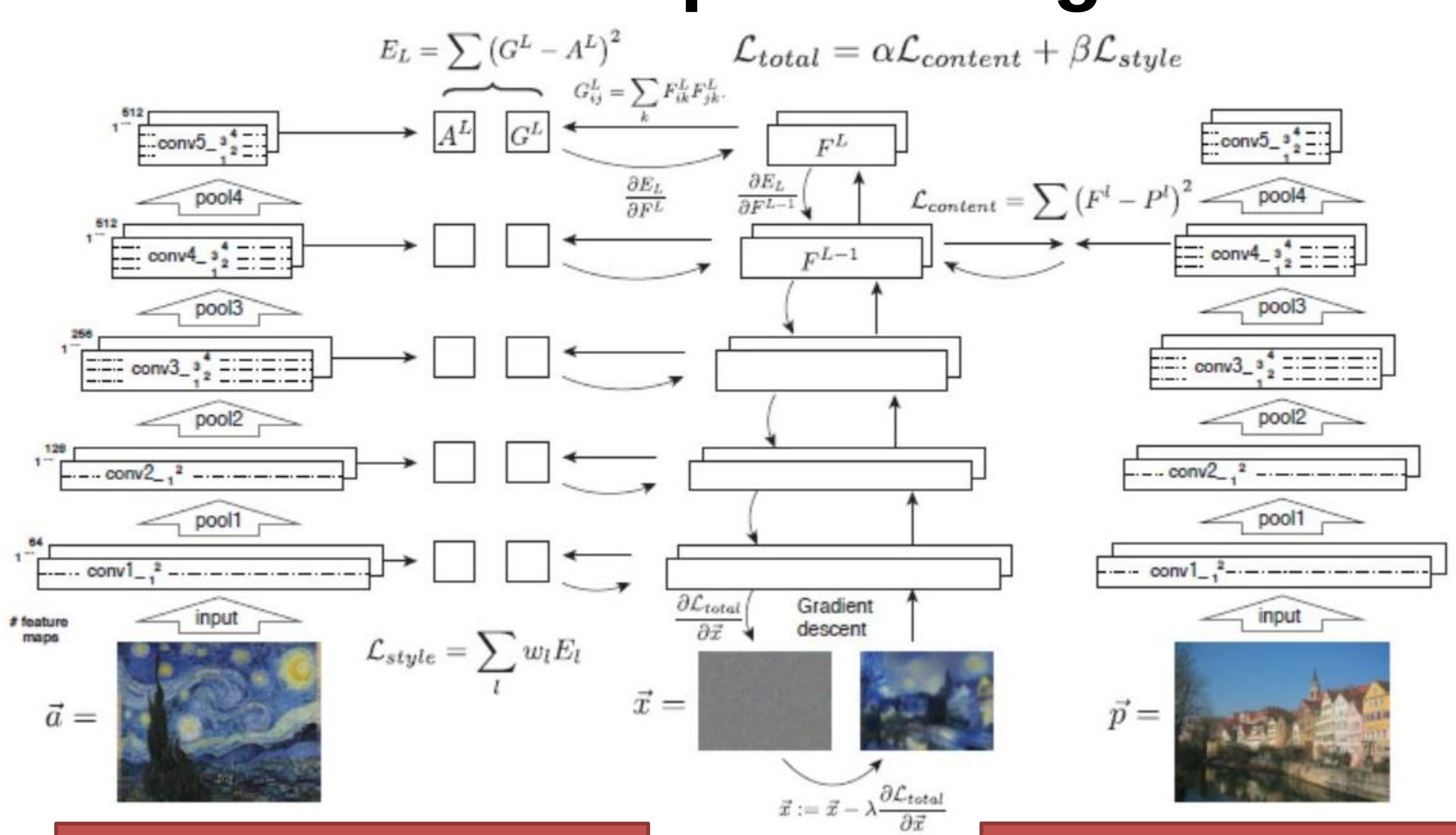
Artistic Style Transfer for Face Portraits

Marcus Pan, Chen Zhu, Daniel Hsu
{mpanj, chen0908, dwhsu}@stanford.edu

Motivation

Our project performs style transfers on a portraits by analyzing content and style in parallel. We use CNNs to learn the content features such as the eyes and nose from the portrait, and style features such as brush strokes from a target style image. Combining these results using gradient descent provides us with the final style transferred portrait.

Deep Learning Model[1]



Platform: MatConvNet

$\vec{x}$:generated image
$\vec{p}$:content image
$\vec{a}$:style image

Feature Space: Normalized 19-layer VGG Network

Gram Matrix:

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

Content Representation

N^l: feature maps of content image at layer l
F^l: feature maps of generated image at layer l

Style Representation

A^l: gram matrix of the style image at layer l
G^l: gram matrix of the generated image at layer l

Loss function

$$L_{content}(\vec{p},\vec{x},l) = \frac{1}{2}\sum_{i,j}\left(F_{ij}^l - P_{ij}^l\right)^2 \qquad L_{style}(\vec{a},\vec{x}) = \sum_{l=0}^{L}\frac{w_l}{4N_l^2 M_l^2}\sum_{i,j}\left(G_{ij}^l - A_{ij}^l\right)^2$$

L-BFGS gradient descent

$$\vec{x} \coloneqq \vec{x} - \lambda\frac{\partial(L_{content} + L_{style})}{\partial \vec{x}}$$

Regularization

We found many output pixel values were outside the (0, 255) range. The following per-pixel loss functions [2] regularize the pixel size and variation:

$$E(x_{i,j})_{size} = x_{i,j}^2 \qquad E(x_{i,j})_{variation} = (x_{i,j} - x_{i+1,j})^2 + (x_{i,j} - x_{i,j+1})^2$$

Reference

[1] Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576.
[2] A. Mahendran and A. Vevaldi (2014). Understanding Deep Image Representations by Inverting Them [Online]. Available: https://arxiv.org/abs/1412.0035

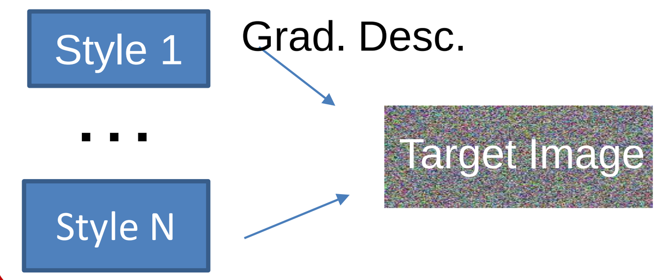
Discussion

We conclude that our style transfer algorithm is able to operate on multiple style images, and use both gradient descent and eigenface techniques. Furthermore, our project produces visually appealing images that keep both the content and style consistent with the original images.

Multiple Style Transfer

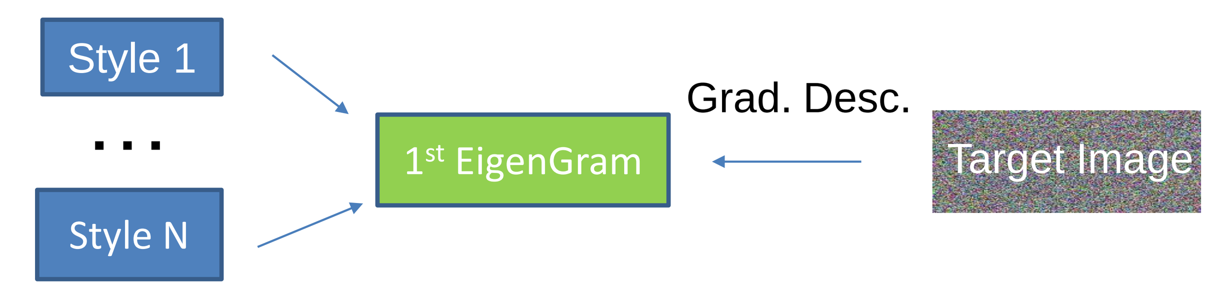
Method 1

Match target image gram matrix with combination of gram matrices of style images:



Method 2

Match target image with 1st eigen-gram of style images (more efficient):



Results and Discussion

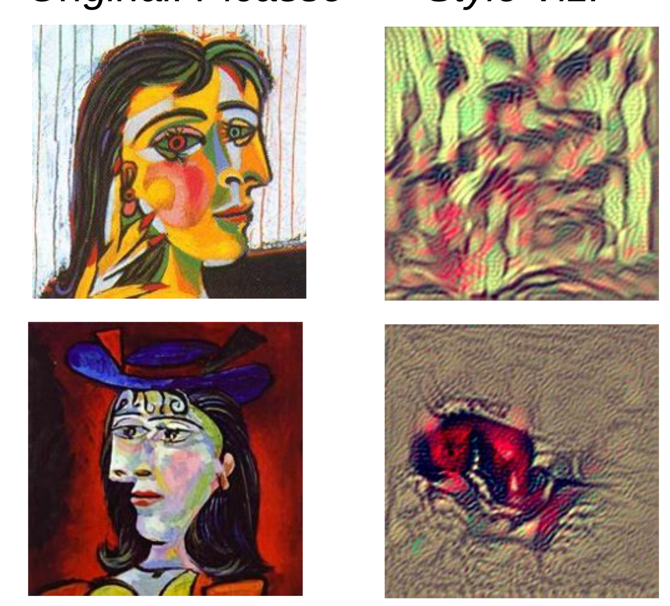
Content: The content representation is visualized below. The deep layer representation preserves the general structure but not specific pixel color.

Original: Chow Yun Fat

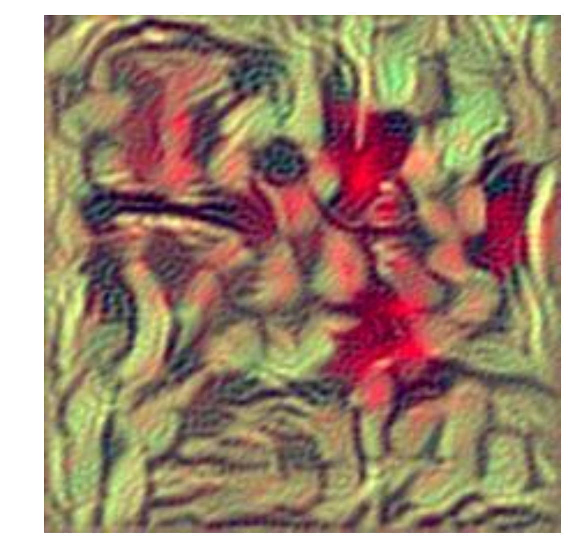
Output Layer 27

Style: The following compares the different types of average style.

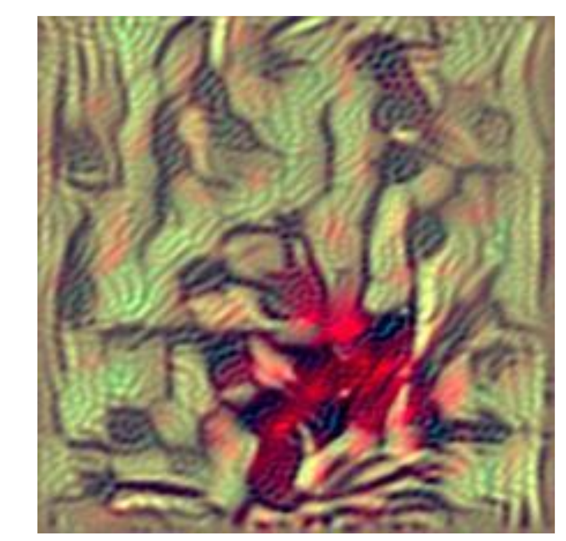
Original: Picasso *Style Viz.*



Grad. Desc Average



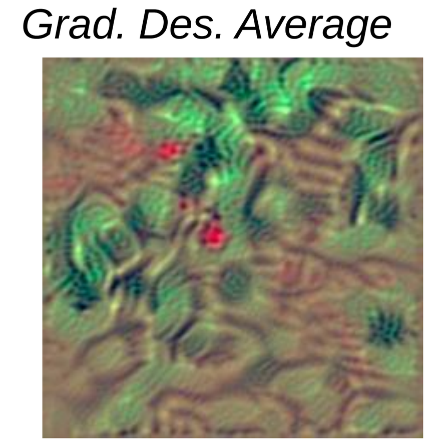
1st Eigen-gram Average



Combination: We combined the content picture with an average style of an artist

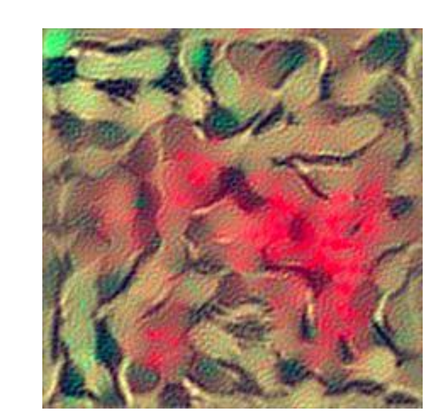
Style Images: Van Gough (Post Impressionist)

Grad. Des. Average

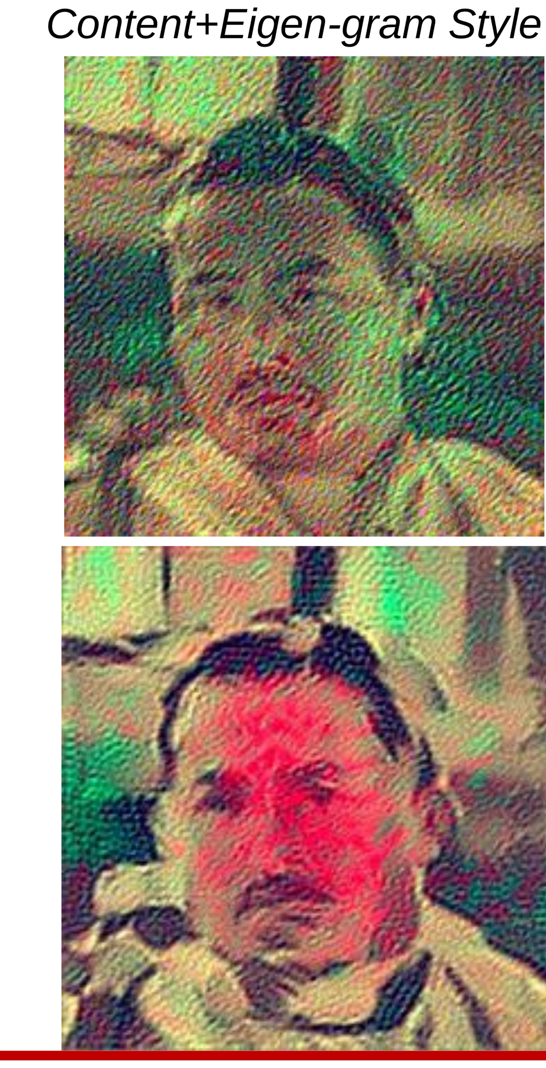


1st Eigen-gram

Style Images: Andy Warhol (Pop Art)



Content+Eigen-gram Style



Future work

Potential future projects include different style transfers on a single image, such transferring the face to Picasso, while changing the background to Warhol. We would also like to explore real time style transfer for augmented reality applications.