Unsupervised Deep Learning Approach for Aliasing Removal of Screen Captured Image

Sohyeon Park, Myungjin Sim
Ewha Womans University

cookiedough0911@ewhain.net kp6069@ewhain.net

Abstract

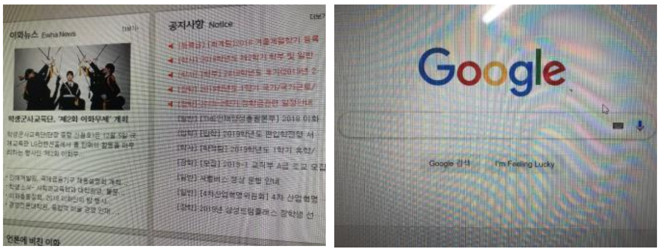
When we take pictures of computer monitors with digital devices such as mobile phones, aliasing patterns occur which causes visual damage to the original image. The idea of the research is to study a way to eliminate the aliasing pattern using deep learning in order to achieve clear images even when we take pictures of computer monitors with a mobile phone. In this paper, we propose a model that removes aliasing patterns by using unsupervised deep learning. With the proposed model which uses CNN(Convolutional Neural Network), the aliasing patterns are removed to the extent where it can be confirmed with naked eye.

Key words : Unsupervised deep learning, CNN, Aliasing pattern, Siamese Network

1. Introduction

As the period of artificial intelligence robot engineering advances, robots are replacing humans in performing various tasks. However, looking at the digital monitors through the robots' eyes may cause their ability to identify objects to deteriorate significantly, since the images are impaired due to aliasing patterns. This can be confirmed in real life, Fu, Xueyang, et al. [3] mentions that the performance of the pedestrian recognition algorithm, which is used very often in computer vision, decreases when applied in rainy days because the rain patterns impairs the pedestrian images.

Aliasing refers to a phenomenon in which noise or signal transformation occurs in the original signal due to the change or transformation of the high frequency component to a low frequency during the sampling process, and this creates certain patterns which are called



[Figure 1] Images where aliasing patterns occurred

aliasing patterns. These aliasing patterns can easily be recognized in real life. For example, as in Figure 1, taking pictures of computer monitors with mobile phones creates many aliasing patterns that can impair the original image unlike what we desired. In this paper, we aim to provide a clean image even when taking pictures of computer monitors with mobile phones, by studying a way to remove such aliasing patterns using unsupervised deep learning.

This paper proposes a method which uses CNN, where CNN(Convolutional Neural Network) is a model that enables people to train images while maintaining the spatial information.

2. Theoretical Background

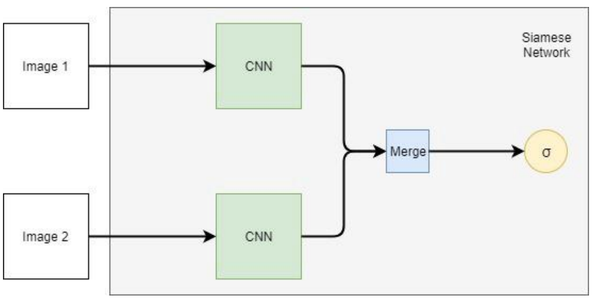
In order to provide a clear original image by removing the aliasing patterns from the image taken on the computer monitor, we decided to apply the unsupervised deep learning approach.

1) Unsupervised Deep Learning

Unsupervised deep learning is a learning method which combines data with similar characteristics and classifies them into groups. It refers to finding a specific pattern among the data without any information about the result or ground truth. The purpose of unsupervised learning is to identify the type of relationship the data consist of. Since unsupervised deep learning can be trained without ground truth, it is possible to train data by computer itself without human intervention. However, it is necessary to input a large amount of input data in order to replace the ground truth during training. In this paper, images of the same monitor screen photographed from various angles will be used as input data.

2) CNN

CNN (Convolutional Neural Network) is an artificial neural network used to process data composed of pixels such as images. Neural network is a software system that patternizes the behavior of neurons in the human brain. Image processing was not ideal because existing neural networks could only process low-resolution images. In contrast, CNN uses a system which is very similar to a multilayer perceptron, that learns nonlinearly separated data by arranging neurons in a manner that covers the entire field of view similar to the frontal lobe, which is an area that processes human visual stimuli.
In this paper, a Siamese network is used to train images, and the Siamese network trains by comparing the difference between the results after putting two input data into the CNN as shown in [Figure 2], which removes the aliasing pattern in the input data.



[Figure 2] Relationship between Siamese Network and CNN

3. Related Work

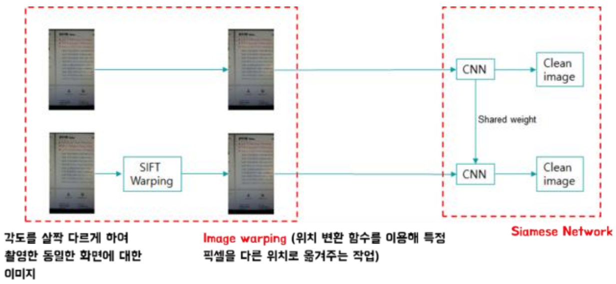
1) Prior research on unsupervised deep learning

Louis Lettry et al. [1] used unsupervised deep learning to extract only the albedo, the eigenvalue of the image, by dividing the image into albedo and shading. In order to replace the ground truth, two images consisting of the same albedo but different shades were input as input data and training was conducted. Using this training technique makes it possible to extract albedo from a single image during testing.

2) Prior research on CNN

Kai Chang et al. [2] added a layer to the existing CNN structure and modified the learning algorithm and normalization method to remove noise in images in a more effective way. Specifically, the training speed and noise removal performance were greatly improved by using residual learning and batch normalization.

4. Design



[Figure 3] Overall Structure

Figure 3 shows the overall structure of the method this paper presents. The overall network is a Siamese network, where two images are independently put into the same CNN model and the difference between the output values is calculated, which then gradually reduces the difference value through training. The two images at the far left refer to the image data set of the same screen obtained by taking a picture with a mobile phone, of a monitor screen. This is a pre-processing process before putting the image data sets into the deep learning model. The two images go through SIFT Warping which makes the positions of the two images completely identical. Image warping refers to the operation of moving a specific pixel to another location using a conversion function.

5. Implementation

1) Development environment

[Figure 4] Software and Hardware Environment

The development environment is as in Figure 4. A total of 2600 sets (5200 photos) of data sets were taken manually using a computer/TV monitor screen and camera. The development language used was Python to utilize Pytorch library, and for the server, GTX 1080Ti GPU server was used.

2) Implementation algorithm

2-1. Unsupervised Learning

The advantage of unsupervised learning is that the model can be trained without the ground truth in order to provide a clean image. Therefore, the corresponding method was adopted, and since unsupervised learning requires multiple images for comparison during learning, an image of the same monitor screen taken from various angles was prepared as a data set. The two images of the same screen were taken at different angles and positions to obtain images with different aliasing patterns.

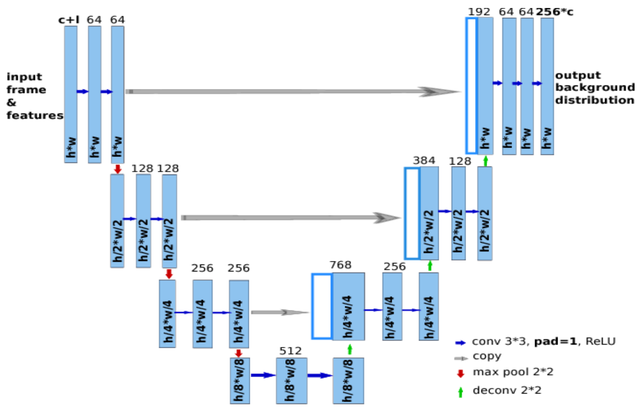
2-2. SIFT Algorithm

The SIFT image matching algorithm was used for warping. SIFT (Scale-Invariant Feature Transform) is an algorithm which extracts features that are invariant to the size and rotation of an image. When SIFT features are extracted from two different images, and the features that are most similar to each other are matched, the corresponding part of the two images is found. In this study, such an algorithm was used to perform image warping.

[Figure 5] SIFT warping result

The image placed in the most right in Figure 5 is the result of image warping, before adding it to the deep learning model, using the modified SIFT algorithm. Image warping was applied to all 2600 sets of data sets, and approximately 10000 sets of data were derived by cropping each of them. As shown in Figure 5, the result image was derived as a result of the source image warped to match the target image. It can be seen that the target image and the result image differ only in its patterns, and the content of the original image is maintained.

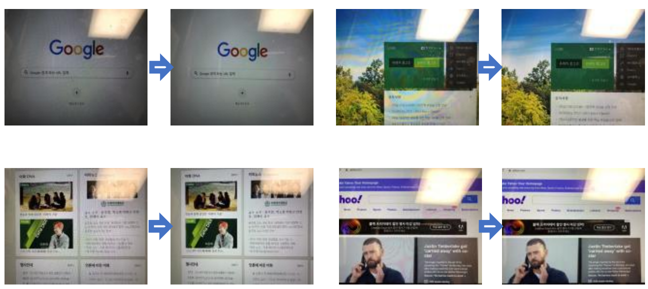
2-3. Unet



[Figure 6] Unet structure

Unet model is used for training, which is an encoder-decoder structure where both the model input value and the output value are forms of an image. The warping-completed image is put into the Unet, and a feature map, that is suitable for extracting features, is generated by processing through each convolution layer, max pooling layer, etc., and the corresponding layer is reversed through deconvolution to restore the original image form. The model referenced the Beyond a Gaussian Denoiser [2] model, transformed it into a Siamese network, then performed tasks such as changing the loss and optimizer. We used l2 loss for the loss variable, which is effective in obtaining the difference between two image pixel values, and adam optimizer was used for the optimizer variable for fast convergence.

2-4. Result



[Figure 7] Test results

As a result of training the model with 2500 sets of training data sets, the loss converged to 0, and Figure 7 shows the output image when a part of the testing data is put as the input data. When compared with the input image, it can be noticed that the aliasing patterns in the original image have been removed without causing any damage to the content of the image as an output.

6. Conclusion

In this paper, we implemented an algorithm that outputs a clean image by removing aliasing patterns which appear when a monitor screen is photographed by mobile phone, using an unsupervised deep learning model. With the recent development of artificial intelligence technology, systems such as autonomous vehicles and autonomous drones are being developed. Techniques such as computer vision-based pedestrian detection and image segmentation used in the systems mentioned above are highly influenced by the aliasing patterns, which has a negative effect on the system extracting features. Therefore, this study can be applied to various artificial intelligence systems.

7. Reference

[1] Louis Lettry, Kenneth Vanhoey, Luc van Gool. "Unsupervised Deep Single-Image Intrinsic Decomposition using Illumination-Varying Image Sequences", 2018.

[2] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, Lei Zhang. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising", 2016.

[3] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, John Paisley. "Removing rain from single images via a deep detail network", 2017.