Our implementation of Gatys’ algorithm to produce some synthetic artwork. We also implement a spatial control extension to Gatys’ algorithm and check how the spatial control extension improves output image quality.

Method

The style representation of an image $x$ at layer $l$ is represented by the Gram matrix $\mathbf{G}^{(l)}(x) = \left[ \mathbf{F}^{(l)}(x) \right] ^T \left[ \mathbf{F}^{(l)}(x) \right]$. 

To transfer the style of an artwork $a$ onto a photograph $p$ and produce a synthesized image $x$, we initialize a random image $x$, and minimize the total loss function defined as a linear combination of content loss, style loss, and total variational loss.

$$L_{\text{total}}(p, a, x) = \alpha L_c(p, x) + \beta L_s(a, x) + \gamma L_v(x)$$

Where $L_c(p, x)$ is the content loss, $L_s(a, x)$ is the style loss, and $L_v(x)$ is the total loss. 

If spatial control is applied, both style and content images are divided into $K$ regions and the style loss above is replaced with normalized style loss:

$$L_s(a, x) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{N_k} \left\| \sum_{l \in K_k} \mathbf{F}^{(l)}(x) - \mathbf{G}^{(l)}(a) \right\|_F^2$$

In our implementation of Gatys’ method we use weights of VGG-19 (Simonyan & Zisserman, 2014), a 19-layer deep convolutional neural network which has been pre-trained on the ImageNet dataset.

Results, cont.

Our addition improves output image quality.

Spatial control:

Difficult case

Discussion

We implemented Gatys’ neural style transfer algorithm, which is capable of arbitrary style transfer and produces high quality output, but has high computational cost (>1 hour CPU runtime for 500x500 image, 300 iterations). Our output images are limited in resolution and contains residual noise from random image initialization. Additionally, each pair of [style, content] images requires custom hyperparameter tuning in order to obtain the most visually impressive result.

➢ Our spatial control method requires the user to generate two additional input masks. It enables us to obtain higher quality output for some examples, but introduces distortion at the mask boundaries, and fails if the art style is too photorealistic.

Future Work

Potentially interesting future work of this project includes:
1. Implement additional control options, such as color control, feature size control, and blending multiple styles.
2. Implement a real-time style transfer algorithm and compare its performance with Gatys method.
3. Explore options for automatic hyperparameter tuning.

References