

---

---

# Neural Style Transfer

— Analysis and Improvement —

---

---

FAI CS5100  
Project

# The Intent



Content Image



Reference Image



Generated Image

# Starting point



Content Image (C)



Style Image (S)



Generated Image (G)

# Concept

## Loss function:

$$J_{\text{total}}(G) = \alpha \cdot J_{\text{content}}(C, G) + \beta \cdot J_{\text{style}}(S, G)$$

$$J_{\text{content}}(C, G, l) = \frac{1}{2} \sum (F_{ij}^l - P_{ij}^l)^2$$

$$J_{\text{style}}(S, G) = \sum w_l E_l$$

$$E_l = 1/(4N_l^2 M_l^2) \sum (G_{ij}^l - A_{ij}^l)$$

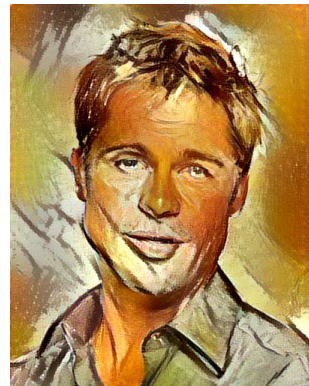
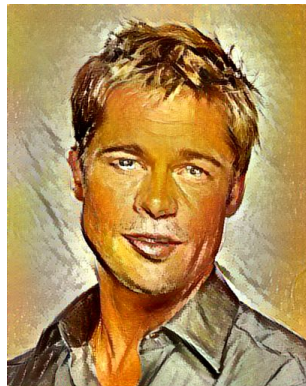
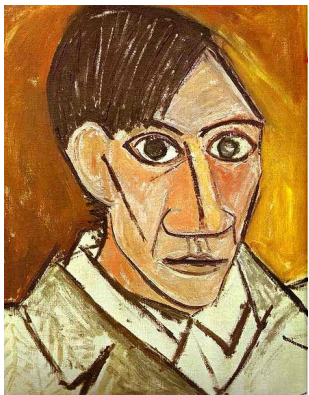
$$G_{ij}^l = \sum (F_{ik}^l F_{jk}^l)$$

## Algorithm:

1. Initialise G randomly  
(here, G initialised to content image)
2. Use **gradient descent** to **minimise J(G)**  
 $G = G - d(J(G)) / (dG)$   
=> Updating pixel value



# Content-style tradeoff



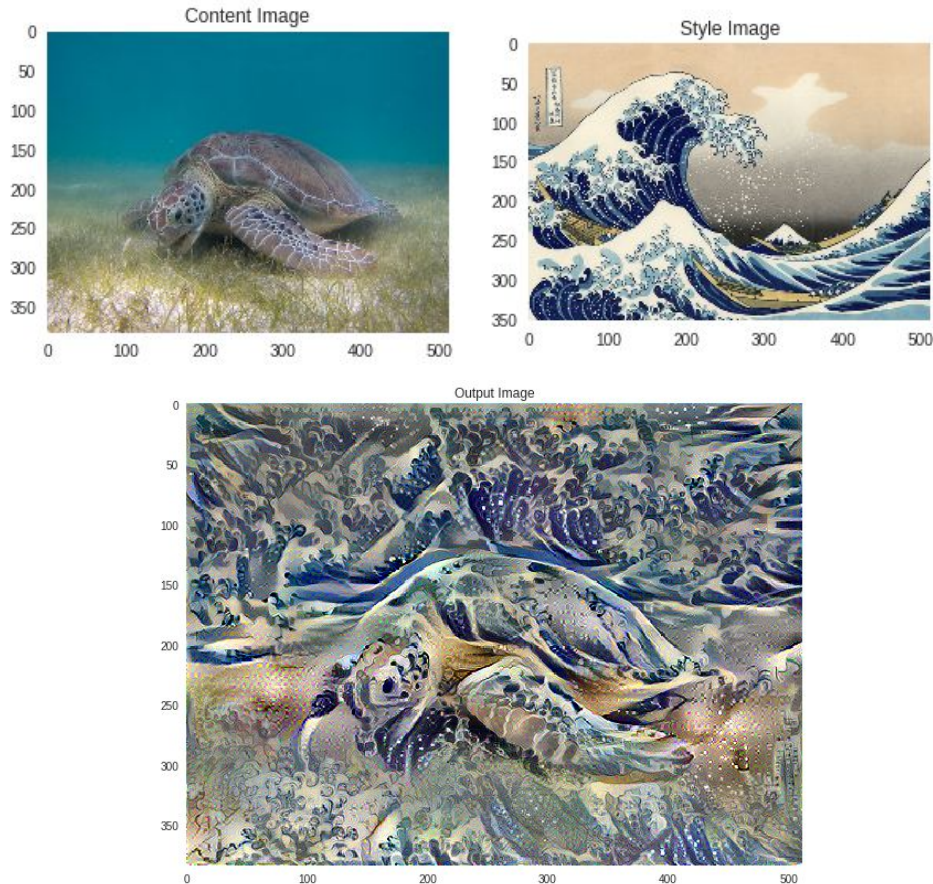


The columns show different relative weightings between the content and style reconstruction.

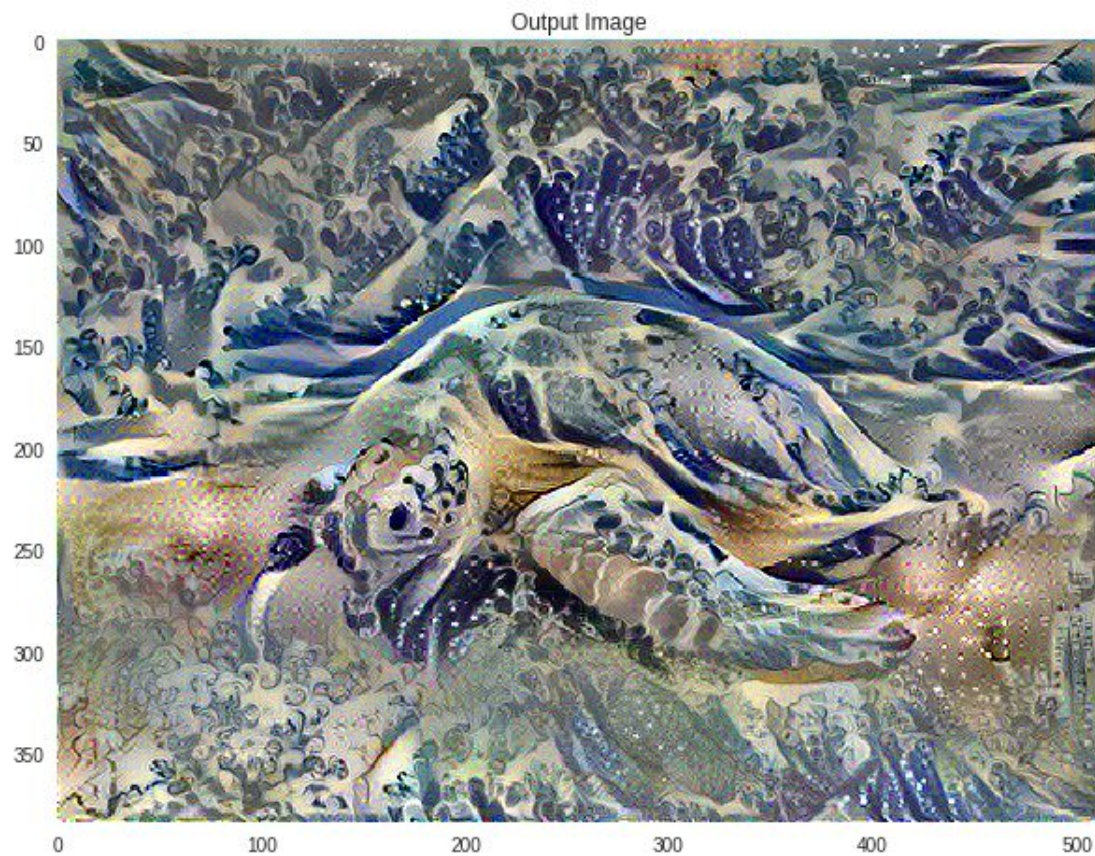
The number above each column indicates the ratio  $\alpha/\beta$  between the emphasis on matching the content of the photograph and the style of the artwork



# Analysis: Using a painting as reference

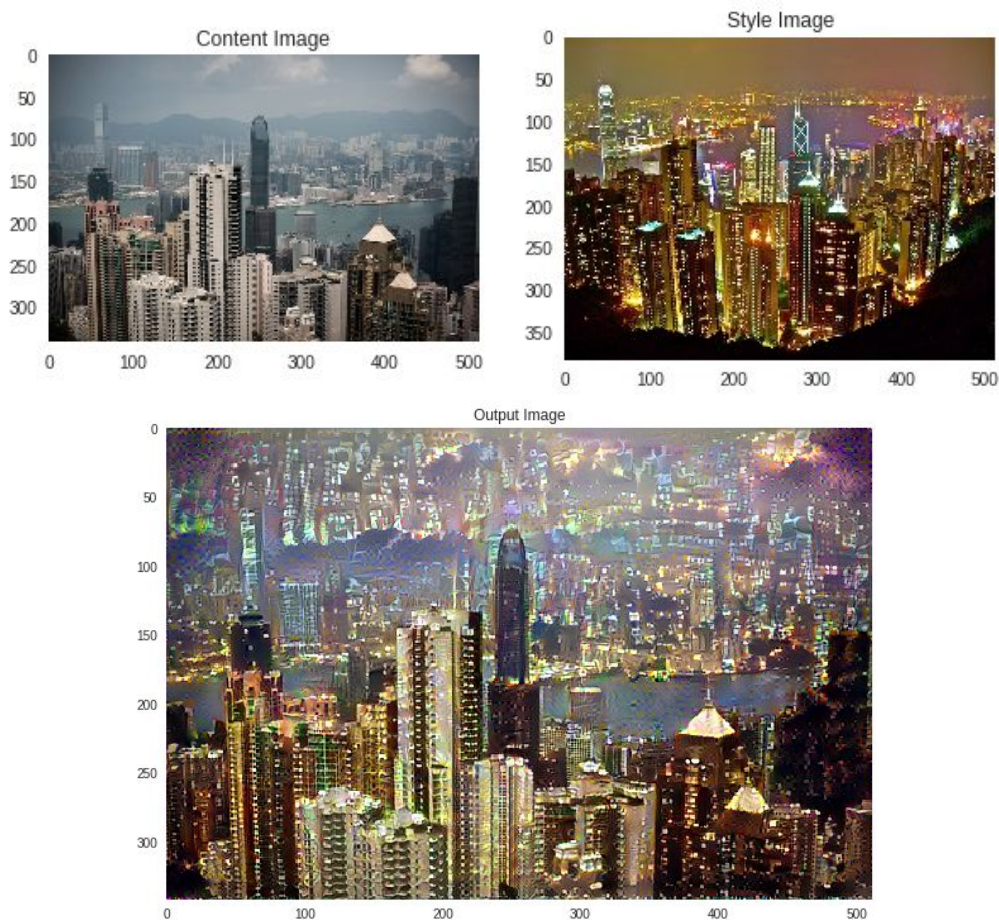






Total loss:  $1.34\text{e}+6$   
(1000 iterations)

# Analysis: Using real image as reference



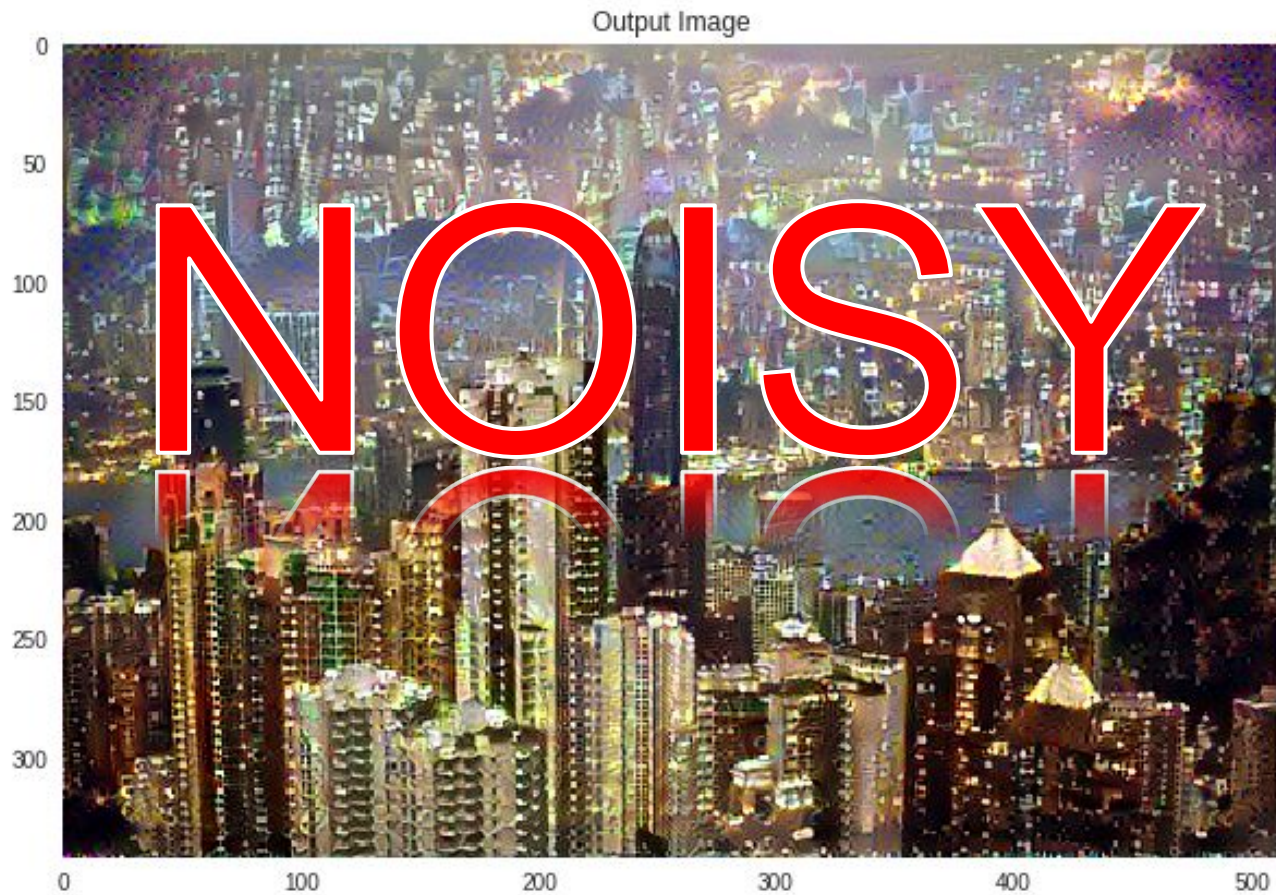


Output Image



Total loss:  $1.35e+6$





Total loss:  $1.35e+6$

# Scope of improvement

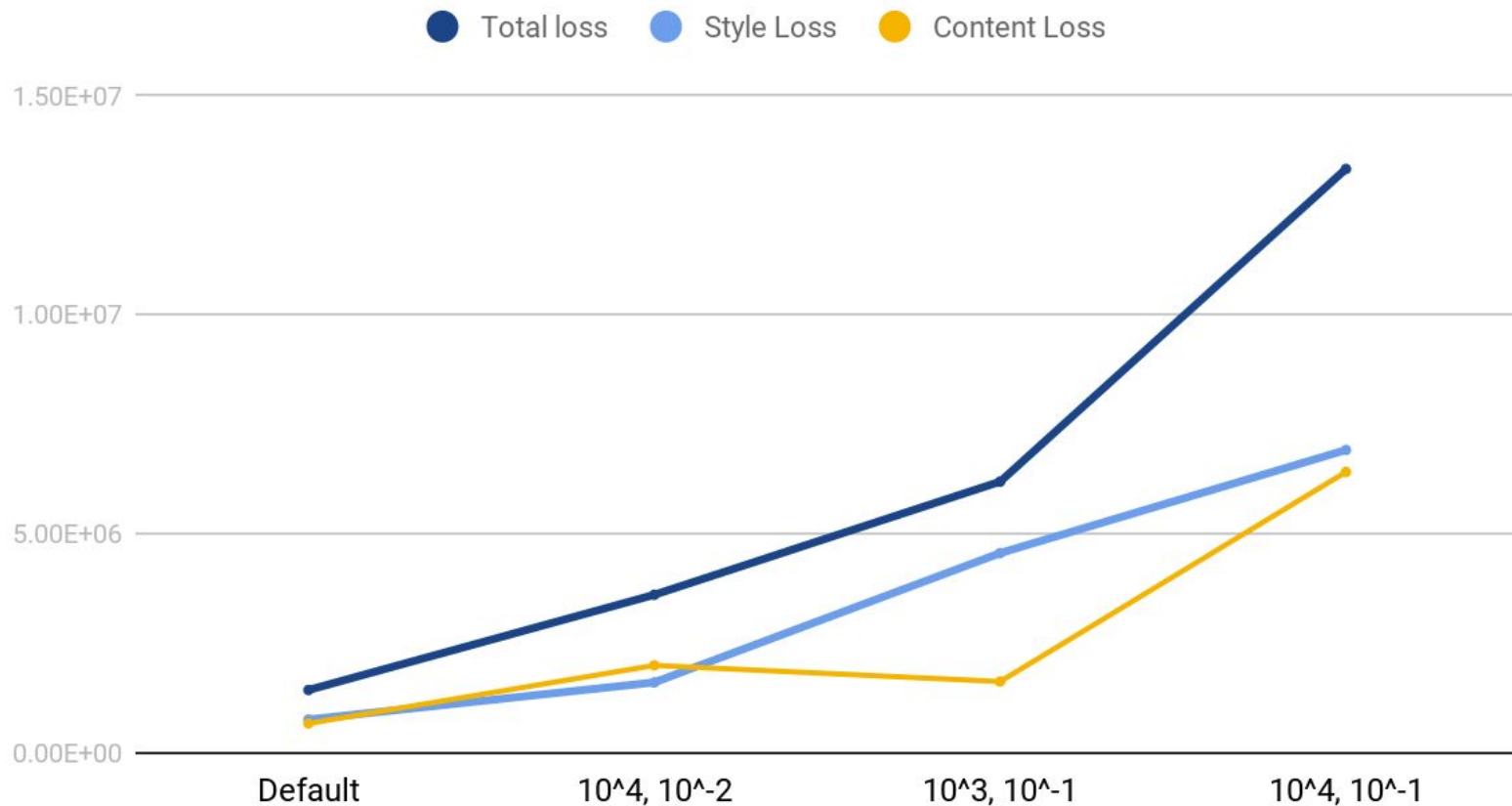
- **Noise & Computation time:** Varying the following may be able achieve the same loss in less number of iterations:
  - Loss function hyperparameters
  - Set of layers for style features
  - Weights of individual style layers
  - Optimizer parameters
  - Image resolution
  - Pre-trained model

*Photo-realistic:*

**Current limitation:** *Style copied fully from reference image, and applied to the whole content image*

**Solution:** *Image segmentation masks*

## Turtle + The Great Wave off Kanagawa: Vary alpha, beta



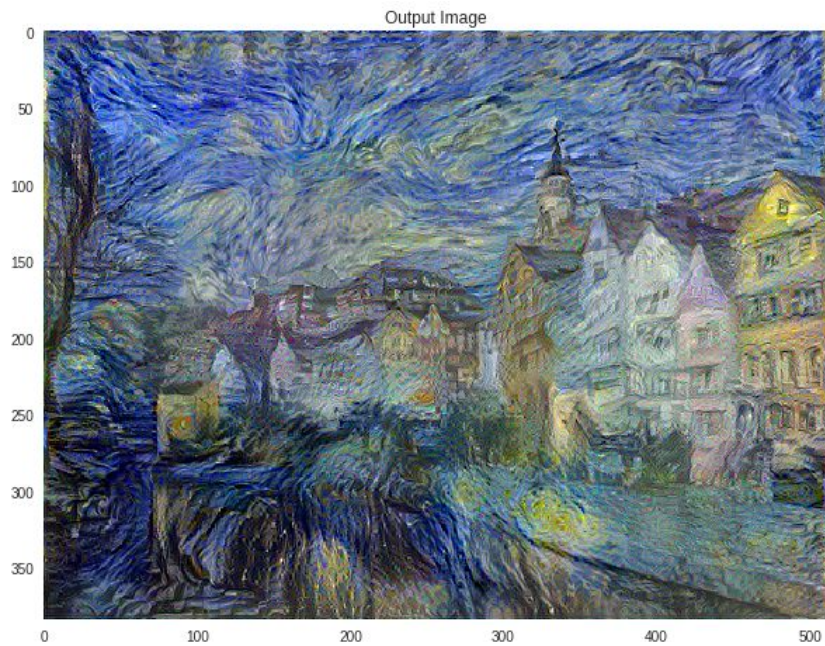


Content Image

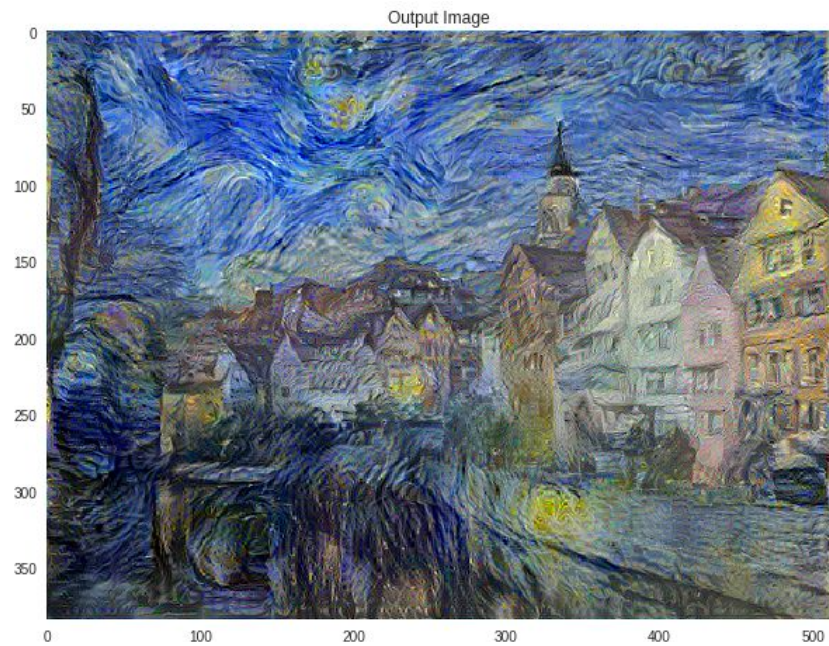


Style Image



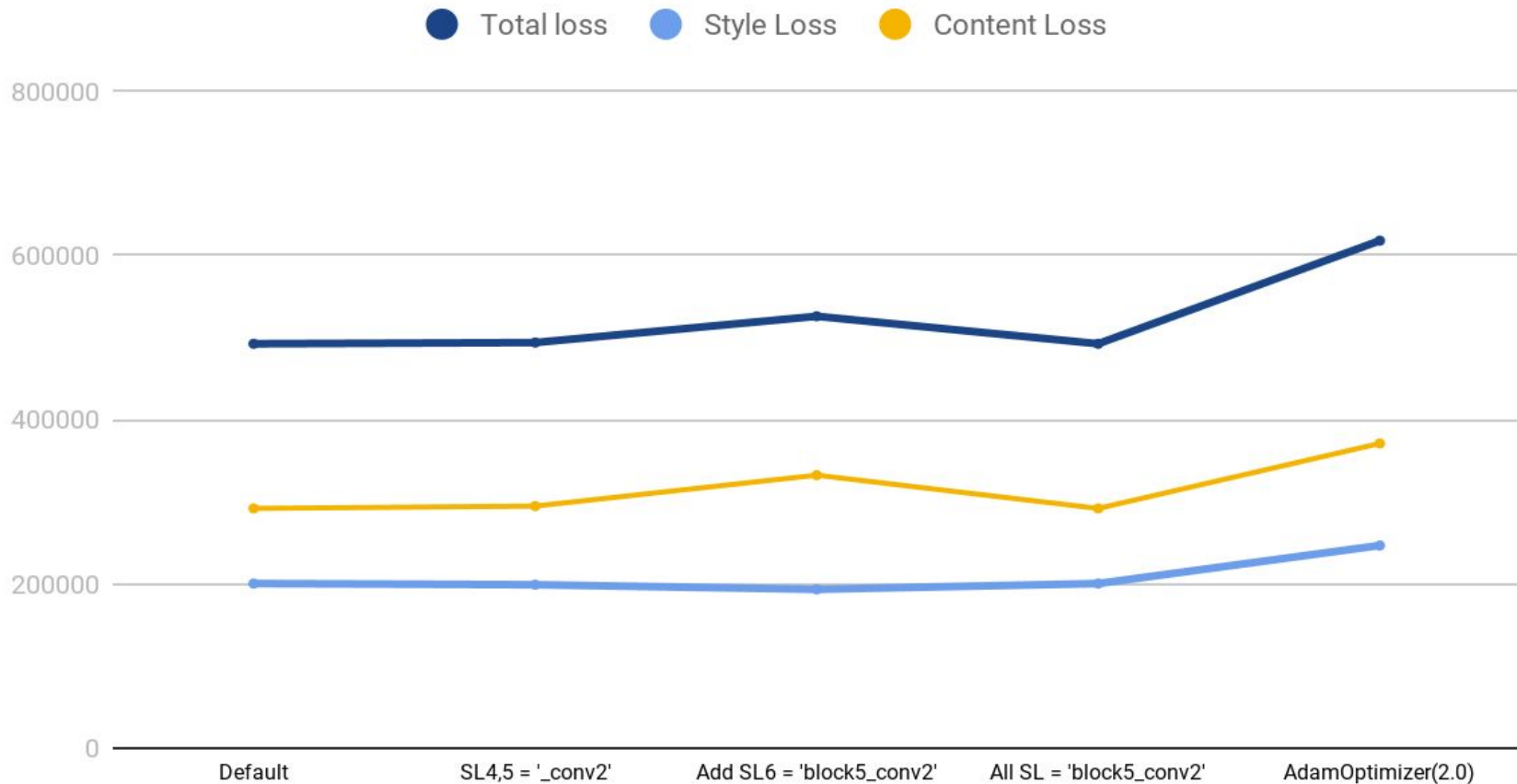


Default



SL4,5=\_conv2

# Starry night + Tuebingen

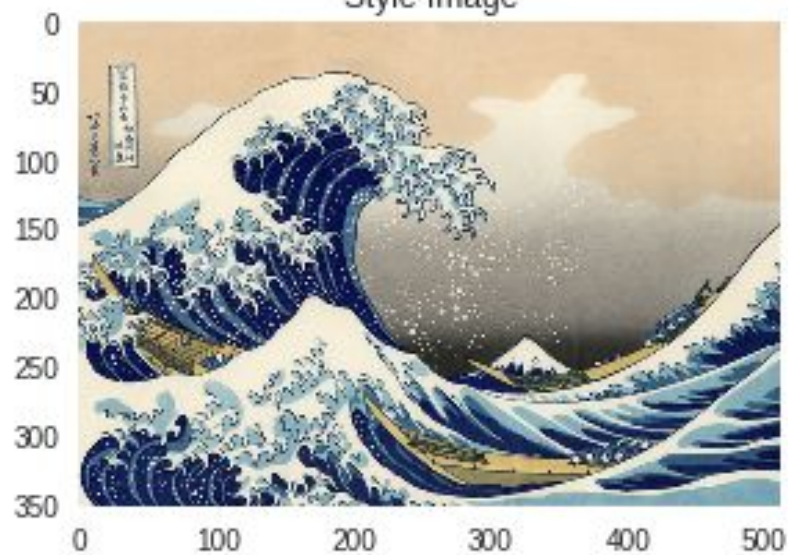


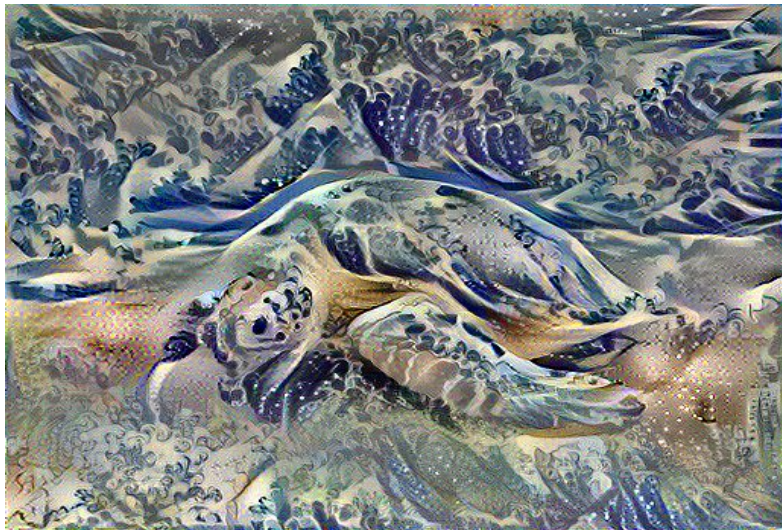


Content Image



Style Image



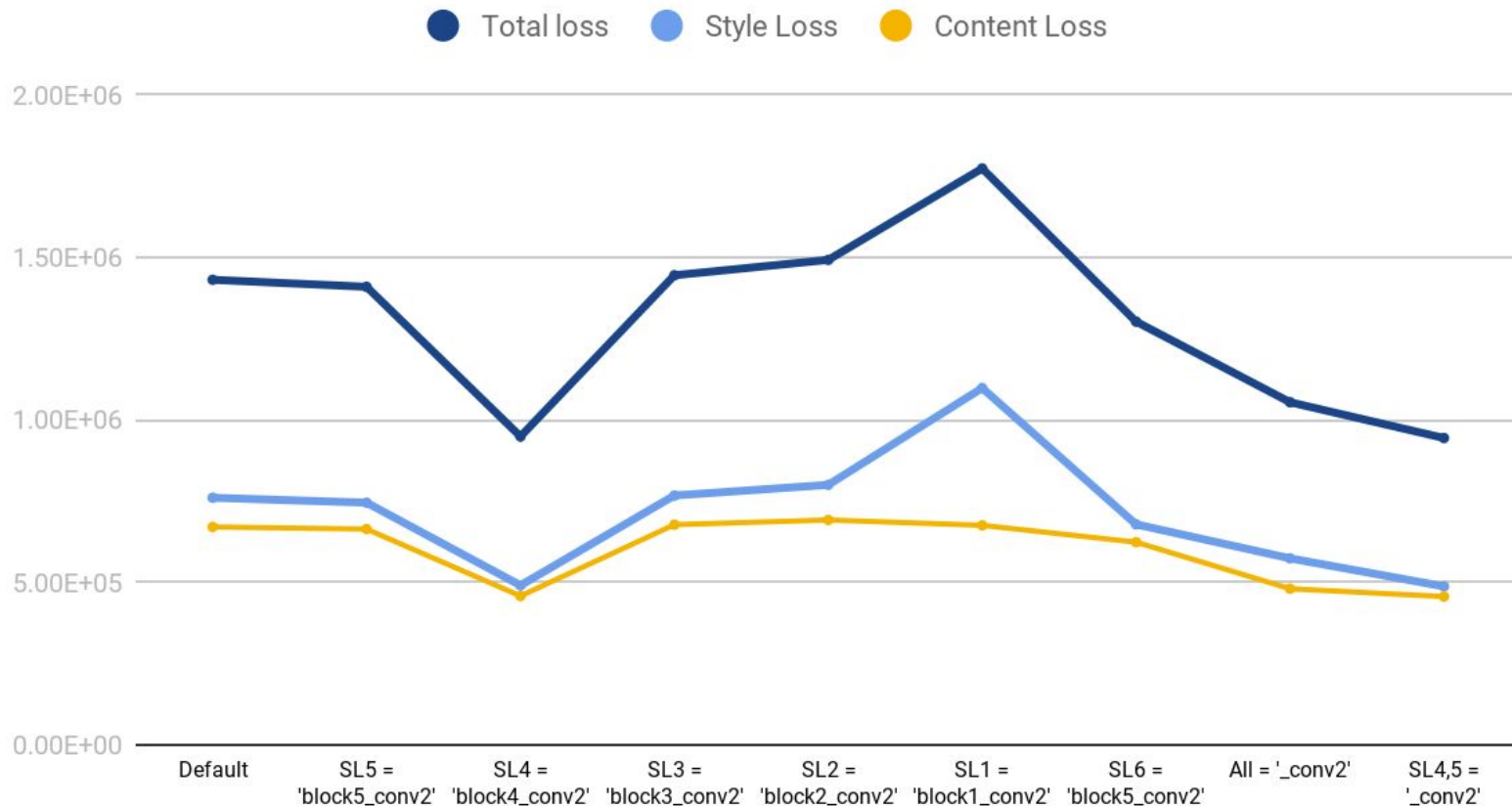


Default



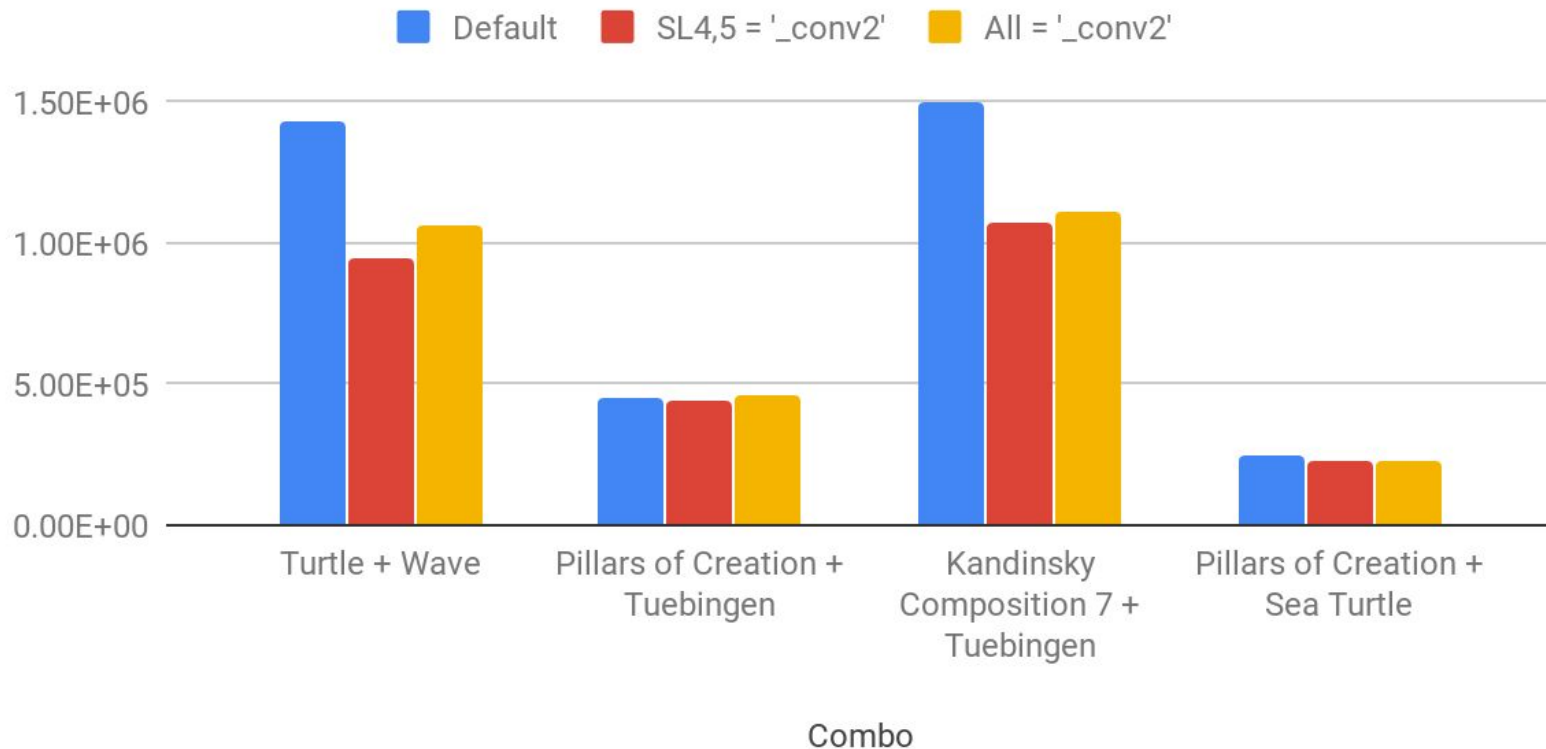
SL4,5=\_conv2

# Turtle + The Great Wave of Kanagawa: Changing the Style Layers





## Effect of different style layers on content-style image combos on total loss

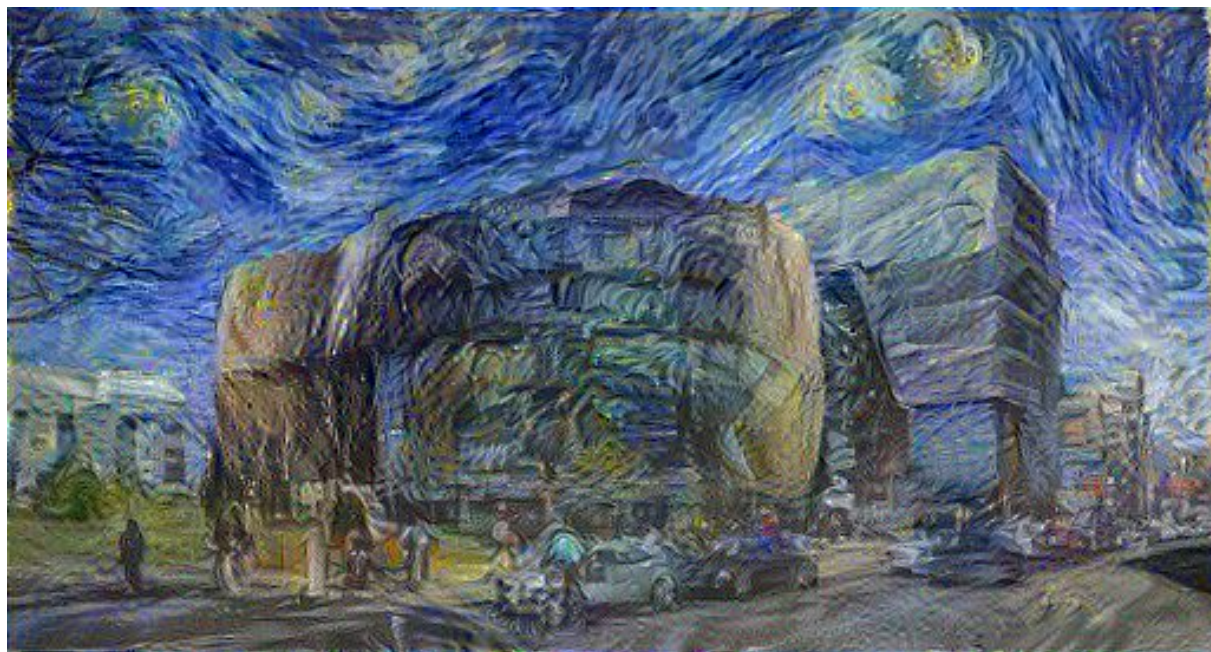


Content Image



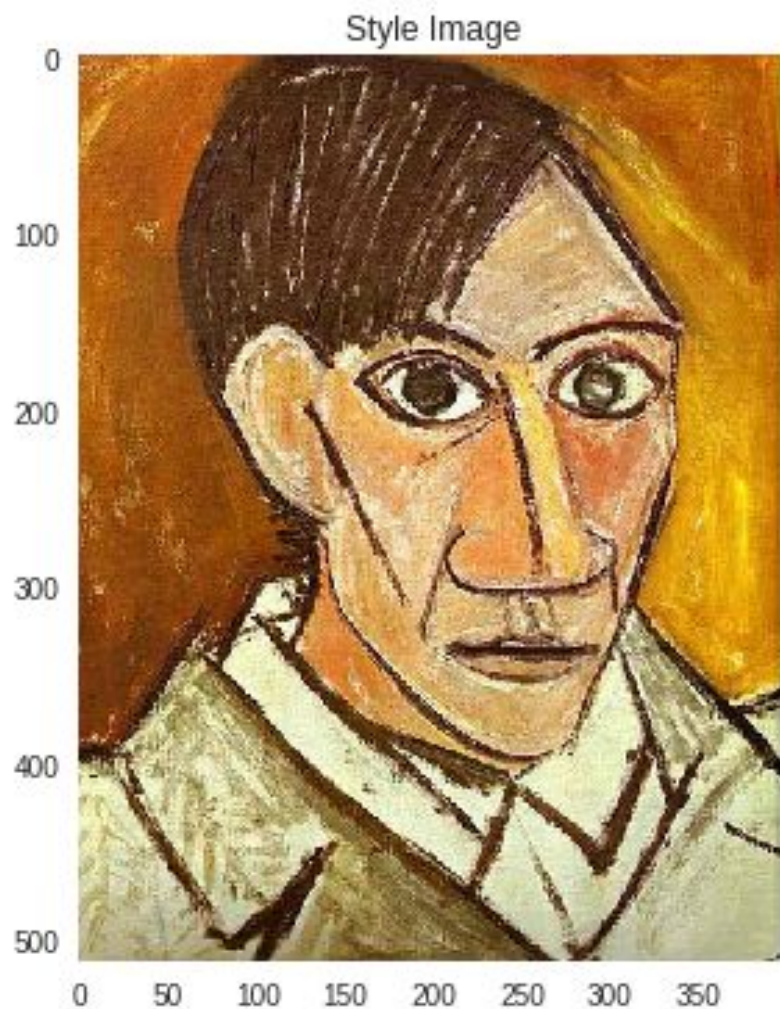
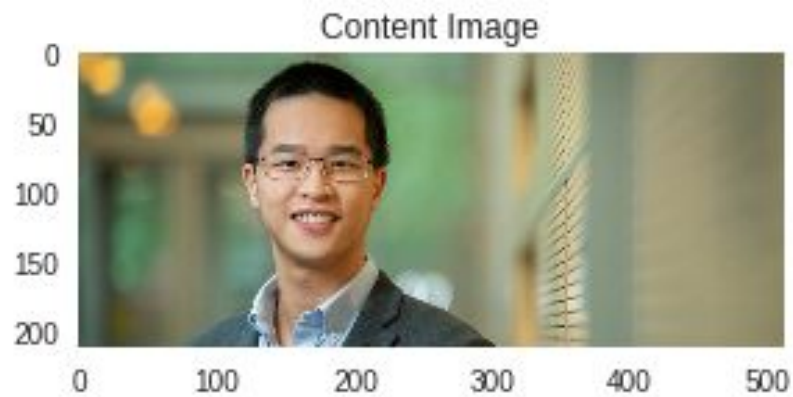
Style Image





SL4,5=\_conv2







Default  
(Total Loss =  $9.9596e+05$ )



Reduced Loss  
by  
18.8%



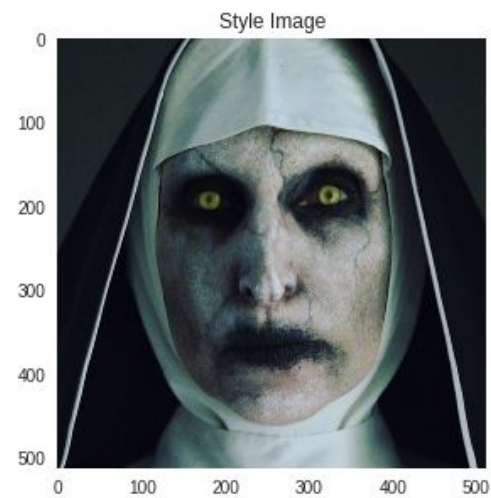
SL4,5=\_conv2  
(Total Loss =  $8.0869e+05$ )











# References

- Gatys et al., 2015 A Neural Algorithm of Artistic Style
- Luan, Fujun et al., 2017 Deep Photo Style Transfer
- <https://github.com/LouieYang/deep-photo-styletransfer-tf>
- <https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/>



# Thank you...

*Presented by*

*Krit Goyal  
Bishwarup Neogy*