Final Project

Neural Style Transfer Explorations

Conor O'Mahony

CSCI S-89 Introduction to Deep Learning
Summer 2024
Harvard Summer School

Goal of Project

- Below is a photograph of me in front of a couple of my paintings.
- I like to paint portraits with a pop-art style:
  - Divide the face into segments based on tone.
  - Use straight lines for the segment borders.
  - Choose bold, contrasting segment colors.
- Goal: take a photograph of a person as input, and generate a pop-art style portrait like these.

A Neural Algorithm of Artistic Style

Leon A. Gatys,[1,2,3*] Alexander S. Ecker,[1,2,4,5] Matthias Bethge[1,2,4]

[1]Werner Reichardt Centre for Integrative Neuroscience
and Institute of Theoretical Physics, University of Tübingen, Germany
[2]Bernstein Center for Computational Neuroscience, Tübingen, Germany
[3]Graduate School for Neural Information Processing, Tübingen, Germany
[4]Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[5]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA
[*]To whom correspondence should be addressed; E-mail: leon.gatys@bethgelab.org

In fine art, especially painting, humans have mastered the skill to create unique visual experiences through composing a complex interplay between the content and style of an image. Thus far the algorithmic basis of this process is unknown and there exists no artificial system with similar capabilities. However, in other key areas of visual perception such as object and face recognition near-human performance was recently demonstrated by a class of biologically inspired vision models called Deep Neural Networks.[1,2] Here we introduce an artificial system based on a Deep Neural Network that creates artistic images

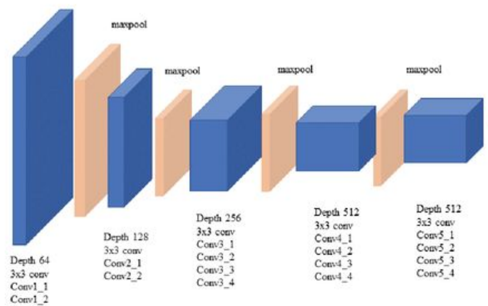
Leon A. Gatys, Alexander S. Ecker, Matthias Bethge. *A Neural Algorithm of Artistic Style*. arXiv (2015). https://arxiv.org/pdf/1508.06576

Neural Style Transfer

- Take the <u>content</u> from a **Base Image**, apply the <u>style</u> from a **Style Image**, and create a **Generated Image**.

Base Image

Style Image



VGG-19 Based Convolutional Neural Network

Generated Image

Separating the Content and Style

- Need: Separately identify the <u>content</u> of an image and the <u>style</u> of an image.
- Need: Measure content and minimize the content loss compared to base image.
- Need: Measure style and minimize the style loss compared to style image.

Content

- Higher layers in CNN capture the high-level content.
- Use a single layer.
- Use squared error loss.

Style

- Style consists of the textures, colors, and visual patterns in the image.
- Combining feature correlations across layers captures the style of an image.
- Use gram matrices.

Combining the Content and Style Loss

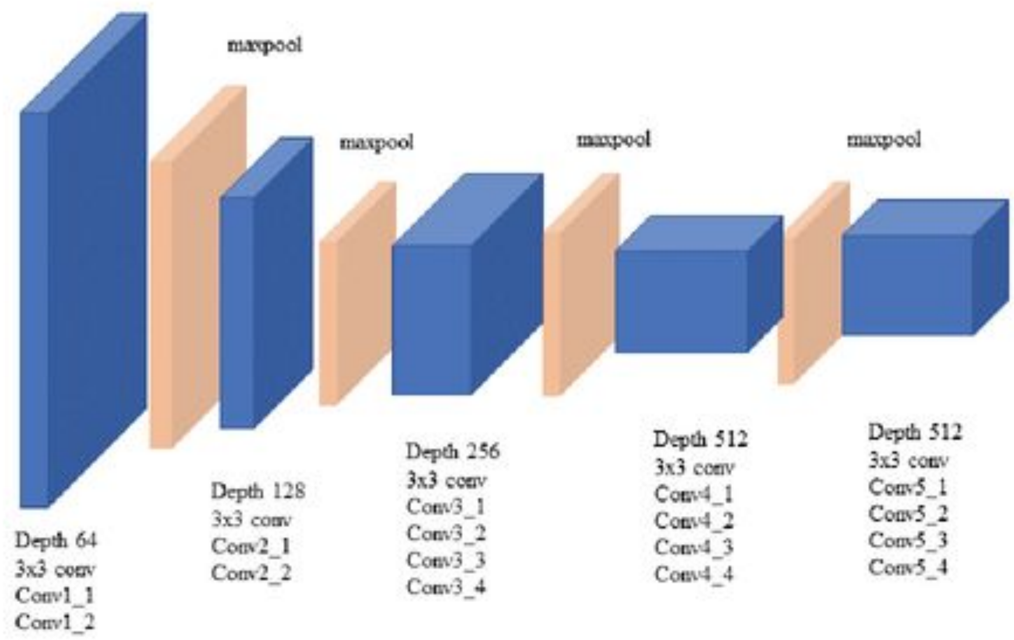
Loss = $\boldsymbol{\alpha}.\mathcal{L}_{\text{content}} + \boldsymbol{\beta}.\mathcal{L}_{\text{style}} + \boldsymbol{\gamma}.\mathcal{L}_{\text{total\_variational}}$

where:

$\boldsymbol{\alpha}$	is the content weight (controls contribution of content loss)
$\mathcal{L}_{\text{content}}$	is the content loss
$\boldsymbol{\beta}$	is the style weight (controls contribution of style loss)
$\mathcal{L}_{\text{style}}$	is the style loss
$\boldsymbol{\gamma}$	is the total variational weight (controls contribution of total variational loss)
$\mathcal{L}_{\text{total\_variational}}$	is the total variational loss

Choice of Neural Network

- Take advantage of transfer learning.
- Use the 16 convolutional and 5 pooling layers of the 19-layer VGG-Network (VGG-19).
- Using the ImageNet weights.



Initial Neural Style Transfer

- Adapted code from Francois Chollet's book: Deep Learning with Python.
- See file: `1_Neural Style Transfer - Baseline.ipynb`

Base Image

Style Image

Generated Image

Francois Chollet. *Deep Learning with Python*, Second Edition. Pages 383-391. Manning Press (2021).

Experimenting with Optimizer

- Original Gatys et al paper uses L-BFGS optimization.
- Chollet uses Stochastic Gradient Descent optimizer.
- Experimented with SGD, Adam, RMSProp, Nadam.
- See file: `2_Neural Style Transfer - Optimizer Experiments.ipynb`

Baseline Image

RMSProp, lr=0.1, decay

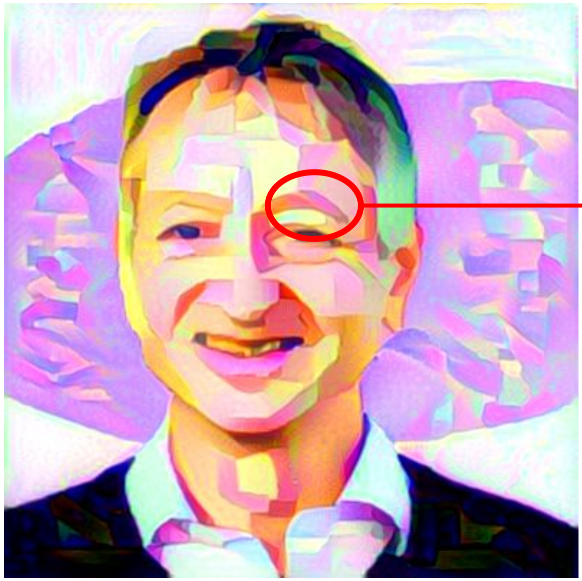
Experimenting with Loss Weights

- Chollet uses the following weights:

```
content_weight = 2.5e-8          # 0.00084
style_weight = 1e-6              # 0.0025
total_variation_weight = 1e-6    # 0.0025
```

- Experimented with weights, getting very slight improvements
- See file: `3_Neural Style Transfer - Loss Experiments.ipynb`

Updated Baseline Image



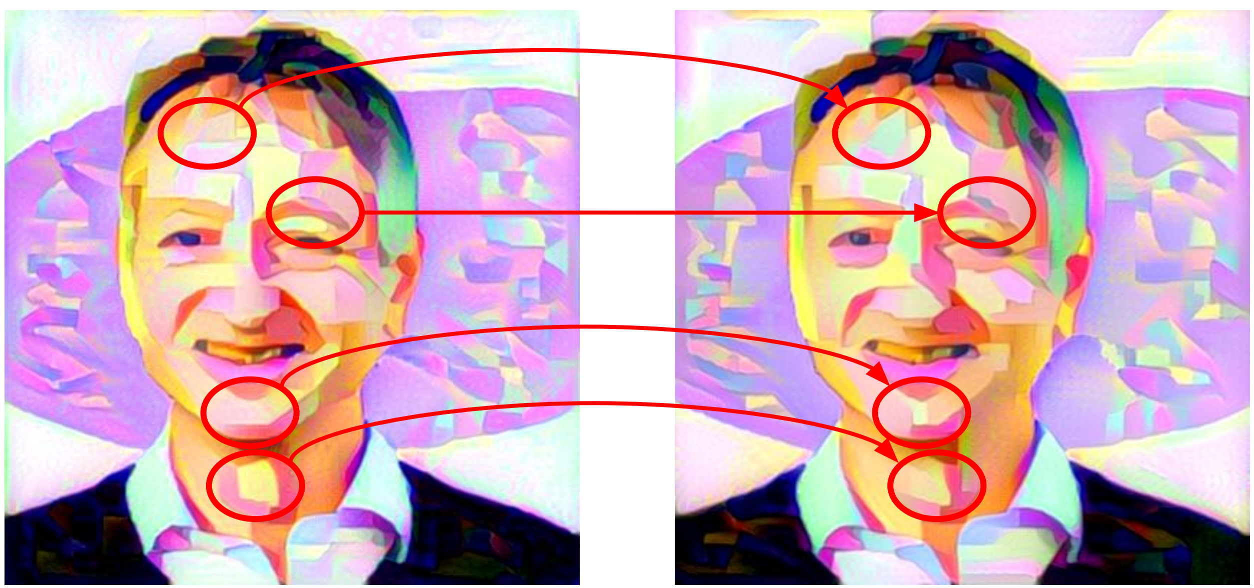
content_weight=2.5e-8, style_weight=1e-2, total_variational_weight=1e-2

Experimenting with Layers

- For the content layer: Chollet uses `block5_conv2`
- For the style layers: Chollet uses `block1_conv1`, `block2_conv1`, `block3_conv1`, `block4_conv1`, and `block5_conv1`
- Experimented with layer choices.
- See file: `4_Neural Style Transfer - Layers Experiments.ipynb`

Updated Baseline Image

Content: `b5c2`
Style: `b1c1`, `b2c2`, `b3c3`, `b4c2`, `b5c3`



Putting it all Together

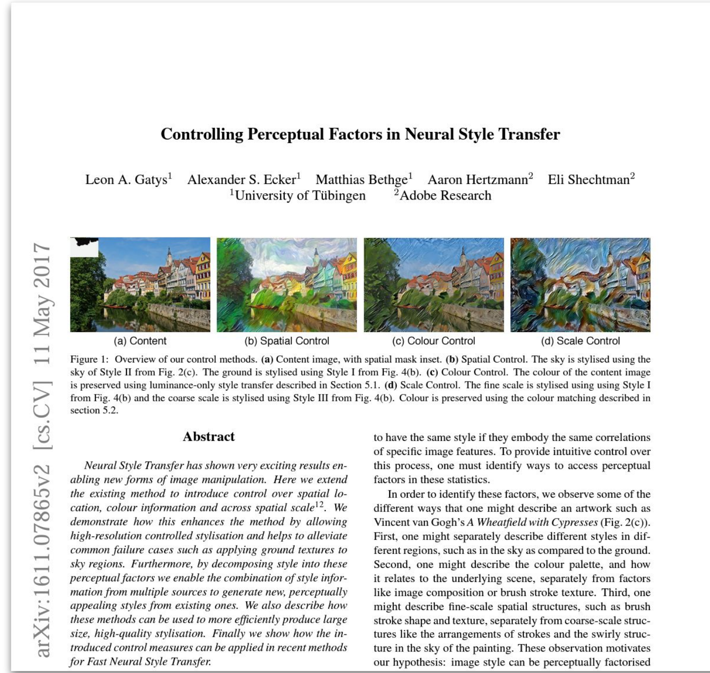
- Choose the best settings from each of these experiments.
- Generate a new baseline image.
- See file: `5_Neural Style Transfer - Optimized.ipynb`

Original Generated Image

New Generated Image

Controlling Perceptual Factors

- In 2016, Leon Gatys et al introduced several enhancements to Neural Style Transfer
- These include:
  - **Spatial control** to control style in different parts of the image. We will use it so we don't style the background.
  - **Color control** to keep the original base image colors, rather than using the style image colors. This does not make sense for our project.
  - **Scale control** to mix different styles at different scales in the generated image. Again, this is not something we want to do for this project.



arXiv:1611.07865v2 [cs.CV] 11 May 2017

Controlling Perceptual Factors in Neural Style Transfer

Leon A. Gatys[1] Alexander S. Ecker[1] Matthias Bethge[1] Aaron Hertzmann[2] Eli Shechtman[2]
[1]University of Tübingen [2]Adobe Research

(a) Content (b) Spatial Control (c) Colour Control (d) Scale Control

Figure 1: Overview of our control methods. **(a)** Content image, with spatial mask inset. **(b)** Spatial Control. The sky is stylised using the sky of Style II from Fig. 2(c). The ground is stylised using Style I from Fig. 4(b). **(c)** Colour Control. The colour of the content image is preserved using luminance-only style transfer described in Section 5.1. **(d)** Scale Control. The fine scale is stylised using using Style I from Fig. 4(b) and the coarse scale is stylised using Style III from Fig. 4(b). Colour is preserved using the colour matching described in section 5.2.

Abstract

Neural Style Transfer has shown very exciting results enabling new forms of image manipulation. Here we extend the existing method to introduce control over spatial location, colour information and across spatial scale[12]. We demonstrate how this enhances the method by allowing high-resolution controlled stylisation and helps to alleviate common failure cases such as applying ground textures to sky regions. Furthermore, by decomposing style into these perceptual factors we enable the combination of style information from multiple sources to generate new, perceptually appealing styles from existing ones. We also describe how these methods can be used to more efficiently produce large size, high-quality stylisation. Finally we show how the introduced control measures can be applied in recent methods for Fast Neural Style Transfer.

to have the same style if they embody the same correlations of specific image features. To provide intuitive control over this process, one must identify ways to access perceptual factors in these statistics.

In order to identify these factors, we observe some of the different ways that one might describe an artwork such as Vincent van Gogh's *A Wheatfield with Cypresses* (Fig. 2(c)). First, one might separately describe different styles in different regions, such as in the sky as compared to the ground. Second, one might describe the colour palette, and how it relates to the underlying scene, separately from factors like image composition or brush stroke texture. Third, one might describe fine-scale spatial structures, such as brush stroke shape and texture, separately from coarse-scale structures like the arrangements of strokes and the swirly structure in the sky of the painting. These observation motivates our hypothesis: image style can be perceptually factorised

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Aaron Hertzmann, Eli Shechtman. *Controlling Perceptual Factors in Neural Style Transfer*. arXiv (2016). https://arxiv.org/pdf/1611.07865

Adding Spatial Control

- Let's focus the application of style to Geoffrey Hinton himself.
- Do not apply style to the background.
- See file: `6_Perceptual Control.ipynb`

Before
Spatial Control

After
Spatial Control

Review

- Applied Neural Style Transfer.
- Experimented with settings.
- Added Spatial Control.

Base Image

Original Generated Image

Final Generated Image

Future Work

- Try different pre-trained CNNs, including VGGFace2.
- Experiment with Selim et al’s efforts to improve Neural Style Transfer for portraits.
- Use the findings from these experiments to implement Real-Time Style Transfer.
- For more details, see the accompanying Report.

Ahmed Selim, Mohamed Elgharib, Linda Doyle. *Painting Style Transfer for Head Portraits using Convolutional Neural Networks*. ACMTrans. Graph 35, 4, Article 129 (July 2016). https://dl.acm.org/doi/10.1145/2897824.2925968

Justin Johnson, Alexandre Alahi, Li Fei-Fei. *Perceptual Losses for Real-time Style Transfer and Super-resolution*. arXiv (2016). https://arxiv.org/pdf/1603.08155

YouTube Video Presentation

YouTube video presentation: https://youtu.be/FIDVFDB5_Xk