Korean Society For Computer Game, Vol. 31, No.2, June 2018, pp.17-24
ISSN 1976-6513 http://dx.doi.org

Generating Pixel Art from Game Characters with Convolutional-Neural Network

Cheolseong Park[1], Heekyung Yang[1], Hyukmin Kwon[2], Kyungha Min[3*]

[1]Dept. of Computer Science, Graduate School, Sangmyung Univ., 20, Hongjimun 2-gil, Jongno-gu, Seoul, Korea
[2]Dept. of Eenrgy Grd, Graduate School, Sangmyung Univ., 20, Hongjimun 2-gil, Jongno-gu, Seoul, Korea
[3]Dept. of Computer Science, Sangmyung Univ., 20, Hongjimun 2-gil, Jongno-gu, Seoul, Korea
(received June 2, 2018; revised June 10, 2018; accepted June 22 2018; published June 30, 2018)

ABSTRACT

Pixel art, which presents low-resolutional images with restricted color palette, has been employed frequently in the early computer games played on low memory capacity and computational performance. Recently, pixel art wides its applications to the area such as puzzle and game. In this paper, we present a pixel art generator from images of game characters. Unlike traditional framework, we employ and modify a Convolutional-Neural Network(CNN) to generate pixel art by placing an up-convolution layer after convolution layers. The up-convolution layer increases the resolution of the result images to satisfy user-required resolution. The network is trained by minimizing the Mean Squared Error(MSE) between ground truth images and generated pixel art images from the input high-resolutional image. Also, we employ artists to produce the ground truth of pixel art for our network and augment the data by rotating and fliping. We partition the ground truth images into three datasets: a training, validation and test dataset. With this dataset, we perform a supervised learning and train our network as the pixel art generator. We show a training process and a training accuracy. Moreover, we test our architecture for a various images as well as the test dataset to prove the excellence of our architecture.

Key words: pixel art, deep learning, image abstraction, non-photorealistic rendering

1. Introduction

The display device outputs discrete colors for each pixel to represent discrete information. A cell phone, a computer monitor, and a television used in daily life can be a good example of a display device.

In the case of computers, in the early days, it was limited to express natural expressions due to limitations of computation performance and storage space.

Especially, in the game, I could see a lot of them, and objects were expressed in the form of a step by taking a dot at each pixel. By now, with the development of hardware and display devices, it has become more natural and clearer to be able to

* Correspondence to: Kyungha Min, Tel.: +82-2-2287-5377 E-mail: minkh@smu.ac.kr

express images and movies, and 4K resolution with about 8 million pixels is being prepared for commercialization.

Although it seemed that the dot-like representation of each pixel in the early games seemed to disappear due to technological advances, it is now called pixel art, and it does not disappear and attracts people's attention.

PixelArt expresses images with a lower resolution and a limited color called a color palette, just as it was in the early games. Pixel art stimulates people's inspiration and sensitivity. It can be said that the role of the arts is also sufficient. In addition, it is often used in games, such as puzzles and handheld games.

Even in the latest games, you can see the application of pixel art. In this paper, we propose a method to transform the input image into pixel art. In particular, this paper presents a method of generating pixel art through a learning process using a CNN (Convolutional-Neural Network) which has excellent results in many fields(Fig. 1).

2. Related research

Research on pixel art is a field of image abstraction in non-photorealistic rendering (NPR), and many researchers have studied it. Inglis[1,2] proposed a study to convert the line art to pixel art, and adjusted the asymmetry or protruded pixel of the result at the same time as pixel art, and showed the result of natural pixel art.

Hu [3] has further refined it, suggesting a way to maintain topological properties such as line connectivity or area in the output. In the pixel art research, choosing a color palette is also one of the research topics. Huang [4] presented the process of selecting a color palette as a real pixel artist.

Gerstner [5,6] proposed a study of pixel art of color images with color palette selection. Kuo [7] presented a method of presenting a design so that pixel art images can be assembled into Lego. In addition to this, they show another application of the field of pixel art by suggesting a research that transforms multiple images into pixel art animation instead of a single image [8].

Pixel artification can be regarded as a process of abstracting an input image. On the other hand, there is also research to restore a pixel art image to a clear image.

As a typical example, Kpof [9] proposed a method for restoring pixel art to a vector image without any resolution.

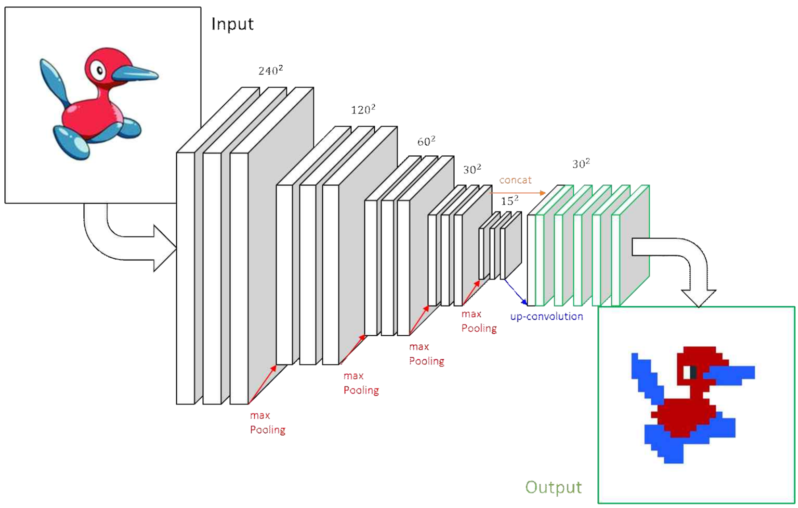

Fig. 1. Feature maps of our network

In this paper, we propose a pixel artification method based on the Composite Neural Network (CNN). CNN published by LeCun [10] showed good results as an image classification model. CNN uses the ground truth data to determine the error of the model result and automatically determines the computational weight of the composite neural network using the back propagation algorithm.

This process is called learning, and after learning, the weight is adjusted so that the developer can produce the desired result. CNN is used not only for image classification, but also for object recognition, area segmentation, etc. in addition to image processing. Especially, in the domain segmentation study, it is possible to show good results on a specific topic by only unsupervised learning. In addition, supervisory learning can also identify objects in the area.

Therefore, most of the domain segmentation researches not only the domain segmentation but also the task of determining which domain the domain is in at the same time.

In the segmentation and discrimination study, each pixel in the corresponding region has a vector value corresponding to each object. This process is very similar to the process in which each pixel has a vector value corresponding to a color in the present study. One example of this segmentation study is Pinheiro [11] and proposed a method of segmentation using CNN.

Long [12] also proposed a neural network for segmentation and object discrimination by modifying the existing CNN structure. Uhrig [13] proposed a method to extract additional information such as image depth using Long [12]. Ronneberger [14] published a neural network called U-Net, which also showed good results in the area segmentation study.

In this study, we modify the structure of neural network proposed by Ronneberger [14], and propose new neural networks and supervised learning for them. Since this study is a pixel art study rather than a domain segmentation study, the structure has been modified, and the details of the neural network structure are described in the next chapter.

3. Neural network structure

The neural network used in this study has a modified U-Net structure proposed by Ronneberger[14]. Fig. 1. shows the feature maps extracted from the neural network. The input image is a 240 × 240 color image, and the output image is a 30 × 30 pixel artifact.

We can see that there are three feature maps of the same size in the neural network, which can be regarded as an operation process of one module. Therefore, one module will have two convolution operations.

Unlike Ronneberger[14]'s neural network, we set the padding to 1 in order to prevent the image size from being reduced in the convolution process. There are 4 max pooling and 1 upsampling process between module and module. When performing upsampling, we combine the feature maps of the same size before to compensate for the missing information in the polling process. After that, it adds 4 convolution additionally. The final feature map is shown in detail in Fig. 2.

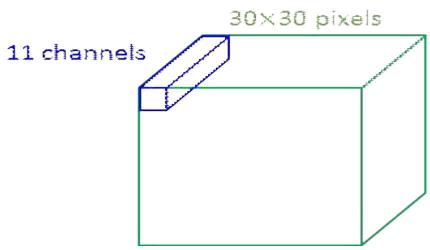

Fig. 2. The last feature map of our network. The size is 30×30 pixels and each pixel has a vector

The feature map in Fig.2 is composed of 11 channels, and it has vector values corresponding to 10 colors and no pixels. Information on the color palette used in this study is summarized in Table 1. Applying softmax to each pixel, a label with a high numerical value is applied to determine the color of the pixel.

Table 1. The color palette

Index	Color	(R,G,B)
0	Empty	(255,255,255)
1	Yellow	(250,200,10)
2	Orange	(214,121,35)
3	Red	(180,0,0)
4	Pink	(255,158,205)
5	Blue	(30,90,255)
6	Green	(88,171,65)
7	Brown	(95,49,9)
8	White	(244,244,244)
9	Gray	(100,100,100)
10	Black	(27,42,52)

4. Dataset

In this study, 240 × 240 input image and 30 × 30 pixel artified ground truth are required to learn neural network. We prepared about 2,000 data sets (Fig. 3).

First, the data set is automatically generated as a pixel art by applying the technology [15-17] possessed by the present research team, and then implemented the correction tool so that the ground truth can be prepared more easily.

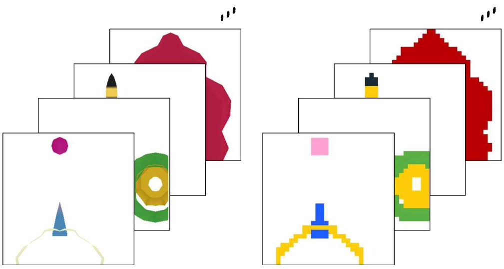

Fig. 3. Our dataset example: Raw images(left) and Ground truth images(right)

The original image and the pixel art image for the original image are paired so that the error can be obtained and the learning can proceed. Therefore, the file name of the original image and 30 × 30 pixel data were managed by csv file.

The image file is loaded with the file name of the original image and the discrete data of the image information is fetched, and the pixel data can be fetched by reading the csv file as it is. However, there are two problems in using this data set for neural network learning. To do this, an additional preprocessing process was required.

First, the frequency of the colors of each pixel in the pixel art output is unbalanced. For the most divergent colors, the difference was about four times. When learning is performed in this state, since the neural network is deflected to a color having a high frequency, it is solved by reducing the data to about 1000 so that the ratios of the respective colors are the same.

The second problem is the number of less data. Although not as many as 2,000, it was further reduced by the first problem.

Therefore, data augmentation was used to solve the problem of small number of data.
Rotation and inversion were applied to the augmentation, and the existing data could be increased 8 times.

In our data set, we divided 70% into learning data, 10% into validation data and the remaining 20% into test data.

5. Learning

The neural network proposed in this paper is implemented in Python 2.7 version and PyTorch 0.2 version. The learning was done by NVIDIA Titan X GPU operation in Ubuntu operating system environment.

The learning progressed to 100 epochs. The mean squared error was obtained and back propagation of the neural network was performed with the Adam optimizer. Accuracy is calculated by the difference between the result image of the neural network and the pixel art of the data set.

The true color of the two images is the same for each pixel. However, if the background is not a pixel, the accuracy is not taken into consideration.

If the background is not a background or a color is assigned, it is reflected in the calculation of accuracy.

Fig. 4 shows the accuracy of learning, verification and test data according to epoch.

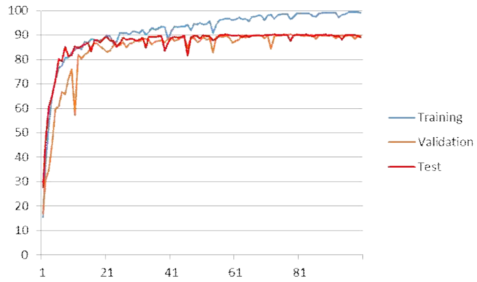

Fig. 4. A accuracy per epoch

6. Result

As can be seen in Figure 4, the accuracy of the test data is about 90%. The pixel artification results for the test data are shown in Fig. 5.

In addition to the test data with ground truth, we also tested other images, and the results are shown in Fig. 6.

7. Conclusion

In this paper, we propose a method using neural network to generate pixel art of image. A new neural network with modified CNN for the existing domain segmentation is presented and the neural network is learned from the data set of the research team. We also tested some images as well as test data.

In this study, white pixels and background pixels are distinguished. Thus, unlike other pixel art studies, we can see that the white pixel and the background are separated (Fig. 5, center column).

The limitation of this study is that the color palette must be fixed. For that reason, future studies will require you to choose a color palette that is appropriate for your situation.

In addition, you can expect to be able to distinguish similar colors if the learning data is sufficient.

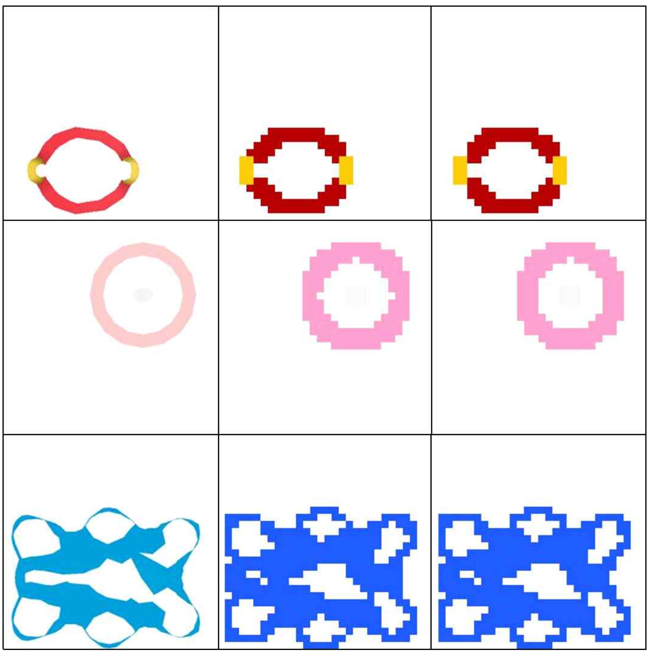

Fig. 5. Our result of test data

Fig. 6. Our result of other images

Acknowledgement

본 논문은 상명대학교 에너지그리드학과에서 수행하는 IoT와 패스트데이터 기반의 스마트에너지 산업 전문 인력 양성 사업(20164030300230)의 지원을 받았습니다. 또한, 한국 연구 재단에서 지원하는 연구 과제 (NRF-2017R1D1A1B03034137 및 NRF-2018R1D1A1A02050292)의 지원을 받았습니다.

Reference

Journal Articles

1. T.C. Inglise and C. S. Kaplan, "Pixelating Vector Line Art" *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering*, pp.21-28, 2012.
2. T.C. Inglise, et al, "Rastering and Antialiasing Vector Line Art in the Pixel Art Style" *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering*, pp.25-32, 2013.
3. Z. Hu and K. Urahama, "Cartesian Resizing of Line Drawing Pictures for Pixel Line Arts" *Journal of IEICE Transactions on Information and Systems, Vol. E97.D,* No. 4, pp.1008-1010, 2014.
4. M.R. Huang and R. R. Lee, "Pixel Art Color Palette Synthesis" *Proceedings of Information Science and Applications,* pp.327-334, 2015.
5. T. Gerstner, et al, "Pixelated Image Abstraction", *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering*, pp.29-36, 2012.
6. T. Gerstner, et al, "Pixelated Image Abstraction with Integrated User Constraints" *Journal of Comput. Graph., Vol. 37,* No. 5, pp.333-347, 2012.

7. M.H. Kuo, et al, “PIXEL2BRICK: Constructing Brick Sculptures from Pixel Art” *Journal of Computer Graphics Forum, Vol. 34,* No. 7, pp.339-348, 2015.
8. M.H. Kuo, et al, “Feature-aware Pixel Art Animation” *Proceedings of the 24th Pacific Conference on Computer Graphics and Applications,* pp.411-420, 2016.
9. J. Kopf and D. Lischinski, “Depixelizing Pixel Art” *Journal of ACM Trans. Graph. Vol. 30,* No. 4, pp.99, 2011.
10. Y. LeCun, et al, “Gradient-based learning applied to document recognition” *Journal of the IEEE. Vol. 86,* No. 11, pp.2278-2324, 1998.
11. P.O. Pinheiro and R. Collobert, “From Image-Level to Puxel-Level Labeling With Convolutional Networks” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR),* pp.1713-1721, 2015.
12. J. Long et al, “Fully Convolutional Networks for Semantic Segmentation” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR),* pp.3431-3440, 2015.
13. J. Uhrig, et al, “Pixel-Level Encoding and Depth Layering for Instance-Level Semantic Labeling” *Proceedings of Pattern Recognition,* pp.14-25, 2016.
14. O. Ronneberger, et al, “U-Net: Convolutional Networks for Biomedical Image Segmentation” *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI),* pp.234-241, 2015.
15. H. Yang, et al, “An Abstraction Technique for Creating Game Backgrounds of Selective Multiscale Struecture” *Journal of Korean Society for Computer Game, Vol. 19,* pp.131-139, 2009.
16. D. Kang, et al, “Implementation of Real-Time Video Abstraction using CUDA” *Journal of Korean Society for Computer Game, Vol. 24,* No. 3, pp.125-135, 2011.
17. G.R. Yun, et al, “A Multiscale Voxelization for Abstraction of Polygonal Model” *Journal of Korean Society for Computer Game, Vol. 29,* No. 3, pp.1-8, 2016.

<국문초록>

합성곱 신경망을 이용한 게임 캐릭터의 픽셀 아트 생성

박철성, 양희경, 권혁민, 민경하

픽셀 아트는 낮은 해상도와 제한된 색 팔레트를 가지고 영상을 표현한다. 픽셀 아트는 낮은 연산 성능과 적은 저장 공간을 가지는 초기 컴퓨터 게임에서 주로 사용되었다. 현대에 이르러, 픽셀 아트는 예술이나 퍼즐, 게임과 같은 다양한 분야에서 찾아볼 수 있게 되었다.

본 논문에서는 게임 캐릭터 영상을 입력으로 받는 픽셀 아트 생성 모델을 제안한다. 기존 방법과는 달리, 합성곱 신경망(CNN:Convolutional-Neural Network)를 픽셀 아트 생성 목적에 맞게 변형하여 이를 이용하는 방법을 제시한다. 기존의 합성곱 연산 후에 upsampling 과정을 추가하여 픽셀 아트가 생성될 수 있도록 하였다. 네트워크는 ground truth와 생성된 픽셀 아트와의 평균 오차 제곱(MSE:Mean Squared Error)을 최소화해나가며 학습을 수행한다.

Ground truth는 실제 아티스트가 생성하도록 하였고, 이미지 회전과 반전 기법을 이용하여 augumentation을 수행하였다. 생성된 데이터 집합은 학습, 검증, 시험 데이터로 나누었다. 이러한 데이터 집합을 기반으로 감독 학습을 실시하여 픽셀 아트 생성 네트워크를 학습하였다. 학습 모델의 학습 과정과 학습 정확도를 제시하고, 시험 데이터 뿐만 아니라 다양한 영상에 대한 픽셀 아트 결과도 함께 제시한다.

<결론 및 향후 연구>

본 논문에서는 영상의 픽셀 아트를 생성하는 신경망을 이용한 방법을 제시하였다. 기존의 영역 분할을 위한 CNN을 변형한 새로운 신경망을 제시하였으며, 본 연구팀의 데이터 셋으로 신경망을 학습시켰다. 또한, 시험데이터 뿐만 아니라, 일부 영상에 대해서도 시험해 보았다. 본 연구에서는 흰색 픽셀과 배경 픽셀을 구분하여 학습시켰다. 때문에, 다른 픽셀 아트 연구와 달리 흰색 픽셀 부분과 배경 부분을 구분한 것을 볼 수가 있다.

본 연구의 한계점으로는 색 팔레트를 고정해야 한다는 점이 있다. 그렇기 때문에, 상황에 맞는 색 팔레트를 선택하도록 하는 것이 향후 연구가 될 수 있다. 또한, 학습 데이터가 충분하다면 비슷한 색들도 구분할 수 있을 거라 기대할 수 있다.

<Author Biography>

Cheonseong Park(박철성)
2010-2016; BS, Sangmyung Univ.
2016-present; MS, Sangmyung Univ.
Research interest; Computer graphics, Game programming, Non-photorealistic rendering, Deep learning

Heekyung Yang(양희경)
2006-2010; BS, Sangmyung Univ.
2010-2012; MS, Sangmyung Univ.
2012-present; Ph.D. candidate, Sangmyung Univ.
Research interest; Computer graphics, Game programming, Non-photorealistic rendering, Deep learning

Hyukmin Kwon(권혁민)
2006-2012; BS,
2016-present; MS, Sangmyung Univ.
Research interest; Game implementation

Kyungha Min(민경하)
1988-1992; BS, KAIST
1992-1994; MS, POSTECH
1994-2000; Ph.D., POSTECH
2006-present; Professor, Sangmyung Univ.
Research interest; Computer graphics, Game programming