Real-Time Image Style Transfer

Zhaoyou Wang, zhaoyou@stanford.edu
Zhaoheng Guo, zhaoheng@stanford.edu
Sophia Lu, sophialu@stanford.edu

Motivation

Image style transfer is a long-standing problem that seeks to transfer the style of a reference style image onto another input picture. Our project has implemented two recent proposed image style transfer algorithms based on convolutional neural networks.

Model

Neural algorithm:

1. Content Loss

$$\mathcal{L}_{\mathrm{c}}^{l}(y_c, y) = \frac{1}{2}\sum_{ij}(F_{ij}^{l}(y_c) - F_{ij}^{l}(y))^2$$

2. Style Loss

$$G_{ij}^{l} = \sum_{k} F_{ik}^{l} F_{jk}^{l}$$

$$E_l(y, y_s) = \frac{1}{4N_l^2 M_l^2}\sum_{ij}(G_{ij}^{l}(y) - G_{ij}^{l}(y_s))^2$$

$$\mathcal{L}_{\mathrm{s}}(y, y_s) = \sum_{l} w_l E_l$$

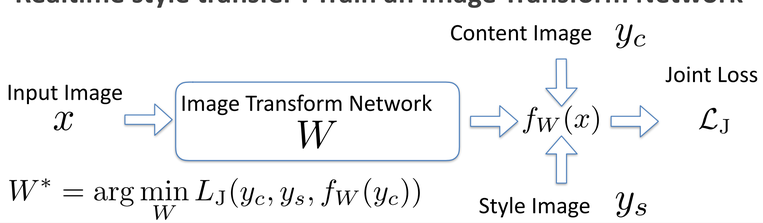
3. Joint Loss Function

$$\mathcal{L}_{\mathrm{J}}(y_c, y_s, y) = \alpha\mathcal{L}_{\mathrm{c}}(y_c, y) + \beta\mathcal{L}_{\mathrm{s}}(y_s, y)$$

4. Output:

$$\hat{y} = \arg\min_{y} \mathcal{L}_{\mathrm{J}}(y_c, y_s, y)$$

Realtime style transfer : Train an Image Transform Network



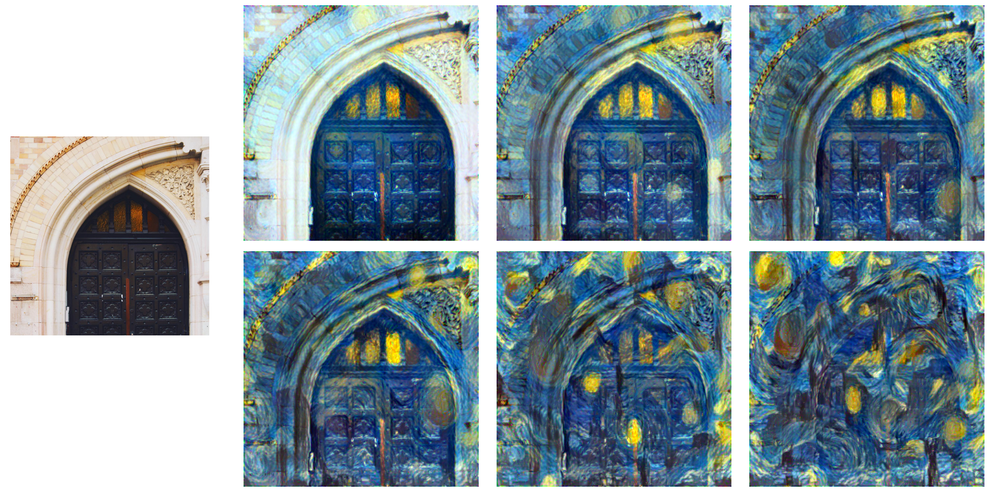
$$W^* = \arg\min_{W} L_{\mathrm{J}}(y_c, y_s, f_W(y_c))$$

Training Details

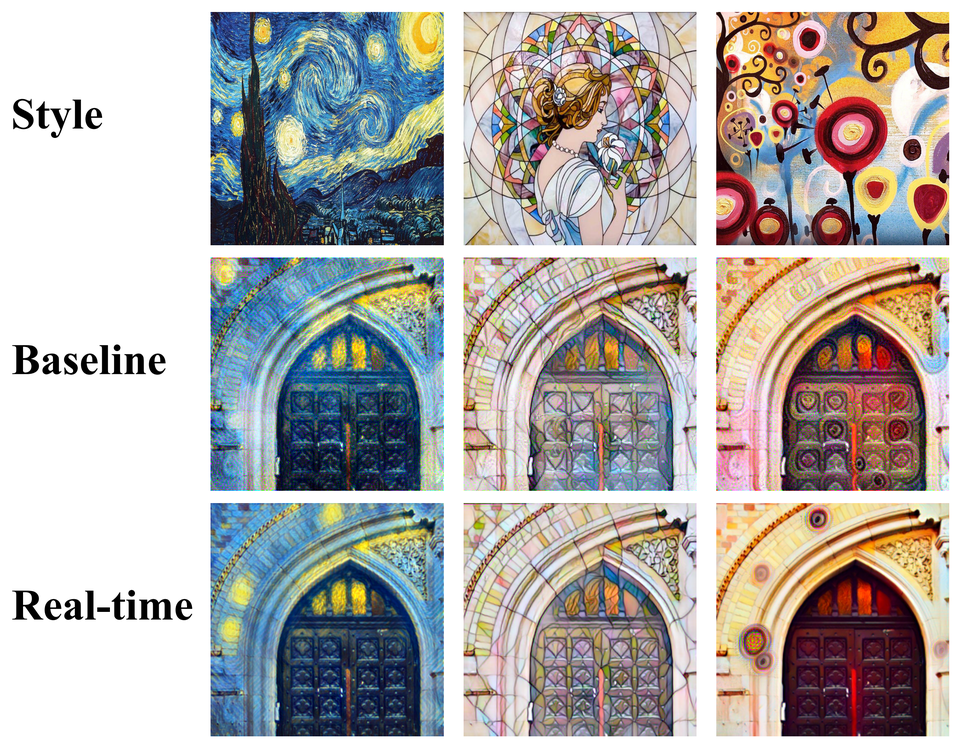
Two epochs over the 80k Microsoft COCO dataset with batch size 4 (resize to 256x256). We use a pre-trained VGG16 network for the loss calculation and an image transformation network with 3 convolution layers, 5 residual layers and 3 deconvolution layers.

Results

Neural algorithm results with different style weights.



Comparison of two methods:



Discussion

Important details for final results:

1. Input images need preprocessing (subtract mean value and normalize) to make loss network work properly.
2. Projecting the values of generated images to range [0, 255] to get rid of some noisy artifacts.
3. During the training process, allowing the image transformation network to work in range [0, 255] generates much better results than than [0, 1].

Future Plans

1. Add photorealistic loss and semantic segmentation to reduce local distortion in original neural algorithm
2. Implement realtime algorithm for photo style transfer
3. Train a style extraction network to realize arbitrary image style transfer

References

1. L. A. Gatys, A. S. Ecker and M. Bethge. A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576, 2015.
2. J. Johnson, A. Alahi and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer, 2016.
3. D. Ulyanov, A. Vedaldi and V. Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.