111062240 李晉方

Neural Style Transfer (NST) Experiment Report

Introduction

Neural Style Transfer (NST), introduced by Gatys et al., combines the content from one image with the style of another to create a unique artistic image. This experiment explores NST using deep neural networks to separate and recombine the semantic content and stylistic features of two images.

Experimental Methods and Tools

- **Python**: as the programming language.
- **PyTorch**: Deep learning framework
- **VGG-19 pretrained network**: identify content and stylefeatures from the images.
- **Adam Optimizer**: Optimization algorithm.
- **PIL & torchvision**: Image handling and preprocessing.

Code overview

You can see the code snippet in my Hackmd: https://hackmd.io/@jacinto5940304/NST

There are some experimental methods:

1. **Image Preprocessing**:
    - Images were resized to match the original dimensions of the content image and normalized according to ImageNet standards.
2. **Feature Extraction**:
    - Extracted content and style features using selected layers from **VGG-19**.
3. **Loss Computation**:
    - **Content loss**: Mean squared error between feature maps of the content and generated images.
    - **Style loss**: Mean squared error between Gram matrices of feature maps from style and generated images.
4. **Image Generation**:
    - Initialized the generated image from the content image and iteratively updated it (5000 steps) to minimize combined losses.
5. **Result Visualization**:
    - Generated images saved after periodic iterations to observe progressive changes.

Something about Math

$$\mathcal{L}_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2} \sum_{i,j} \left(F_{ij}^{l} - P_{ij}^{l} \right)^2$$

- Content loss

$$\frac{\partial \mathcal{L}_{\text{content}}}{\partial F_{ij}^{l}} = \begin{cases} \left(F^{l} - P^{l} \right)_{ij} & \text{if } F_{ij}^{l} > 0 \\ 0 & \text{if } F_{ij}^{l} < 0 \end{cases}$$

- The derivative of this loss with respect to the activations in layer i

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

- Gram matric

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} \left(G_{ij}^{l} - A_{ij}^{l} \right)^2$$

$$\mathcal{L}_{\text{style}}(\vec{a}, \vec{x}) = \sum_{l=0}^{L} w_l E_l$$

- Total style loss

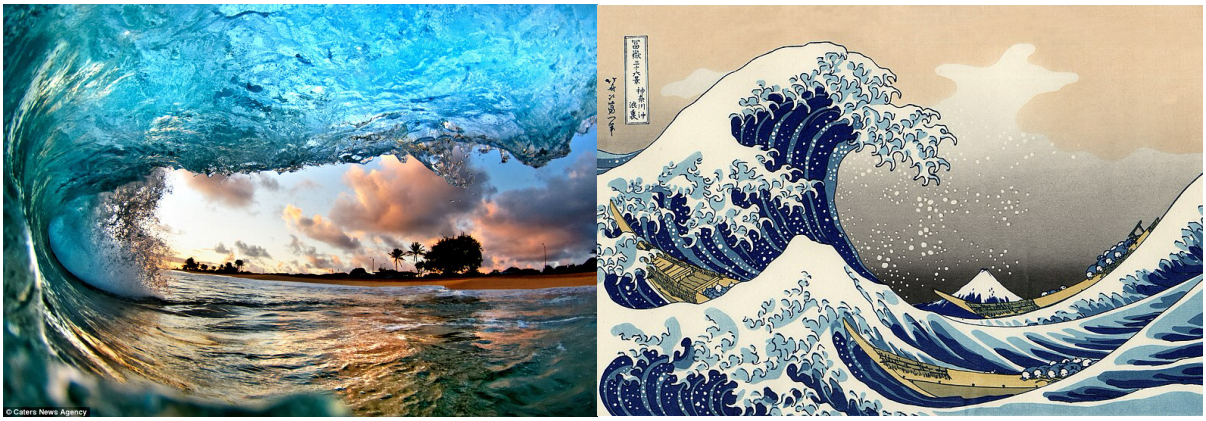
$$\mathcal{L}_{\text{total}}(\vec{p}, \vec{a}, \vec{x}) = \alpha \mathcal{L}_{\text{content}}(\vec{p}, \vec{x}) + \beta \mathcal{L}_{\text{style}}(\vec{a}, \vec{x})$$

- Content loss plus style loss with different weighting factors, α and β

In the paper, "A Neural Algorithm of Artistic Style." The author introduces the formulas above. We culculte the content loss & style loss, and minimise total loss. Owing to length constraints, you can check the Loss function implementation in my code in Hackmd.

Results

The generated NST image successfully combined the realistic content structure of the original photo (an ocean wave) with artistic features and textures from the chosen style image (the famous artwork "神奈川沖浪裏, The Great Wave off Kanagawa").



Content picture(left one) and style picture(right one)

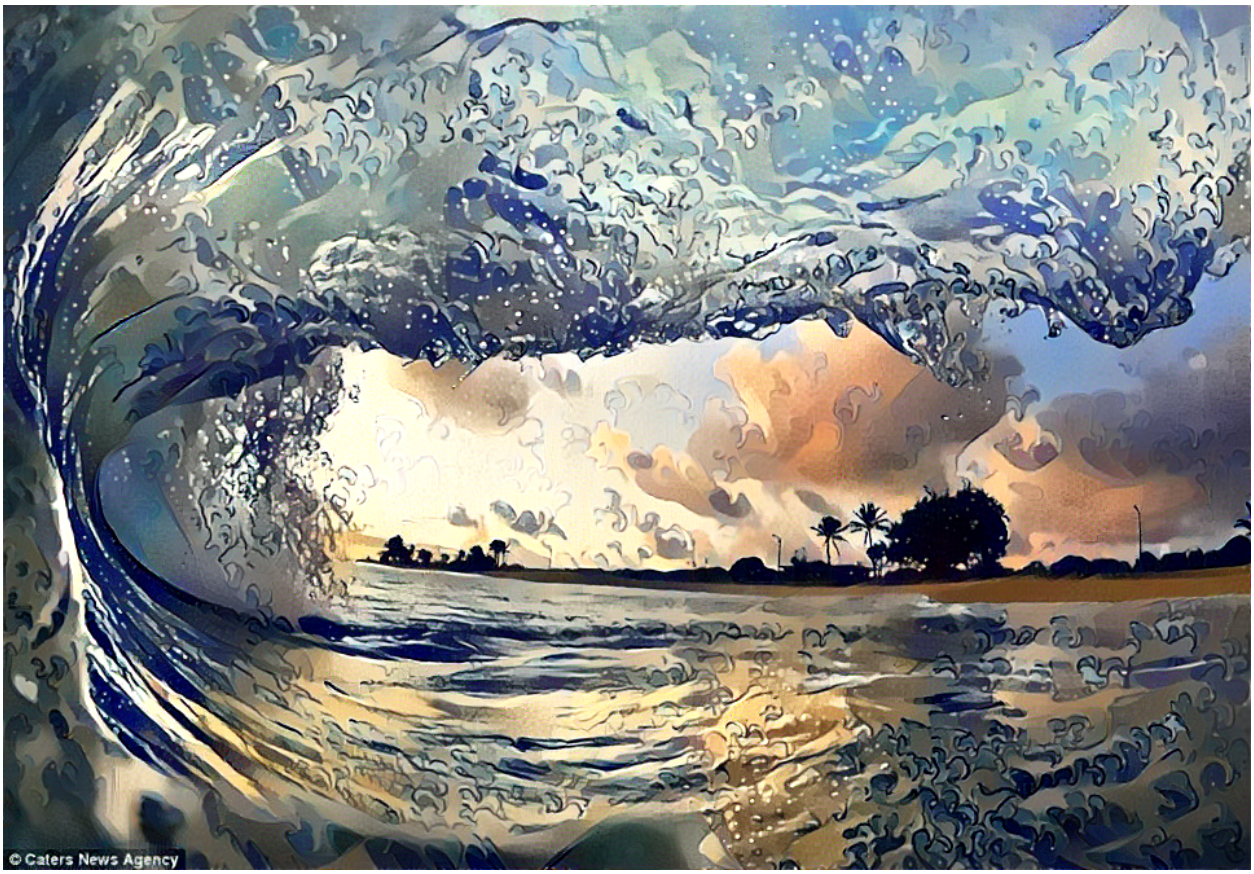
Iteration 1000

Iteration 2000

Iteration 3000

Iteration 4000

Training process



Final result

Conclusions

Through this experiment, I have clearly understood the concepts of style loss and content loss. Style loss captures the texture and artistic style through Gram matrices, while content loss ensures the preservation of essential structures and details of the original content image.

I also became familiar with the practical operations of training neural networks, such as iterative optimization, tuning parameters like alpha and beta, and evaluating intermediate results.

Overall, this experiment has strengthened my understanding of NST and deep learning techniques, enhancing my capability to utilize neural networks creatively.

Reference

[1] Leon A. Gatys. Image Style Transfer Using Convolutional Neural Networks.

[2] Aldo Ferlatti. Neural Style Transfer (NST) — theory and implementation 取自 https://medium.com/@ferlatti.aldo/neural-style-transfer-nst-theory-and-implementation-c26728cf969d