Final Project CPSC8810

Uma Maheswara R Meleti

Problem Statement: Style Transfer from pixels to parameterized brush strokes

- The pixel domain is unnatural for representing artistic styles.
- Paintings are composed of brushstrokes, not individual pixels
- Shift from pixel-based representations to parameterized brushstrokes, which align better with artistic practices.

Related Work: Style

- Initially, Efros and Freeman performed texture synthesis and transfer using image quilting
- Hertzmann et al. used a pair of images - one being a filtered version of the other - to learn a filter, which can then be applied to a new image. Wang et al. introduced a method for synthesizing directional textures.
- Gatys et al. proposed a method that optimizes pixels to combine content and style by minimizing content and style losses.
- Li et al., Huang and Belongie, and others have improved style transfer using feed-forward networks and arbitrary style transfer.
- Sanakoyeu et al. and Kotovenko et al. introduced style-aware content losses to improve stylization quality.

Related Work: Brush Stroke

- Early works include an interactive method by Haeberli - A program follows the cursor across the canvas, obtains a color by point sampling the source image, and then paints a brush of that color.

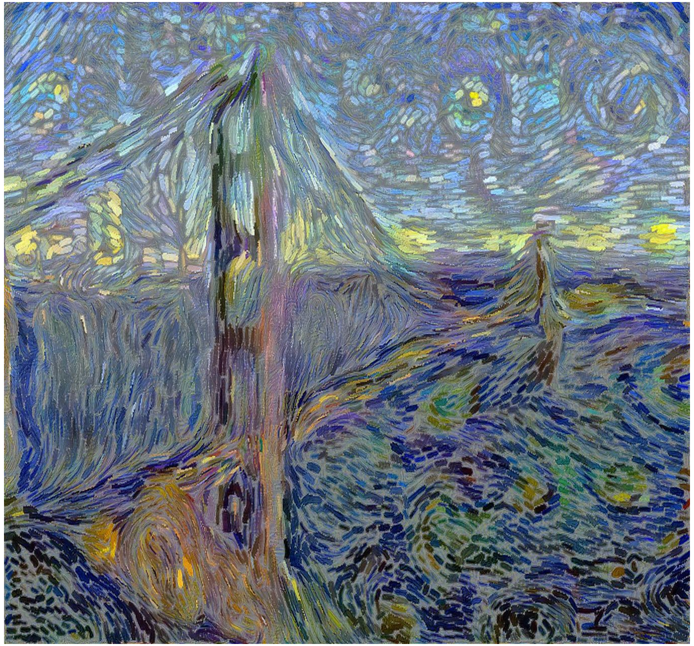
- Hertzmann extended the previous approach using an automated algorithm that takes a source image and a list of brush sizes, and then paints a series of layers, one for each brush size, on a canvas in order to recreate the source image with a hand painted appearance

- Conversely to SBR methods, there have been attempts to detect and extract brush strokes from a given painting. These methods generally utilize edge detection and clustering-based segmentation or other classical computer vision techniques and have been used to analyze paintings.

- Recent work relies on neural networks to predict brush stroke parameters that approximate a given image, using a variety of architectures and training paradigms.

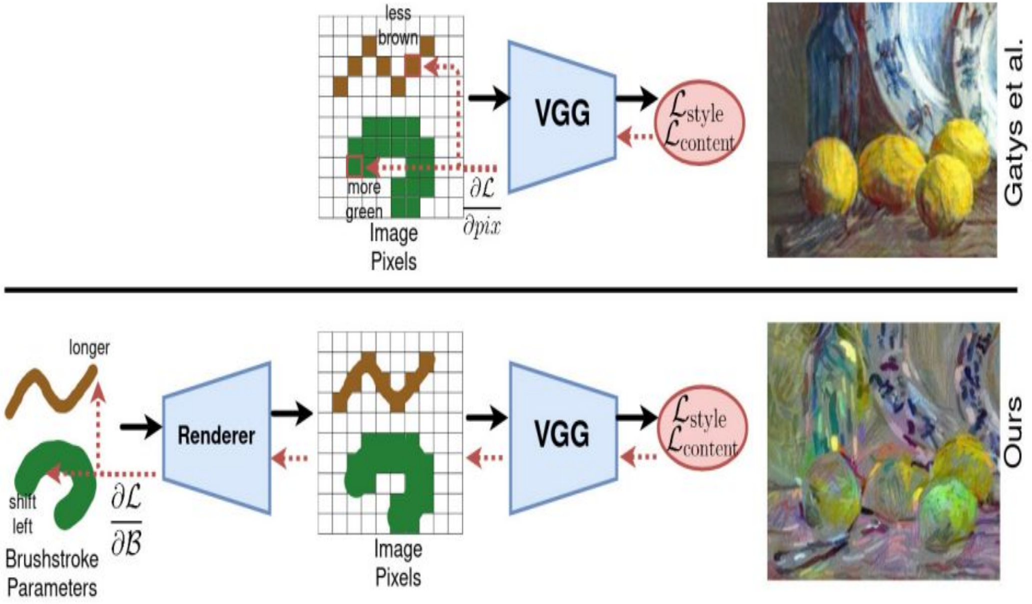
Previous approach Vs Parameterized Brush Strokes

Gaty’s Style Transfer



Current Method

Approach



- Represent image through parameterized brushstrokes, which include parameters for location, color, width, and shape (Bézier curves for shape)
- A Differentiable renderer maps brushstroke parameters into the pixel domain.
- Perform a pixel optimization step after iterative process using Gaty’s algorithm for smoother and natural representation of image.

Differentiable Renderer

- A neural network trained to generate brushstrokes

- The renderer is a differentiable function which transforms a collection of brushstrokes parameterized by location, shape, width and color into pixel values on a canvas.

- In other words, Renderer is a mapping mapping function

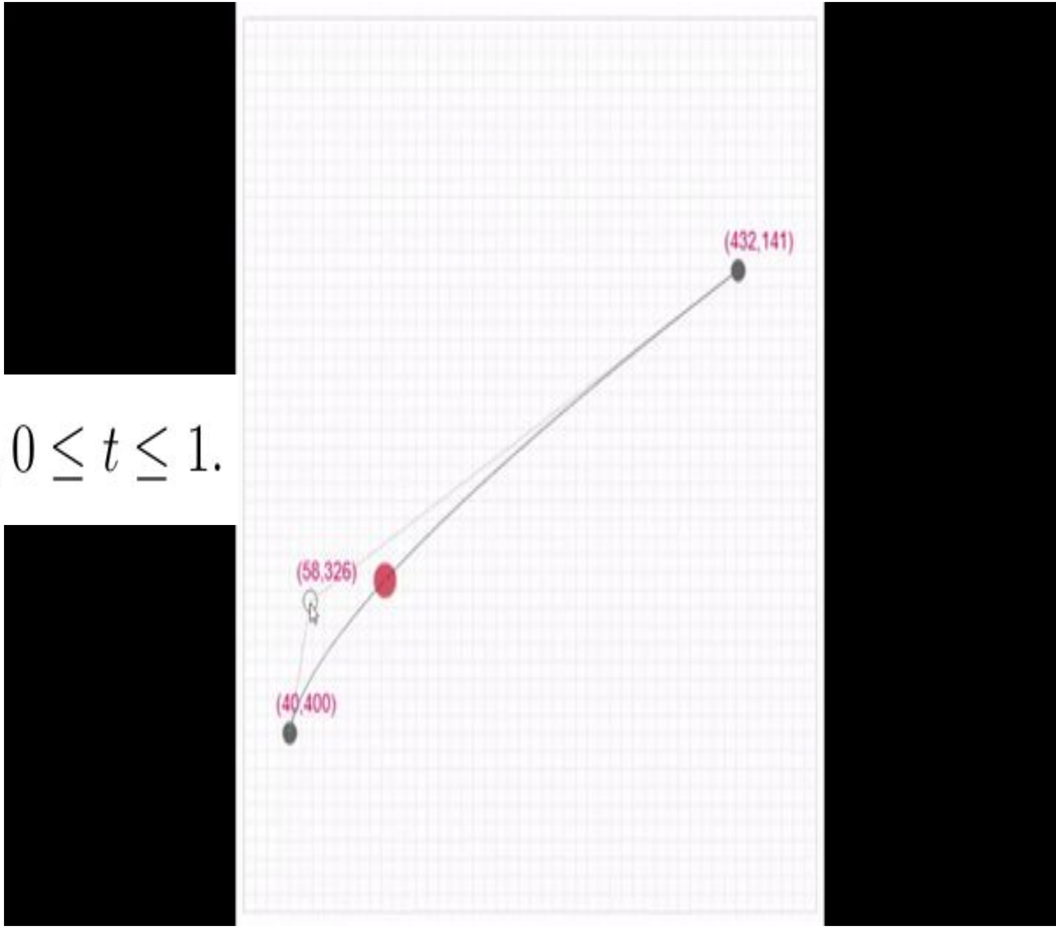
$$\mathcal{R} : \mathbb{R}^{N \times F} \to \mathbb{R}^{H \times W \times 3}$$

- N is number of Brush strokes of dimension F.

- F is 12: Bezier Curve: 6 (Start: 2, End:2, control:2), Location: 2, color: 3, Width: 1

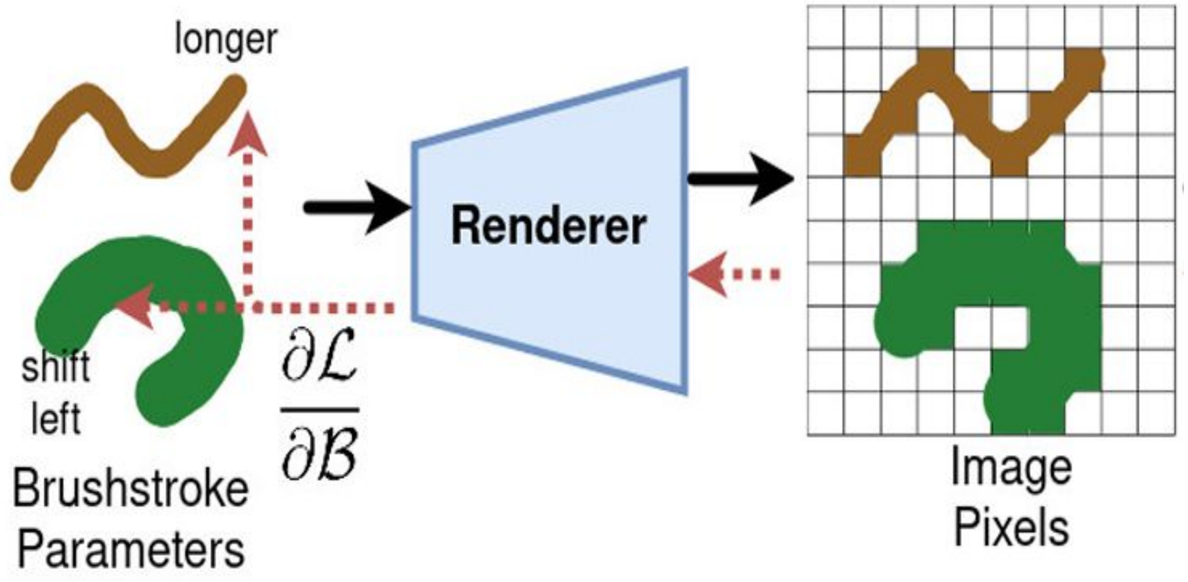
Bezier curve

$$\mathbf{B}(t) = (1-t)^2\mathbf{P}_0 + 2(1-t)t\mathbf{P}_1 + t^2\mathbf{P}_2 \,,\ 0 \le t \le 1.$$



https://ytyt.github.io/siiiimple-bezier/

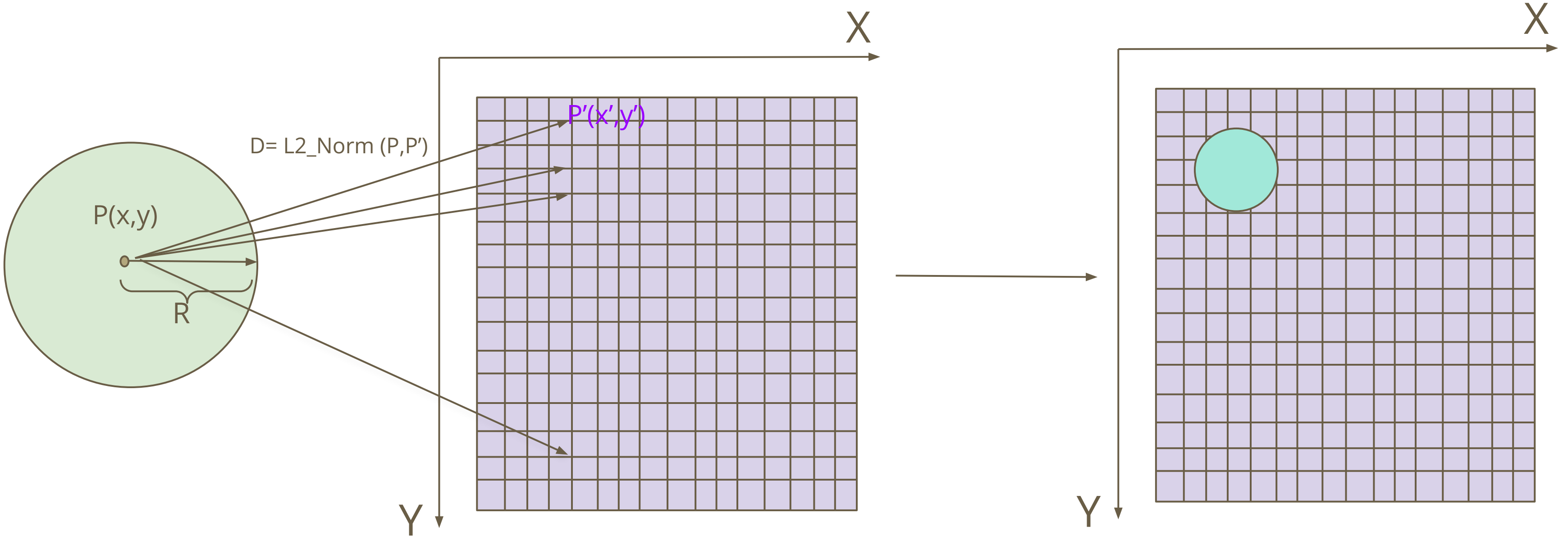
Differentiable Renderer



N x F: 10,000 x curve, location, color, width

HxWx3: RGB Image

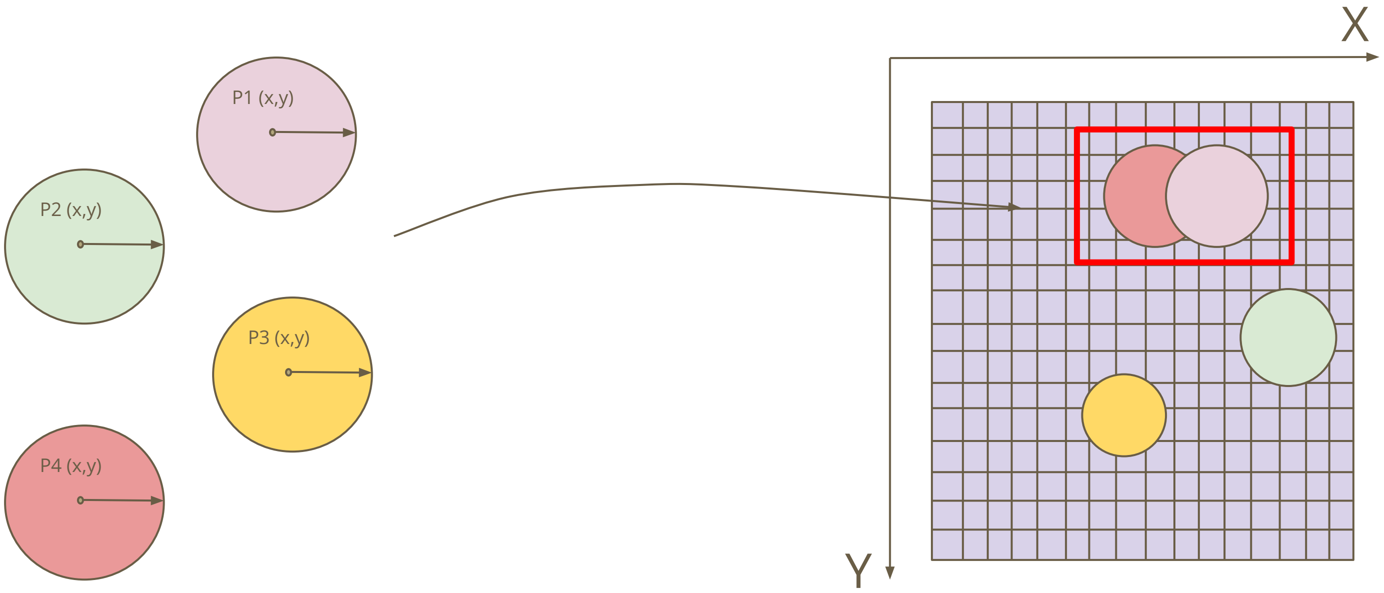
How rendering works: Disk Example



Distance Matrix of canvas
coordinates w.r.t to disc centre

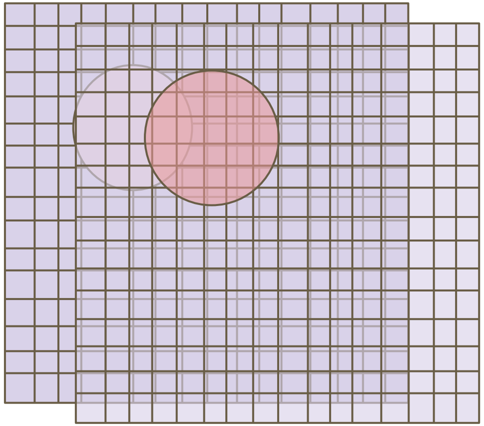
Mask Pixels on the canvas
with given color where D < R

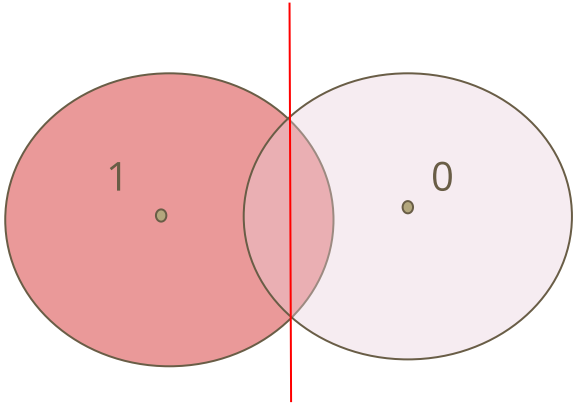
Disk Example: Overlapping Case



Mapping N Discs

Disk Example: Overlapping Case





$$A(i,j,n) := \begin{cases} 1 & \text{if } D_n(i,j) < D_k(i,j)\ \forall k \neq n, \\ 0 & \text{otherwise.} \end{cases}$$

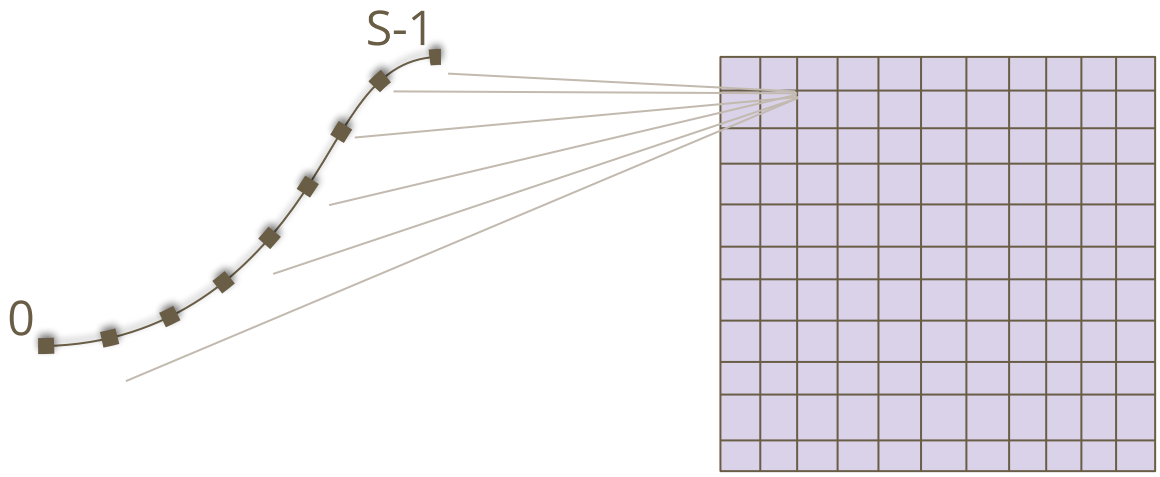
$$A := \{1 \text{ if } D_1 \leq D_2,\ 0 \text{ otherwise}\} \in \mathbb{R}^{H \times W}$$

$$I(i,j) := \sum_{n=1}^{N} I_n(i,j) * A(i,j,n)$$

$$A\colon\ I := I_1 * A + I_2 * (\mathbf{1} - A)$$

Brush Stroke Rendering:

- Mask points that are closer than the brushstroke width and multiply them by a color value.
- Distance from point P on canvas to a Bezier curve is approximated by sampling S equidistant points P’1…..P’s along the curve and computing minimum pairwise distance



Implementation Details

- Computation of assignment matrix and masking operation are both discontinuous.
- Masking operation is implemented with a sigmoid function.
- Assignment matrix is replaced by a softmax operation with high temperature to make it computationally differentiable.
- Computation of distances from a pixel to all the brushstrokes is limited to K nearest brushstrokes to eliminate computational complexity as the brushstrokes affects the nearby areas of canvas.

Loss Functions:

Content Loss: Euclidean distance between rendered image and content image in VGG feature space (layers - conv4_2 and conv5_2).

$$\mathcal{L}_{\text{content}} = ||\phi_l(I_{\text{r}}) - \phi_l(I_{\text{c}})||_2$$

Style Loss: Calculated from the feature maps of pre-trained CNN (layers - conv1_1, conv2_1, conv3_1, conv4_1, and conv5_1)

$$\mathcal{L}_{\text{style}} = \sum_{l=0}^{L} w_l E_l \qquad E_l = \frac{1}{N_l^2 M_l^2} ||G_r^l - G_s^l||_F$$

G's are gram matrices of N feature maps at layer l with an area of M

Experiments: Deception Rate

- A network is trained to classify paintings into artists, deception rate is the fraction of stylized images that the network has assigned to the artist.
- high deception score indicates high similarity to the target image
- A human subject was shown 4 crop outs where he/she was asked to assign the cropout to real artwork or generated art.

Method	Mean deception score ↑	Mean human deception rate ↑
AdaIN [18]	0.08	0.035
WCT [32]	0.11	0.091
Gatys et al. [11]	0.389	0.139
AST [44]	0.451	0.146
Ours	**0.588**	**0.268**
Wikiart test	0.687	-
Photos	0.002	-

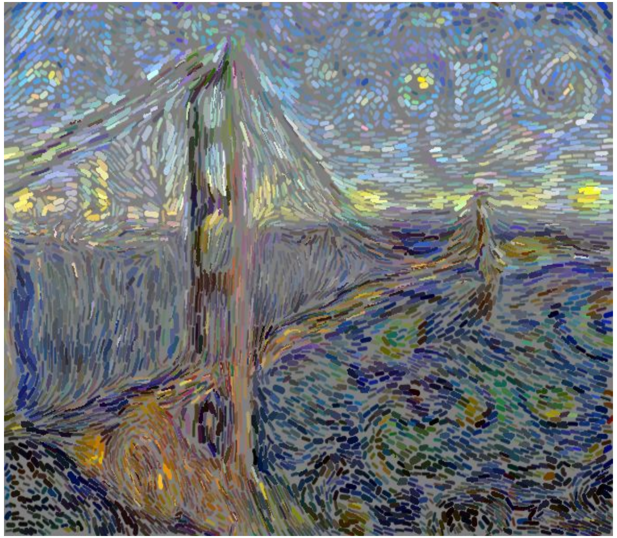
Qualitative Results

Content Image
(Golden Gate Bridge)

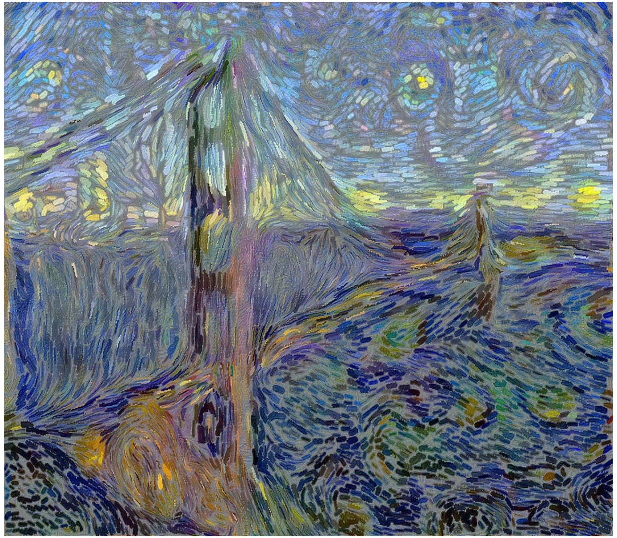
Style Image
van Gogh Starry Night

Qualitative Results

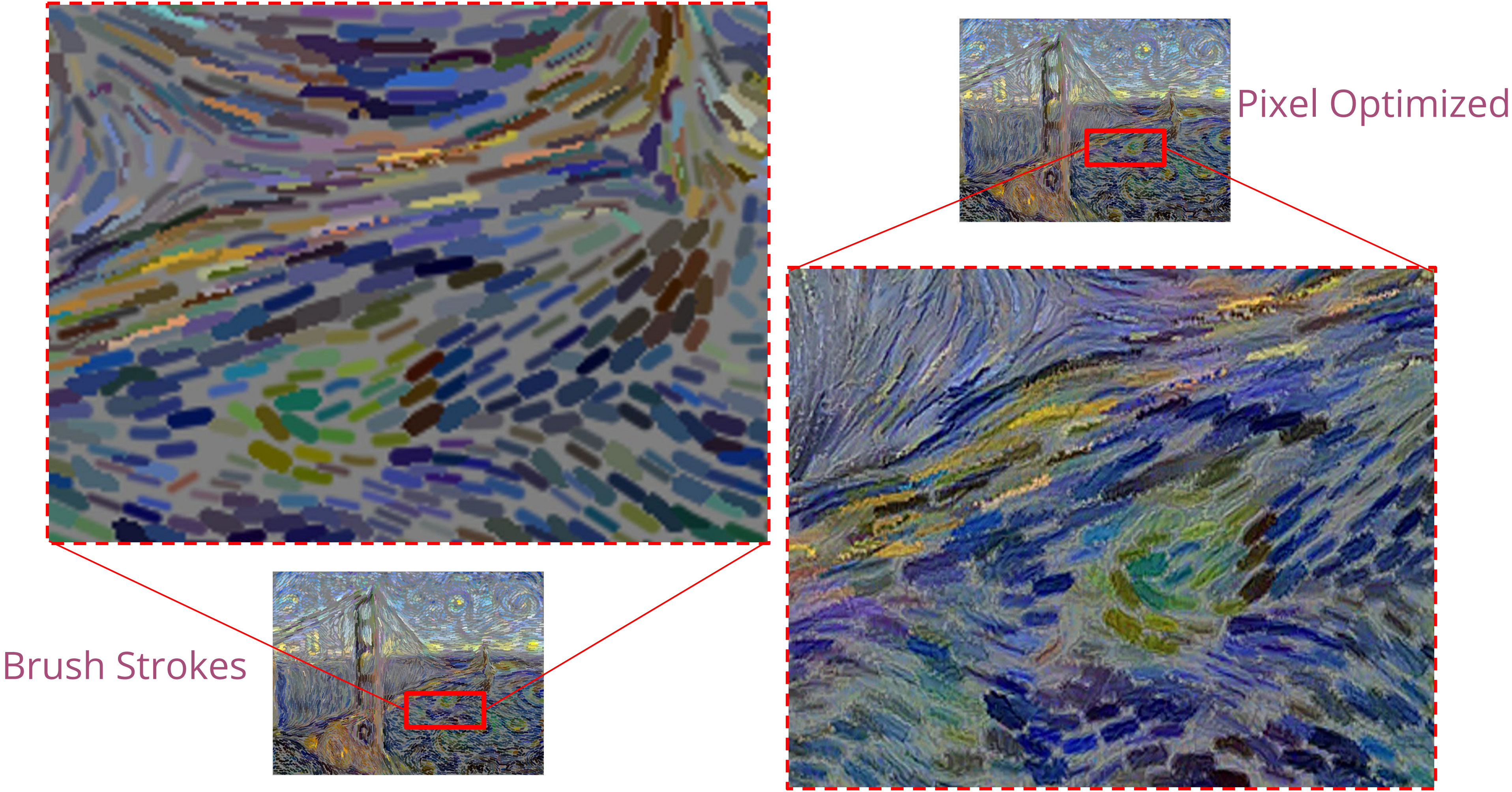
Initial Brush Strokes (using SLIC)

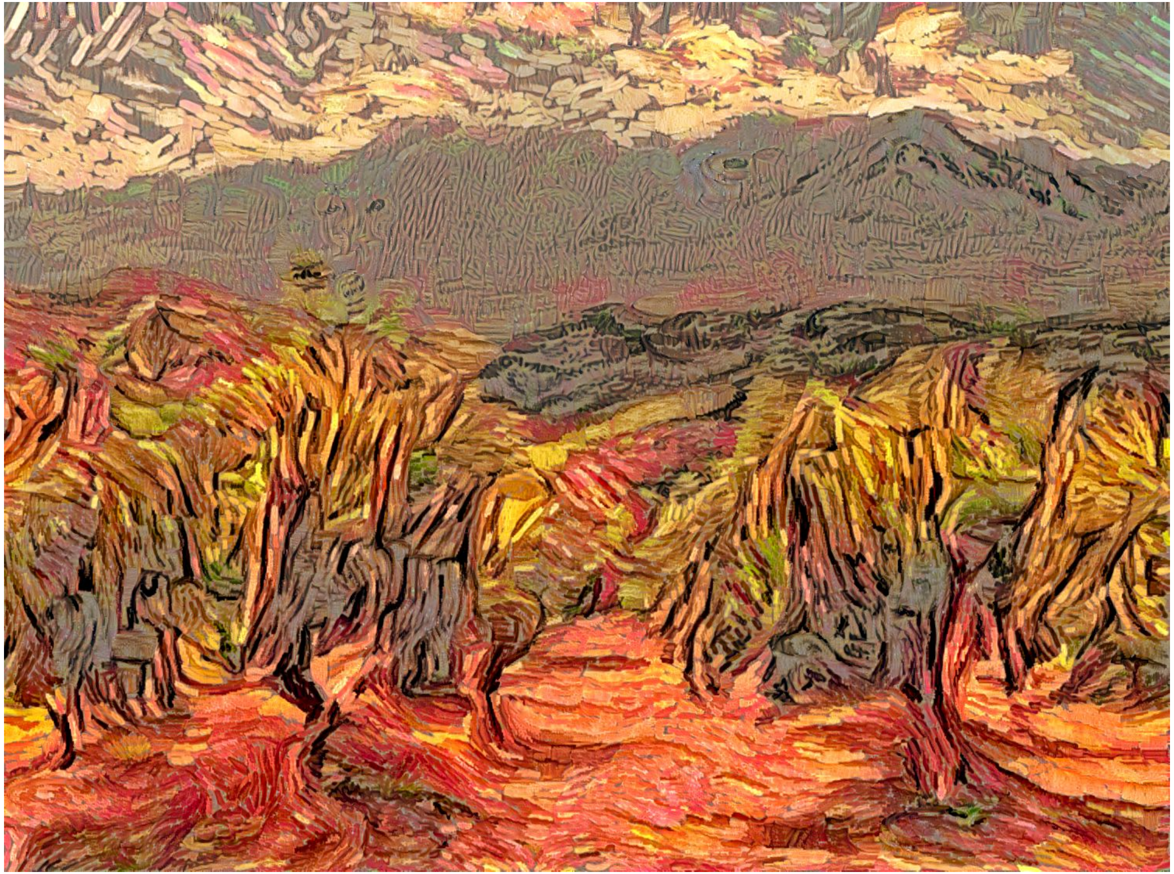


After Rendering



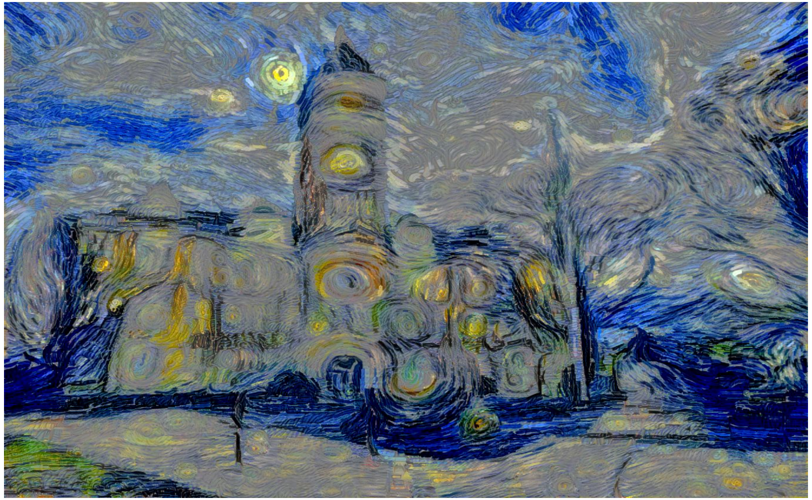
After Pixel Optimization


Pixel Optimized
Brush Strokes



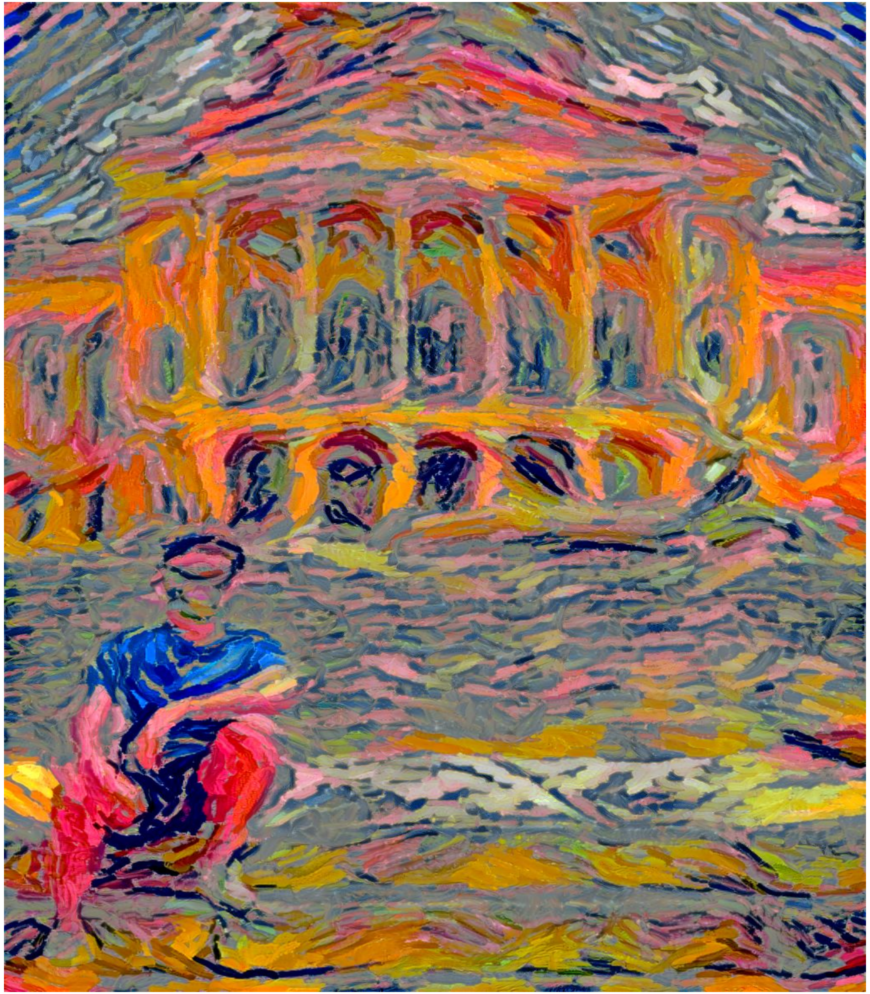
The Stone Bench in Garden of Saint-Paul Hospital - Van Gogh

Clemson University




Me - Iowa State Capitol


Lady On Diwan - Kirchner, Ernst Ludwig



Conclusion and Future work

- Switching to representation for style transfer from pixels to parameterized brush strokes is more natural for artistic style transfer.

- An explicit rendering mechanism is implemented that can be applied beyond style transfer

- The current is best for artistic style where brushstrokes are clearly visible. More sophisticated brush stroke blending procedures can be investigated in future.

- More detailed representation of deep features in content image can be explored.

Thank You

Experiments

- Reproducing the paper in pytorch framework

- Evaluation
    - a network trained to classify paintings into artists. The deception rate is the fraction of stylized images that the network has assigned to the artist, whose artwork has been used for stylization.