ART AND MACHINE LEARNING
CMU 2019 SPRING
PROJECT 1

The Urban Jungle

"The Reclamation" - Kensington Market, Toronto

Jenn Choi, Julie Kim, Jean-Baptiste Lamare, Teven Le Scao, Sophia Yoo

Concept

Our project focuses on the juxtaposition between industrial and natural elements in metropolitan areas across the world. The class discussions on art and technology sparked an interest in framing urban photos with a heightened awareness of natural vs. man-made objects. Namely, we wanted to exhibit how challenging it can be to strike a balance between the two in this age of technology. Our objective is to highlight the “hidden” natural beauty in cities, thus creating an Urban Jungle. The inspiration behind this project came from Vicente Munoz’s *Sublimis* series [1], which portrays the “inevitable struggle between man and nature.” Munoz captured photos with color infrared film, filtering all natural objects in his shots with infrared radiation. His photography depicts what we “still have left unconsumed.” With the same frame in mind, we applied various machine learning techniques to transform cities to present nature as the leading object. Our revisioned landscapes, using our own photos of cities from Pittsburgh to Jerusalem, lend greater importance to the few pockets of greenery left in modern cities.

Techniques

We used Style Transfer, DeepDream and Neural Doodle in this project. However, we only focus here on the variations/adaptations of the original algorithms we have worked on since the three original models were described in class.

Masked Style Transfer

Our objective being to highlight nature in cities, one of our ideas was to turn city buildings into trees, cities into a forest, thus creating an “urban jungle”. Because creating a forest is mostly about texture, we decided to use style transfer to use the style from the picture of a forest and apply it to the picture of a city. However, we only wanted to apply style transfer on a part of the image. This is why we use masked style transfer, described for example in [2] and [3]. We tried two techniques described in [3]: Naive masked style transfer, and masked style loss. We here skip the theoretical details that are described in the references.

Segmented DeepDream

Since Style Transfer was melding buildings with the natural environment but not necessarily bringing the cities’ greenery to the forefront, we decided to try applying Deepdream to that part of the cityscapes to make it stand out. We developed a technique very similar to [3] in order to restrict the effects of dreaming to some parts of the photo. In order to produce an image *y* out of an image *x*, we also create a mask *c* of the same size as the image composed of coefficients between 0 and 1. Then, at every step, in addition to minimizing the standard Deepdream loss *f(y)* of the new image, we add a regularization term, weighted by a coefficient α, as follows :

$$l = -f(y) - \alpha \sum_{(i,j)} c_{i,j} \cdot \left| x_{i,j} - y_{i,j} \right|$$

This makes sure that the variations focus on the parts of the image where $c_{i,j}$ is small.

Process

Neural Doodle (Sophia & Julie)

In transforming an image of a city with a nature-forward focus, we began with Neural Doodle. The idea of using Neural Doodle was to take a style image of a city, such as of downtown Pittsburgh in Figure 1. We would then mask all natural and industrial objects in order to produce an image with fewer buildings and more trees, to envision a city with more nature. We first experimented with a bird's eye vision of Pittsburgh and saw that there was too much noise in our photo (such as people and cars). We ran it on a more broad photograph of Pittsburgh to experiment (Figure 1). A challenge that we came across was the abstract nature of the output, even after changing the model parameters to be more aware of semantic borders and style image patches. Neural Doodle works more effectively on paintings than on photographs, as the model is better at producing "impressions" of objects rather than distinctive images and images tend to have more noise.

Artistic Style Transfer (Jenn)

After Neural Doodle, we experimented with Style Transfer with two paintings – *A Sunday Afternoon on the Island of La Grande Jatte* by Georges Seurat and *Looking Forward* by Monet. We experimented with style weights and number of iterations. However, despite multiple tunings, increasing the number of iterations didn't result in much variance, and greater style weight resulted in a blurry photograph (Figure 2).

Masked Style Transfer (Jean-Baptiste)

We thus decided to focus on two points: 1) to only apply style transfer to a part of the image using a mask and 2) to use a realistic image that mostly illustrates texture, since style transfer cannot reproduce precise shapes). We first compared the naive masked transfer and masked style loss, as described before (Figure 3). As was expected, the masked style loss version wass smoother and more visually pleasing. Thus, we decided to focus on this type of loss for following experiments with several pictures you can find in the results. We also tried a lot of different style images you can find on our Github. This brought to light something we have noticed all across our experiments: the choice of the style image highly depends on the starting image, its content, scale, shapes, and colors. The style transfer algorithm has mostly two hyperparameters: the number of iterations, and the ratio between the style loss and content loss. We tried to tune both but found they had relatively little influence on the final output (as shown by Figure 4), and that the initial parameters gave the most satisfactory results.

Segmented DeepDream (Teven)

Our modification to DeepDream worked as expected, with the completely masked ($c_{i,j}$ significant) parts of the image left unchanged, the unmasked ($c_{i,j}$ = 0) parts of image receiving "dreamy" modifications, and the softly-masked ($c_{i,j}$ small but not 0) being sometimes used to complete the start of a pattern that had begun in the unmasked parts. We also used blurred masks instead of the sharp ones we had created for masked style transfer. In order to better control the direction the algorithm was going into, we used a guided loss, as detailed in [5]. In this version of the algorithm, instead of trying to maximize the activations of all neurons in a given layer, we feed the network a "guide" image, then tweak the loss function of DeepDream to maximize only the activations that respond to the guide image; in effect, this asks the network to emphasize details that "remind" it of the guide image. For example, in figure 5.1 of the appendix, we managed to have a bird figure appear instead of the trees ; in figure 5.2, the face of a bear.

Style Transfer and DeepDream Combination (Jenn)

After implementing masked style transfer and segmented deep dream algorithm, we tried combining them. We generated masks that highlighted specific infrastructures to transfer different artworks using mask style transfer. Then, we initially planned to apply a segmented DeepDream algorithm to the sky of the photograph. However, the application of DeepDream to the sky was messy and rather unpleasant. We concluded that each technique brought a different visual touch and message and that using both in a single image would overflow the spectator and make the message less clear. (Figure 6).

Result & Reflection

We settled on two methods of image transformation that utilized our modified loss function – Style Transfer and Deep Dream. Using original photos of cities that we have traveled to around the world, we produced two kinds of images. The first is of cities where the "greenery" is accentuated using Deep Dream, usually on a small scale. The second is of cities where buildings were converted to trees using Style Transfer, on a much bigger scale. This was effective, as DeepDream can develop details, while Style Transfer is better at creating a general texture. Through those methods, we emphasize the coexistence of man-made and natural elements even in the world of men, and draw the audience to reconsider urban spaces.

From this, we chose our final series of images. One of them was on the first page of this report and we attach two more below. More from our series are available in Results (link below).

"The Oasis" - Lumphini Park, Bangkok

"The Promised Land" - Palestinian Side of Jerusalem

Code: https://github.com/TevenLeScao/artXml/

Results: https://github.com/TevenLeScao/artXml/tree/master/results

Appendix

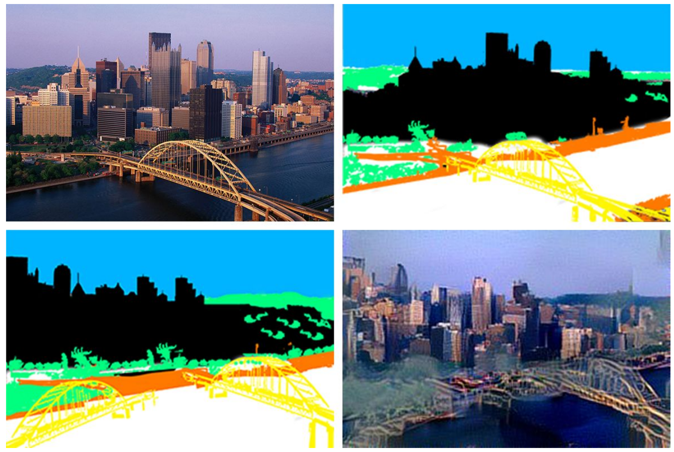
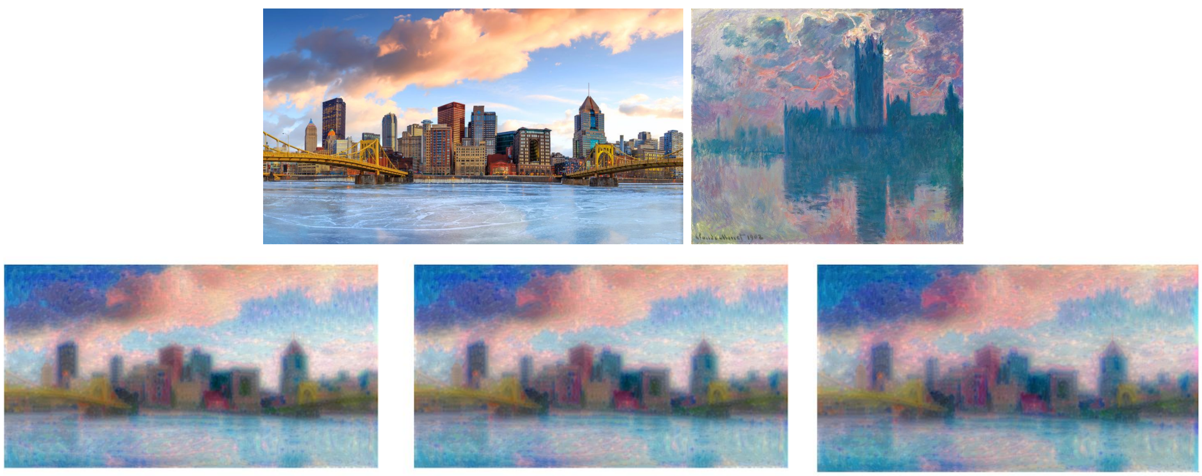


Figure 1: Neural Doodle, Downtown Pittsburgh - *variation weight = 6,000, style weight= 2*



Monet - Iteration 30 Monet - Iteration 50 Monet - Iteration 70

Figure 2: Style Transfer, Monet, Downtown Pittsburgh

Figure 3: Bloomfield, Pittsburgh
Content image, Mask, and outputs of the Naive and Masked Style Loss algorithms (left to right, top to bottom)

Figure 4: Jerusalem, Palestinian side
Outputs with an increasing style loss weight - the third one has the default parameters we kept

Figures 5.1 and 5.2
Outputs of Segmented DeepDream guided by animal photos

Step 1: Bridge Style Transfer | Step 2: Building Naive Style Transfer | Step 2: Building Style Transfer

Figure 6: Masked Style Transfer, Downtown Pittsburgh

References

[1] Vicente Muñoz – Sublimis
http://vicentemunoz.xyz/projects/sublimis
[2] Show, Divide and Neural: Weighted Style Transfer - Ethan Chan, Rishabh Bhargava, 2016
http://cs231n.stanford.edu/reports/2016/pdfs/208_Report.pdf
[3] Localized Style Transfer - Alex Wells, Jeremy Wood, Minna Xiao, 2016
http://cs231n.stanford.edu/reports/2017/pdfs/416.pdf
[4] Neural Transfer Using PyTorch
https://pytorch.org/tutorials/advanced/neural_style_tutorial.html
[5] Convolutional Network Visualizations & Deep Dream
https://www.kaggle.com/carloalbertobarbano/convolutional-network-visualizations-deep-dream
[6] Pre-trained PyTorch Monkeys: A Deep Dream
https://www.kaggle.com/paultimothymooney/pre-trained-pytorch-monkeys-a-deep-dream