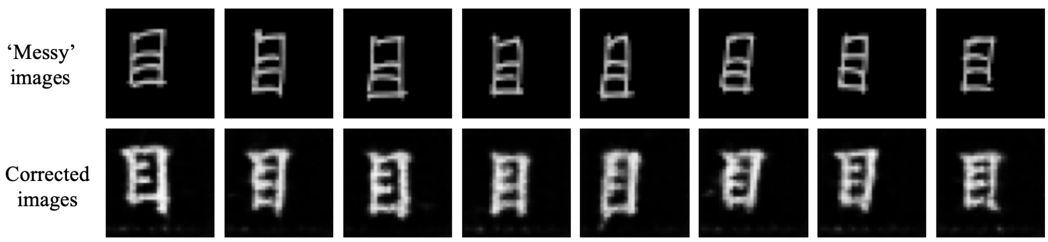

Figure 8. Comparison images of the same character

8, it can be seen that the correction is slightly different even for the same character. From these results, it can be said that the generator does not just replace the characters, but recognizes the parts of the characters and makes appropriate corrections.

6. **Conclusion.** In this research, we proposed an image processing method to correct 'messy' characters into well-written characters using GAN. By using an autoencoder-based generator, it was able to output the same characters that were input and more beautiful handwritings. It can be seen that the important features of Kanji, 'Tome', 'Hane' and 'Harai' are added to the input 'messy' handwritings. In addition to replacing the characters, it turned out that the way of correction was applied to each character. As future work, we would like to develop applications that can perform real-time handwriting correction on tablets and smartphones.

REFERENCES

[1] M. Sasaki, J. Saito and S. Nakamura, Improving visibility and reducing resistance of writers to fusion of handwritten and type characters, in *Collaboration Technologies and Social Computing*, H. Egi, T.

Yuizono, N. Baloian, T. Yoshino, S. Ichimura and A. Rodrigues (eds.), Cham, Springer International Publishing, 2018.

[2] O. Hideki and A. Masao, Kanji use in seventy contemporary magazines and Joyo Kanji: A preliminary study for application of text corpora to Japanese language policy, *Japanese Linguistics*, vol.22, pp.125-146, 2007.

[3] K. Takumi, I. Soichiro, W. Cai and N. Masaki, A learning system to learn to write Kanji characters correctly and neatly, *IPSJ SIG Notes*, vol.2014-CE-124, no.14, pp.1-7, 2014.

[4] Y. Matayoshi, R. Oshima and S. Nakamura, Mojirage: Average handwritten note, *AVI'18: Proc. of the 2018 International Conference on Advanced Visual Interfaces*, New York, NY, USA, 2018.

[5] A. Radford, L. Metz and S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, *ICLR*, 2016.

[6] C. L. Zitnick, Handwriting beautification using token means, *ACM Trans. Graph.*, vol.32, no.4, 2013.

[7] B. Chang, Q. Zhang, S. Pan and L. Meng, Generating handwritten Chinese characters using Cycle-GAN, *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2018.

[8] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, Generative adversarial nets, in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence and K. Q. Weinberger (eds.), Curran Associates, Inc., 2014.