Neural Style Transfer
Mid-Term Report

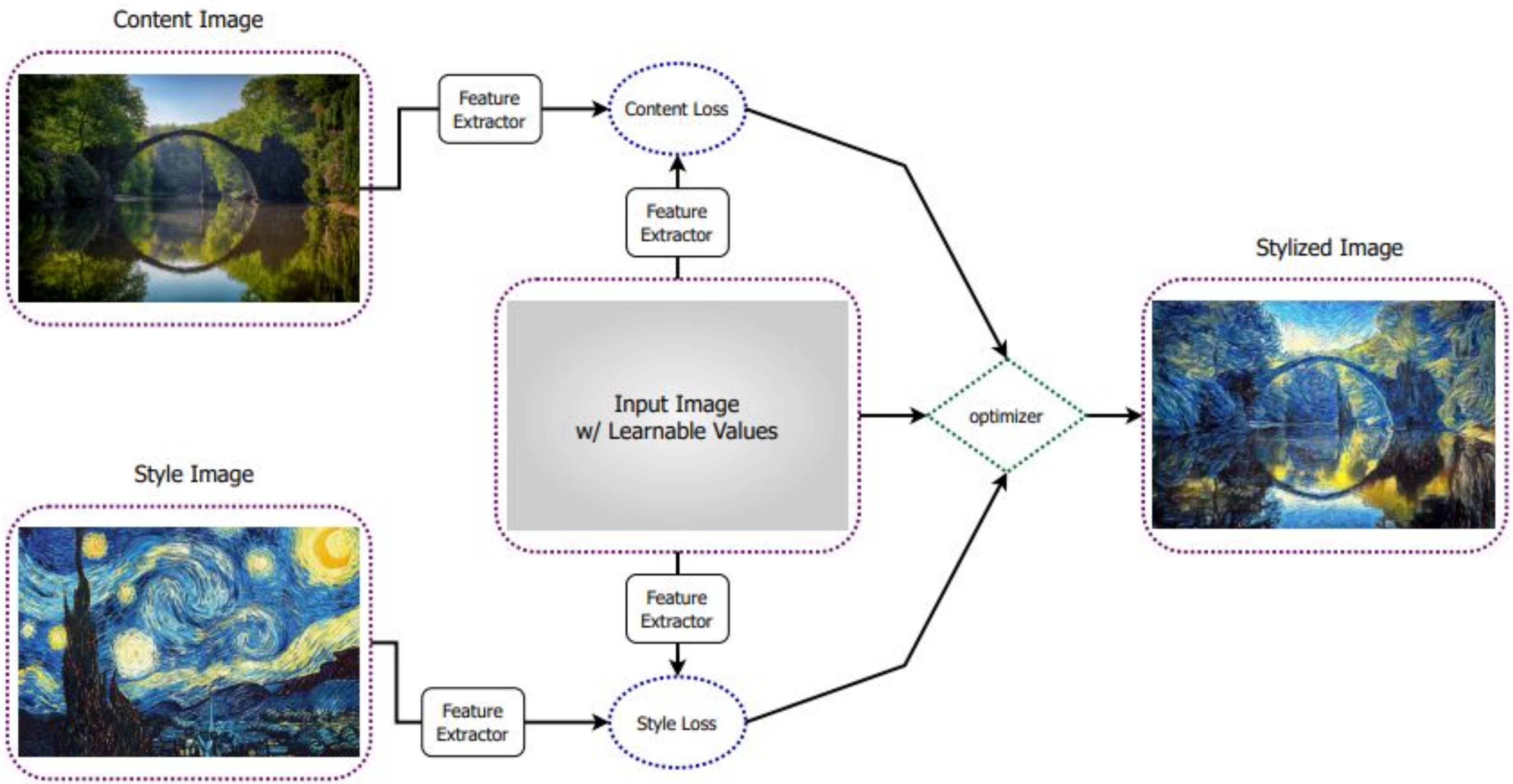
CSN-526: ML Project

Group-1:
Team Members

- Jitesh Jain 19114039
- Gagan Sharma 19114032
- Divyansh Agarwal 19115055
- R Chinmay 19114067
- Md Junaid Mahmood 19116040

Introduction

- **Neural Style Transfer** is one of the earliest application of deep learning in the field of creative art.
- Our method is mainly inspired by the method first proposed in the paper: **A Neural Algorithm of Artistic Style,** by Gatys et al., 2015.
- Neural Style Transfer is an optimization problem, which takes in 3 images as input: content image, style image and a learnable input image.
- The goal is to learn the input image such that the final result consists of the **content of the Content image and style of the Style image.**



$$L_{total} = \lambda_{content} L_{content} + \lambda_{style} L_{style}$$

where $\lambda_{content}$ and λ_{style} are hyperparameters.

We set $\lambda_{content} = 1 \times e^{0}$ and $\lambda_{style} = 1 \times e^{6}$ for all our experiments.

Implementation and Evaluation

- Implemented in PyTorch, using publicly available models on the torchvision model zoo.
- For training and evaluating our methods, we choose **5 pairs of content and style images** from the web.
- Since there exists no real ground truth for the stylized image, we use the results from the Adaptive Instance Normalization based method proposed by Huang et al. as **pseudo ground truths**.
- **Metrics** used to compare our results with pseudo GTs:
    - Root Mean Square Error (**RMSE**)
    - Peak Signal-to-Noise Ratio (**PSNR**)
    - Structural Similarity Index Measure (**SSIM**) metrics

Training Data Pairs:

We use 5 pairs of content and style images taken from the web for all our experiments.

Ablation on Feature Extractors

- We experiment with three different feature extractors - **VGG-16, VGG-19 and ResNet-18**.
- We observe that **ResNet-18** gives the best performance. Thus, we use ResNet-18 as our feature extractor for further experiments.

Metric	VGG-16	VGG-19	ResNet-18
RMSE	0.01213	0.01231	**0.1187**
PSNR	38.336	38.211	**38.528**
SSIM	0.8798	0.8767	**0.8864**

Ablation on Optimizers

- We experiment with three different optimizers - **LBFGS, Adam and AdamW.**
  - In order to find the best learning rate (lr) in case of Adam and AdamW, we tune the lr hyper-parameter on Pair-2.
  - We find lr = 1e −1 and lr = 1e −2 to be the best setting for Adam and AdamW respectively.
- We find **LBFGS** to show the best performance and hence, use LBFGS for further experiments.

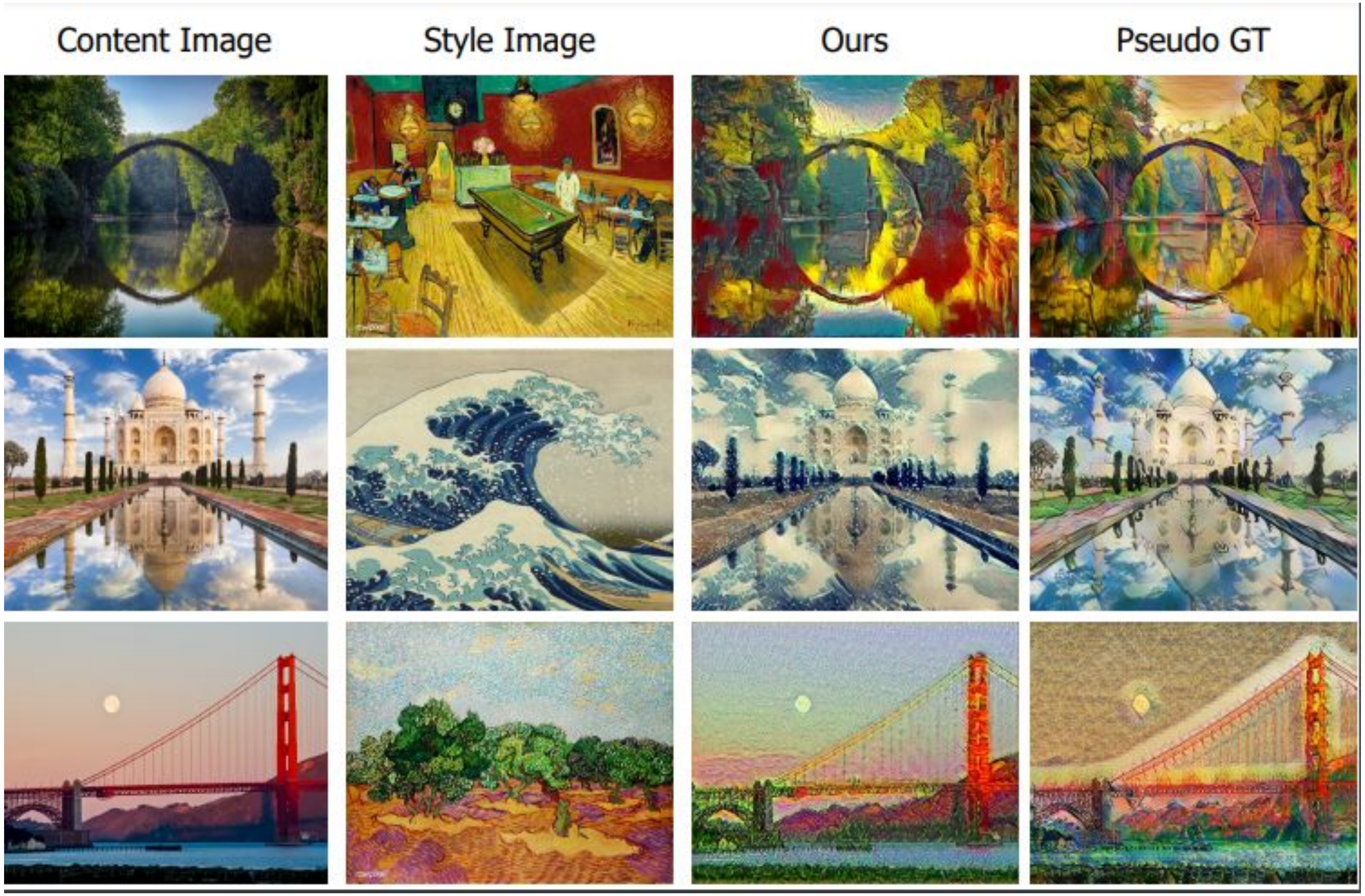
Metric	LBFGS	Adam	AdamW
RMSE	**0.011875678**	0.012709627	0.0123813321813
PSNR	**38.52880338**	38.00075105	38.235296284465
SSIM	**0.886413618**	0.878492873	0.8800718183334

Ablation on Number of Iterations

- We experiment with three different values of number of iterations - **1000, 1500 and 2000.**
- We find that **1000 iterations** gives the best performance. Thus, we set the number of iterations to 1000 for further experiments.

Metric	**1000**	**1500**	**2000**
RMSE	**0.01187**	0.01191	0.01191
PSNR	**38.528**	38.504	38.501
SSIM	**0.8864**	0.8859	0.8860

Results



Future Work

- We plan to work on more ablation experiments to study the effects of initialization techniques of the input image, initializing it with content image (default), or style image or any random image.
- We plan to include tv-loss to the total loss term.
- We may try more feature extractors like ResNet-34.

References

- Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. **A neural algorithm of artistic style**. arXiv, 2015. 2 4
- [2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. **Deep residual learning for image recognition**. In 2016, 2016. 1, 3, 4
- Xun Huang and Serge Belongie. **Arbitrary style transfer in real-time with adaptive instance normalization**. In ICCV, 2017. 2, 3
- Diederik P. Kingma and Jimmy Ba. **Adam: A method for stochastic optimization**. In ICLR, 2015. 3
- Shaohua Li, Xinxing Xu, Liqiang Nie, and Tat-Seng Chua. **Laplacian-steered neural style transfer**. In ACMMM, 2017. 4
- Ilya Loshchilov and Frank Hutter. **Decoupled weight decay regularization**. In ICLR, 2019. 3
- Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. **Pytorch: An imperative style, high-performance deep learning library**. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, Advances in Neural Information Processing Systems 32, pages 8024–8035. Curran Associates, Inc., 2019. 3
- Karen Simonyan and Andrew Zisserman. **Very deep convolutional networks for large-scale image recognition**. In ICLR, 2015.

THANK YOU