Paper Review

Author: Yunfei Luo

May 12, 2020

A Neural Algorithm of Artistic Style, by Gatys, Ecker, and Bethge
Reference Paper URL: https://arxiv.org/abs/1508.06576

Introduction/Main Goal

The paper was published in 2015. The main goal of the paper is to introduce an artificial system that could create artistic images like what human could do with painting. The core of the system is a neural algorithm based on Deep Neural Networks. A sub-goal of the paper is to describe how human create and perceive artistic images, with algorithmic approach.

What's New/Improvement

The authors are not trying improve, but rather, explore a new field of application of Deep Neural Networks. They notice that, during the image processing by the neural networks, the actual features (the mapping of different filtered versions of image) and the correlation between different filters can be individually visualized at each layer. They called them "content representation" and "style representation" respectively. So they are thinking about combining the representations of content and style to make something out.

Observations

As a paper on the neural networks, the input and output in this case are both images, with the normal image structures, i.e. the 2-D matrix of pixels. More specifically, the inputs are two images, one for capturing the content representation, one for extracting style representation. The output is an image that combine the content and style representation through image reconstruction. Moreover, there do exists a hidden state. Since during the match of content and style, if emphasizing more on the content, it will be obvious to detect what was the original photograph with less variant in the style. Same as if emphasizing more on the style, then content of the image might be hard to see, i.e. the image would be too abstractionism. So the hidden state would be some reconstructed image with "just-right" weights for content and style.

Result

In order to dealing with this tasks, the authors done it through a optimization problem on minimizing the linear combination of the loss functions of content and style. The weights, or the coefficients of the loss functions could be understand as the "regularizers" that serves for the extent of how much the image emphasize on content and style. For the result that the authors got (page 5 in the paper):
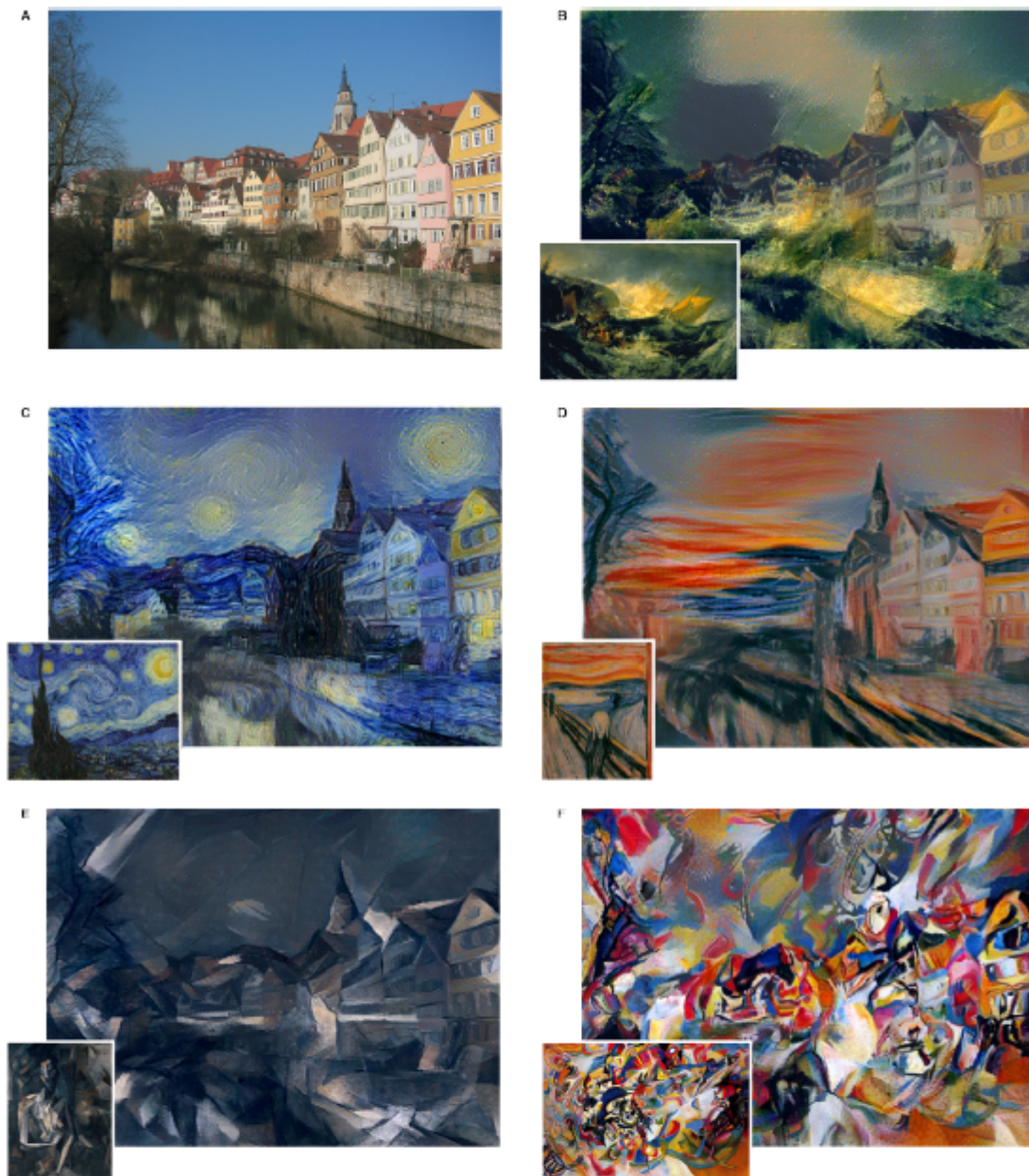
Figure 2: Images that combine the content of a photograph with the style of several well-known artworks. The images were created by finding an image that simultaneously matches the content representation of the photograph and the style representation of the artwork (see Methods). The original photograph depicting the Neckarfront in Tübingen, Germany, is shown in **A** (Photo: Andreas Praefcke). The painting that provided the style for the respective generated image is shown in the bottom left corner of each panel. **B** *The Shipwreck of the Minotaur* by J.M.W. Turner, 1805. **C** *The Starry Night* by Vincent van Gogh, 1889. **D** *Der Schrei* by Edvard Munch, 1893. **E** *Femme nue assise* by Pablo Picasso, 1910. **F** *Composition VII* by Wassily Kandinsky, 1913.

It seems like the Deep Neural Networks perform pretty well on applying different artistic styles, from learning of other artworks, on the photo shown in **A**. To intuitively interpret these results, the neural algorithm do effectively facsimile the photograph with artistic styles of some well-known artists, such as J.M.W. Turner and Vincent Van Gogh.

Extension/Follow up

What would happen if combining content with content, and style with style? An intuitive instance would be, suppose that there is a photo of a cat, and another photo of a dog, what if we use the content representations of these two images from the Deep Neural Network for image reconstruction? How to construct the optimization problem in this case, in order to made an image that have the key content in both original images with relatively "just right" weights and positions? These follow-up questions offers the possible extension for the work presented in the paper.