Proc. of the 4th International Conference on Electrical, Communication and Computer Engineering (ICECCE)
30-31 December 2023, Dubai, UAE

2023 International Conference on Electrical, Communication and Computer Engineering (ICECCE) | 979-8-3503-6969-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICECCE61019.2023.10442795

A Study on the Comparison and Measurement of Improved CartoonGAN Performance for High-Quality Image Generation

Ga Eun Park*,
School of Ict Convergence Enginnering,
Kangnam University,
South Korea,
maybday@naver.com

Hae Won Lee,
School of Ict Convergence Enginnering,
Kangnam University,
South Korea,
heewon9715@gmail.com

Ju Yeon Oh,
School of Ict Convergence Enginnering,
Kangnam University,
South Korea,
ojuyeon88@gmail.com

Seong Geon Bae,
School of Ict Convergence Enginnering,
Kangnam University,
South Korea,
sgbae@kangnam.ac.kr

***Abstract*— With the rapid growth of computer vision, among them, generative artificial intelligence is being used in various industries. CartoonGAN, one of the generating artificial intelligence, can create new images by changing the original image into a cartoon style, which can be used in various fields such as webtoons and animations. To this end, a process of improving the performance of CartoonGAN is necessary, and research for this should be conducted. In this study, to solve the performance degradation that occurs when CartoonGAN is applied, we would like to suggest a direction to improve by applying technologies such as SRCNN and Tiling.**

Keywords—CartoonGAN, SRCNN, Tiling, Model, Computer Vision

I. INTRODUCTION

Computer vision is a technology that interprets and understands visual information through image and video processing. In line with this, the computer vision field is rapidly growing[1, 2]. Computer vision includes various fields such as pattern recognition, object detection, and image generation. Among them, Generative AI refers to a technology that generates or converts images, voices, and texts. Recently, generative AI has developed rapidly and is being used in art, education, writing, image generation, and various industries. Generative AI provides a wider range of experiences by providing experiences that are difficult for users to express or have not existed before. CartoonGAN, one of the generating AI technologies, is a technology that creates new images by converting images into one artist's style [3]. CartoonGAN is used in webtoons or animations because it allows you to learn models and create pictures in a new style. CartoonGAN has greatly reduced labor and costs spent on image generation and provided a variety of content. However, the existing CartoonGAN had a problem of generating inaccurate results when there was an error in the creation process. In particular, problems were evident in image conversion with low resolution and image conversion with noise. This paper applies SRCNN and Image Tilling to improve the image quality and performance of CartoonGAN, and compares them to propose a CartoonGAN with more improved performance. Through this, we aim to create high-definition images and minimize errors.

II. PREVIOUS STUDIES

Previous studies have proposed an algorithm to improve the resolution of input images by combining CartoonGAN and SRCNN. CartoonGAN is a GAN that can change the image of reality into a cartoon style. GAN is a generative model that learns two models simultaneously: creation and discrimination. GAN generates high-quality data by performing adversarial learning by confronting the constructor and the discriminator. Currently, GAN is applied in various fields and presents various possibilities for application depending on the field and purpose [4].

One of the earliest models of applying deep learning to the ultra-resolution field, SRCNN led a deep learning approach for single-image Super-Resolution (SISR) tasks [5]. The SRCNN model consists of three convolutional layers, which upscales the low-resolution version of the image to high resolution and restores the details. Through the initial up-sampling phase, it was shown to be effective in converting low-resolution input images to larger resolutions and then improving the resolution of images using the learned Convolutional Neural Network (CNN) model.

To improve CartoonGAN's Cartoon-style image resolution and quality, this study intends to apply several deep learning technologies such as SRCNN, Image Tiling, and Resizing and compare the performance of each technology to present a new perspective on image quality improvement.

III. METHODOLOGY

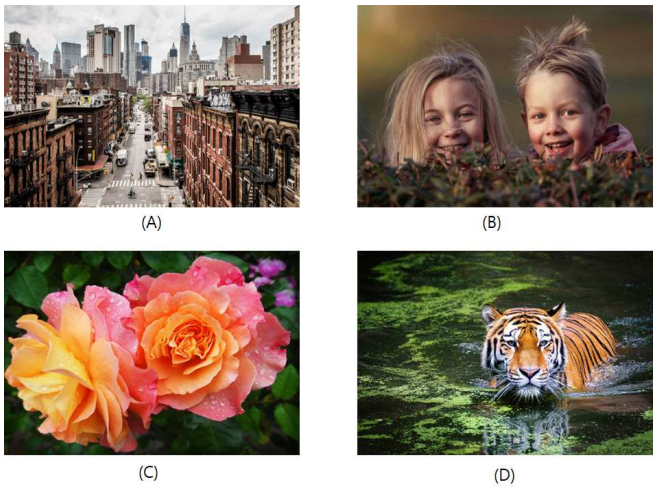

Fig. 1 Selected images for each category

This study applies a model to improve resolution and image quality to CartoonGAN to finally improve image quality, so the same image is provided to each model. Four image sample categories were selected as shown in Figure 1. (A) The images are plant, (B) the images are people, (C) the images are landscape, (D) the images are animals, and the sizes of the input image samples are uniformed to 250 ppi. Images created by CartoonGAN are often limited in improving resolution. In order to overcome the limitations of CartoonGAN and improve the quality and details of images, technologies such as Super-Resolution Convolutional Neural Network (SRCNN) and Image Tiling are additionally utilized to convert low-performance images into high-performance images. To measure the performance of the converted image, the CORREL method is used to represent the common part of the image from -1 to 1. The measurement map is represented by a number from -1 to 1, indicating that the closer the value is to 1, the more similar the image is.

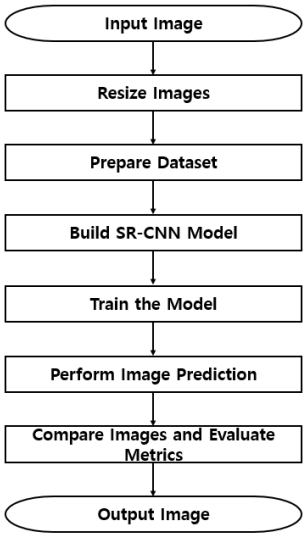

Fig. 2 Algorithm block diagram with CartoonGAN and SRCNN

If the input image is entered into CartoonGAN, the input image is downsampled during cartonization to reduce spatial resolution, and then the data is expanded and interpolated again to make it high-resolution. In this process, the resolution of images with CartoonGAN applied appears to be mostly lower than that of the original. After obtaining the generated cartoon style image, it is inputted to the SRCNN model. After that, SRCNN is used to generate a high-resolution image and performance measurement is performed through image similarity comparison based on the original image to measure improved high-resolution performance compared to the generated low-resolution image.

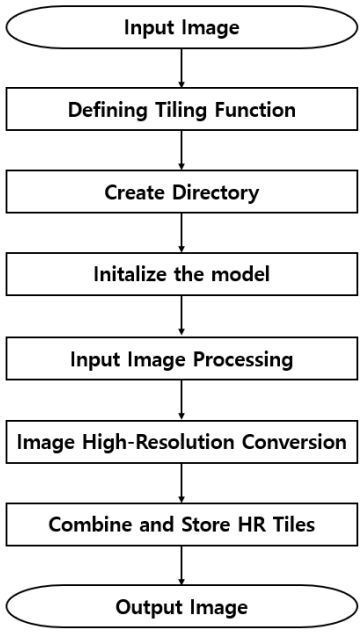

Fig. 3 Algorithm block diagram with CartoonGAN and Tilling

Image Tilling is applied to CartoonGAN as a way to improve image quality. Image Tiling is a technology that processes the original image by dividing it into small parts. This technology allows processing for each tile individually, preserving the details of the entire image. Since the loss of detailed information may be minimized, the overall image quality of the image may be improved. In order to measure the improved high-definition performance compared to the low-quality image generated through the Tiling process, performance measurement is conducted through image similarity comparison based on the original image.

IV. Experiments and Results

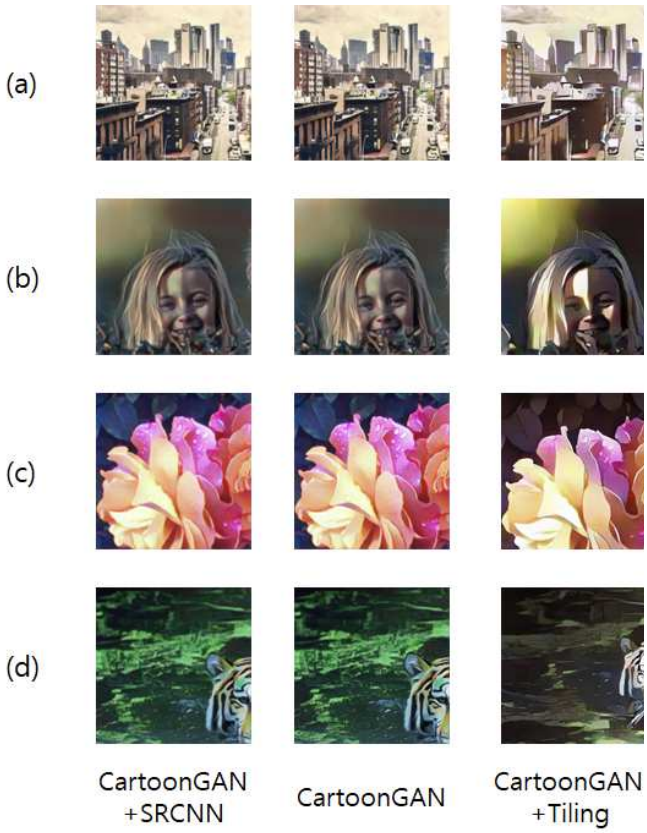

Fig. 4 Image results applying CartoonGAN, CartoonGAN and SRCNN, and CartoonGAN and Tilling

TABLE I CARTOONGAN, CARTOONGAN AND SRCNN, CARTOONGAN AND TILING IMAGE PERFORMANCE EVALUATION RESULTS

Image	CartoonGAN +SRCNN	CartoonGAN	CartoonGAN +Tiling
A	0.21	0.11	0.29
B	0.12	0.1	0.14
C	0.12	0.1	0.16
D	-0.02	-0.1	0.02

As in the categories previously selected in [Figure 4], column (a) is a plant, column (b) is a person, column (c) is a landscape, and column (d) is a result of enlarging images that apply animal images to each model. Table 1 summarizes the results of performance evaluation by comparing the original image with the image to which each model is applied. A Image shows that the original image with CartoonGAN is as similar as 0.11; the image with CartoonGAN and SRCNN is as similar as 0.21, and the image with CartoonGAN and Tilling is as similar as 0.29. B Image shows that the image with CartoonGAN applied to the original is similar by 0.1, the image with CartoonGAN and SRCNN is similar by 0.12, and the image with CartoonGAN and Tilling is similar by 0.14. C Image shows that images with CartoonGAN applied to the original are as similar as 0.1, images with CartoonGAN and SRCNN are as similar as 0.12, and images with CartoonGAN and Tilling are as similar as 0.16. D Image shows that images with CartoonGAN applied to the original are similar by -0.1, images with CartoonGAN and SRCNN are similar by -0.02, and images with CartoonGAN and Tilling are similar by 0.02. These results show that the performance of each model is improved in the order of CartoonGAN and Tilling, CartoonGAN and SRCNN, and CartoonGAN.

V. Conclusion

Computer vision is a technology that interprets and understands visual information through image and video processing and is growing rapidly.[1,2]. Computer vision includes various fields, among which generated artificial intelligence has developed rapidly and is being used in various industries. CartoonGAN, one of the generating artificial intelligence, converts the original image into a cartoon style to create a new image. It is used in various fields such as webtoons and animation because it can be created in a new style by learning models and greatly reduces the labor and cost spent on creating cartoon images.

In this study, in order to solve the problem of deteriorating performance when applying CartoonGAN, various technologies were applied to compare them after performance improvement. As for the technologies used, SRCNN and Tiling were applied to improve performance, and performance evaluation was conducted by comparing them with the original image. It was found that when SRCNN was applied together than when CartoonGAN was applied, a high-resolution image was provided. In addition, it can be seen that the image quality of the image when CartoonGAN is applied improves when Tiling is applied. To compare the performance of the two performance-enhancing models, it was found that images with Tilling performed better than images with SRCNN compared to the original images.

Through these results, CartoonGAN's performance can be improved to significantly reduce the labor and cost consumed to create cartoon images than before. In addition, more detailed cartooning can be carried out, so it can be applied to various fields other than webtoons and animations. In future research, we aim to modify existing algorithms to improve the performance of CartoonGAN and to learn and apply new models rather than existing ones.

References

[1] Hong, Y.-r., Choi, Y.-j., Oh, J.-y., Bae, S.-g. (2023). A study on sound effects through the development of in-door and outdoor detection algorithms in YOLOv3. In The 13th International Conference on Convergence Technology, pp. 119-120.

[2] Park, G. E., Lee, H. W., Cho, S. B., & Bae, S. G. (2023). A study on the visual effect of classification according to indoor/outdoor changes using

YOLOv5 in metaverse space. In The 13th International Conference on Convergence Technology in 2023, 13(1), ISSN 2287-3252.

[3] Chen, Y., Lai, Y.-K., Liu, Y.-J. (2018). Cartoongan: Generative adversarial networks for photo cartoonization. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 9565-9474.

[4] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. (2014). Generative Adversarial Networks. In Advances in neural information processing systems, Vol.27, pp. 2672-2680.

[5] Oh, S., Oh, S., Shin, H., & Kim, J. (2021). A Study on Super Resolution of SRCNN according to Early Up-sampling Method. In The Journal of Contents Computing, 3(2), 375-380.