National ChengChi University, Taipei, Taiwan, May 5-6, 2017

第5屆台灣工業與應用數學會年會

(海報論文：大專生組)

Style Transfer using DeepLearning

Hsien_Hao Chen(陳先灝) **Han-Yi Shing**(邢恒毅) **Zi-Yu Huang**(黃梓育) **Yi-Hsun Lin** (林奕勳) **Bo-Kuan Wu**(吳柏寬) **Yuan-Hao Fan**(范原豪)
Department of Mathematical Sciences, National Chengchi University, Taiwan (102701008@nccu.edu.tw)
Advisor: Prof. Yen-Lung Tsai (蔡炎龍)

Abstract

In fine art, humans are good at creating paintings in different styles. Computer scientists have tried different ways to approach this. According to the paper from Gatys et al.[1] , we built a style-transfer model using DeepLearning. However, we find it sometimes comes out with an unsatisfying result, so we tried to get a better output through giving different parameters or adjusting the weights of the function.

Problem description

We give two images as input, one attributes the content, and the other attributes the style. Our objective is to transfer the style of one image to another. First, we generate an image with random noise, then put the images, including the two as inputs, into a pre-trained model such as VGG-16 or VGG-19. After that we calculate the loss between the generated image, the content image and the style image. We then optimize the image with the help of gradient descent.

The problem we encountered is that the image we optimized didn't satisfy us. The style is not transferred properly. We believe that the tragedy occurred due to our loss is stuck in an local minimum during the process of gradient descending. In order to solve this situation, we started to change the parameters in our model hoping to get a better transferred result. The parameters we modified include the pre-trained models, loss functions, optimizers and some other parameters in the neural network.

Results and discussion

According to our survey on facebook, some people thinks that the brushwork was too delicate compared to the style image.

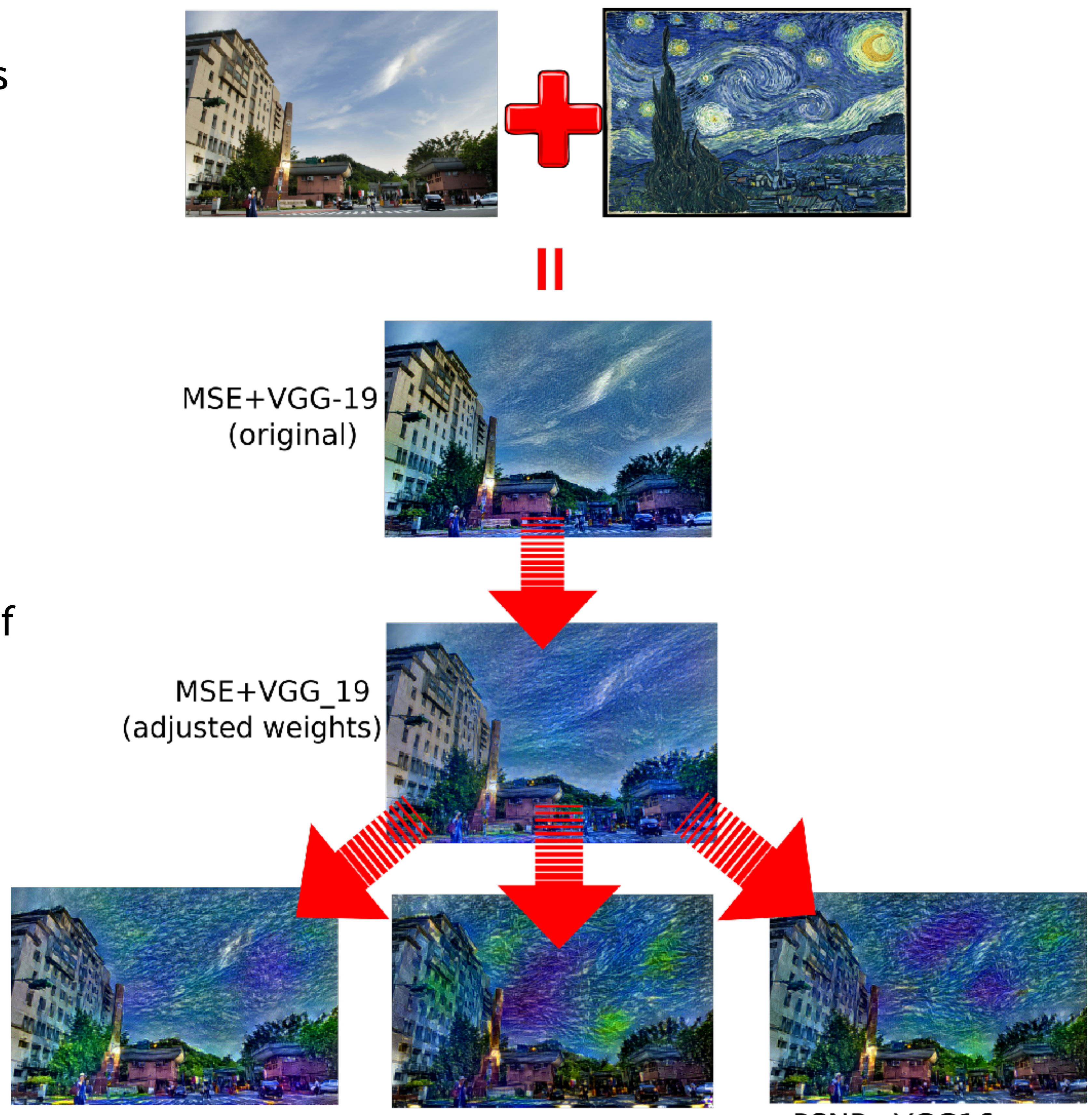
We discussed several possible reasons for this result.

The first is that the coefficient of the content-weight is too high.Secondly, VGG-19 might not be an ideal pre-train model in this case. The third reason is that the loss function for content-loss in the original paper is mean square error, which has been found to be a poor method to calculate the distance between two images.

For the first reason, we gave a style-weight ten times larger than the original one. The result was better in terms of style, but we consider it to be overfitting to the style. The content is gone in some parts of the result, and the image doesn't seem to be smooth enough. Therefore, we increased the content-weight and the total-variational-weight and decreased the style-weight. After these modification, we got a relatively better outcome.

Fot the second reason, we changed the pre-train model to VGG-16.We found that VGG-16 somehow did a better job in identifying features.

For the third reason, we tried PSNR and SSIM as our content-loss, and get some interesting output.



Conclusions

After all these experiments, we have the following conclusions:

1. A balanced weight is required. The style-weight should be 10-20 times larger than content-weight. The total-variation-weight should be 0.5-2 times of the style-weight.
2. VGG-16 preserves more style features during the transfer, while VGG-19 preserves more content and smoothness.
3. In this case, because we update our weight pixelwise, MSE strike a better balance than other assession like PSNR or SSIM, though they are better loss function than MSE in most of case.

References

1.Gatys, L.A., Ecker, A.S., Bethge, M.: A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576 (2015)
2.Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution. arXiv preprint arXiv:1603.08155 (2016)
3.fchollet/neural_style_transfer https://github.com/fchollet/keras/blob/master/examples/neural_style_transfer.py