

Transferência de Estilo: Usando Redes Neurais para Criar Arte

Adilson Medronha
Luís Lima

Trabalho II de Aprendizado de Máquina

MALTA

Machine Learning Theory
and Applications Lab



Sumário

- I. Introdução
- II. Fundamentação Teórica
- III. Detalhes da Implementação
- IV. Demonstração

Introdução

Então, qual é o problema?



Starry Night



Poppy Field

+

=

?

Introdução

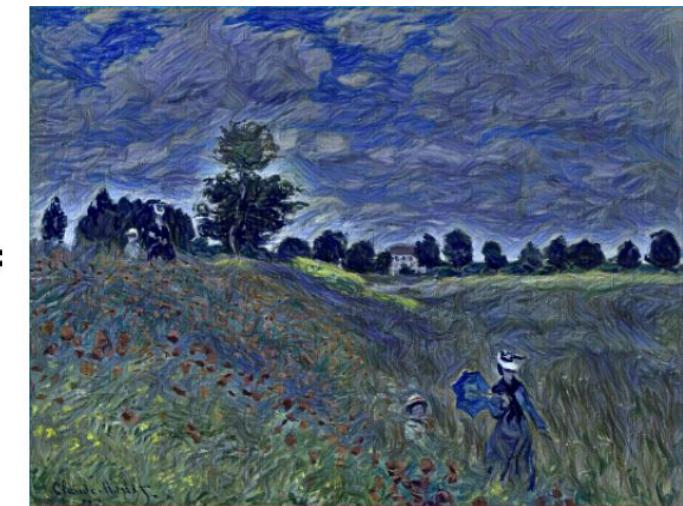
Então, qual é o problema? As redes neurais resolvem



Starry Night



Poppy Field



Van Gogh + Monet

Introdução



Conteúdo

+



Estilo

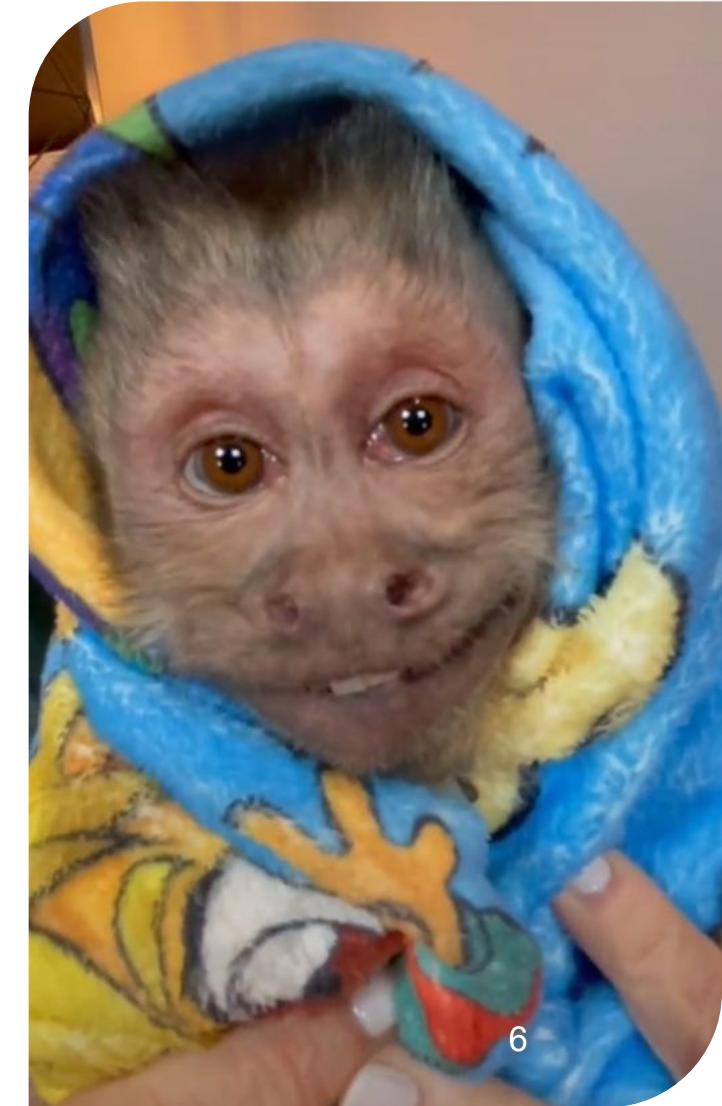
=



Resultado

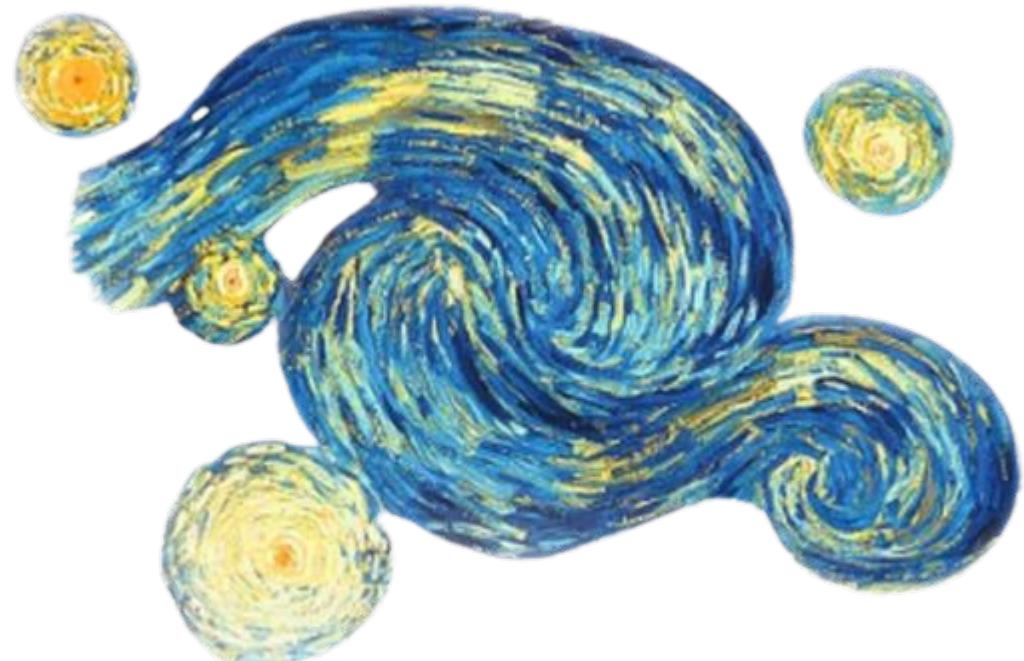
Fundamentação Teórica

**A essência da transferência de estilo
utilizando redes neurais...**



Fundamentação Teórica

Afinal, o que são estilos?

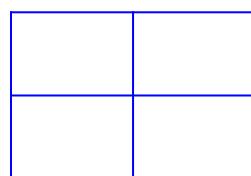


Fundamentação Teórica

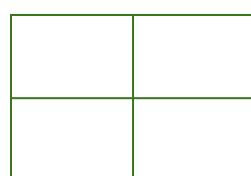
Afinal, o que são estilos?



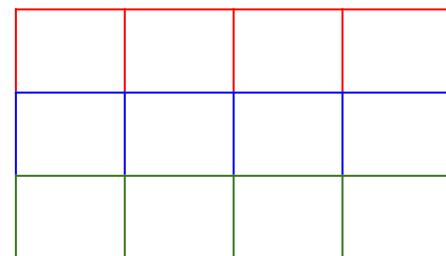
Pontos



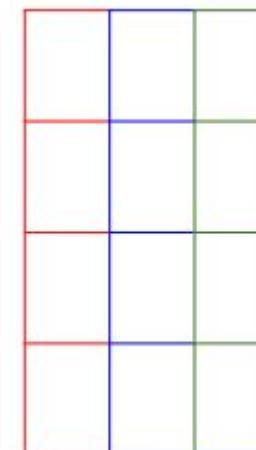
Linhas



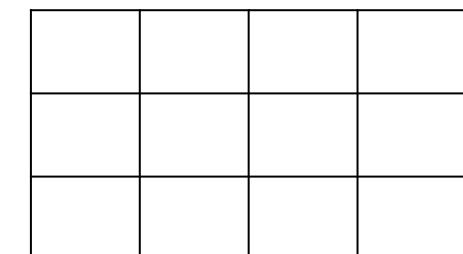
Círculos



*



=



Gram matrix

Fundamentação Teórica

Afinal, o que são estilos?



Fundamentação Teórica

Função de Custo do Estilo

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

$$S_l(o, \hat{g}) = \frac{1}{2} \sum_{i,j} (O_{ij}^l - G_{ij}^l)^2$$

Fundamentação Teórica

Função de Custo do Estilo

$$S_l(o, \hat{g}) = \frac{1}{2} \sum_{i,j} (O_{ij}^l - G_{ij}^l)^2$$
$$L_{estilo}(o, \hat{g}) = \sum_{l=0}^L s_l S_l(o, \hat{g})$$

Fundamentação Teórica

Função de Custo do Conteúdo

$$C_l(o, \hat{g}) = \frac{1}{2} \sum_{i,j} (O_{ij}^l - G_{ij}^l)$$

Fundamentação Teórica

Função de Custo do Conteúdo

$$C_l(o, \hat{g}) = \frac{1}{2} \sum_{i,j} (O_{ij}^l - G_{ij}^l)$$

$$L_{conteudo}(o, \hat{g}) = \sum_{l=0}^L c_l C_l(o, \hat{g})$$

Fundamentação Teórica

Função de Custo do Total

$$L_{total}(p, a, x) = \alpha L_{conteudo} + \beta L_{estilo}$$

Hiperparâmetros

Fundamentação Teórica

Função de Custo do Total

$$L_{total}(p, a, x) = \alpha L_{conteudo} + \beta L_{estilo}$$

Hiperparâmetros

$$\frac{\alpha}{\beta} = r = \begin{cases} \text{prioriza o conteúdo,} & \text{se } r = \infty \\ \text{prioriza o estilo,} & \text{se } r = -\infty \end{cases}$$

Fundamentação Teórica

Função de Custo do Total

$$L_{total}(p, a, x) = \alpha L_{conteudo} + \beta L_{estilo}$$

Hiperparâmetros

$$\frac{\alpha}{\beta} = \frac{1e^5}{1e^{10}} = 1e^{-5}$$

Implementação

Configuração do ambiente de execução

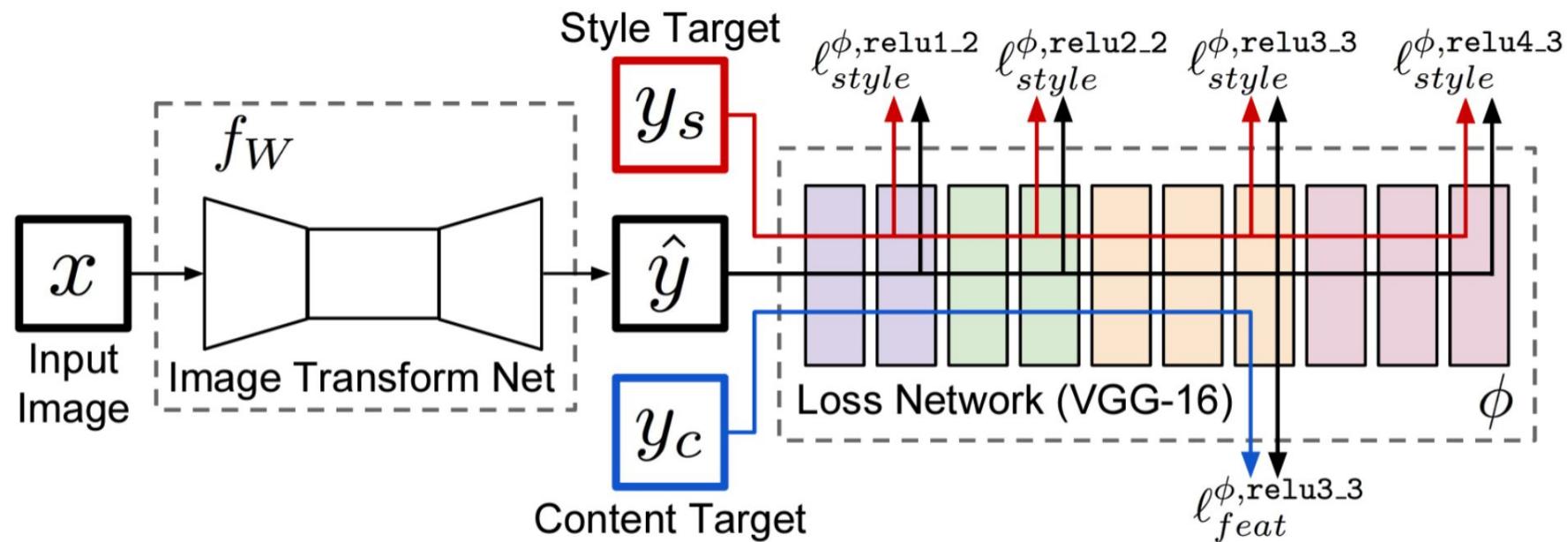
- I. TITAN XP 12gb
- II. CPU 32 cores
- III. 128gb RAM

Configuração do nosso treinamento

- *Dataset*: COCO2017 - 118k
- *Batch size* = 16
- *Img-size* = 256x256
- *Seed* = 42
- *Content-weight* = 1e5
- *Style-weight* = 1e10
- *Learning-rate* = 1e-3

Implementação

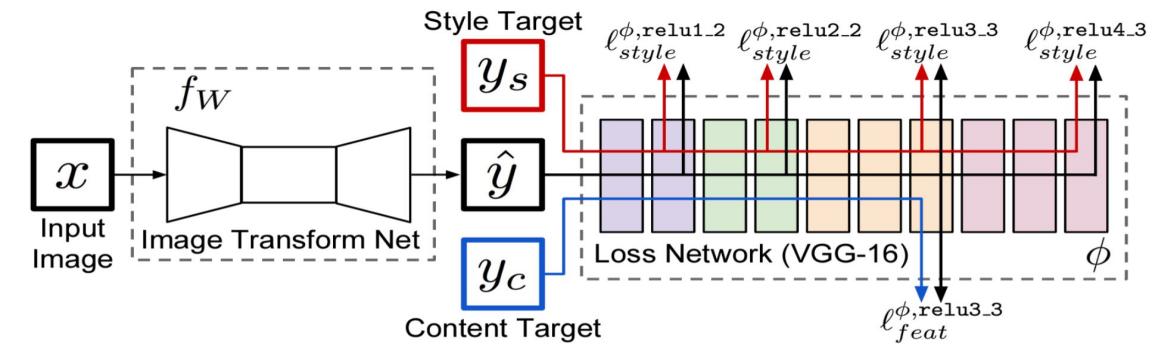
Neural Style Transfer



Implementação

Exemplificando...

y_c



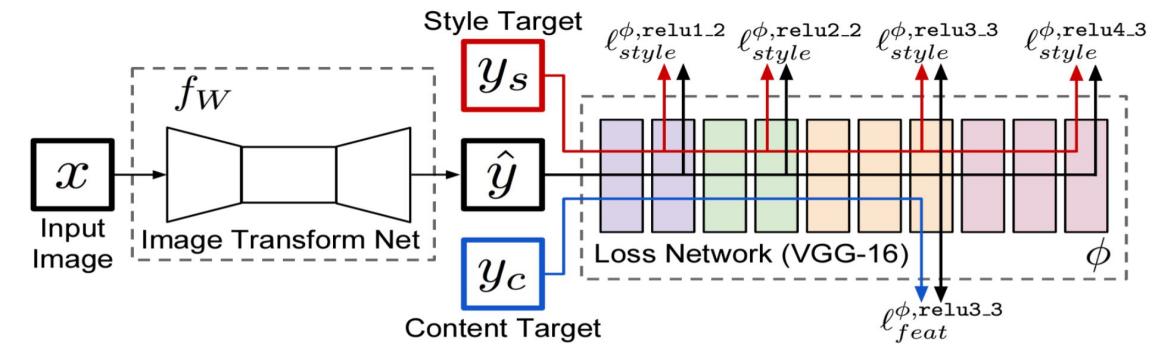
Implementação

Exemplificando...

y_c



y_s



Implementação

Exemplificando...

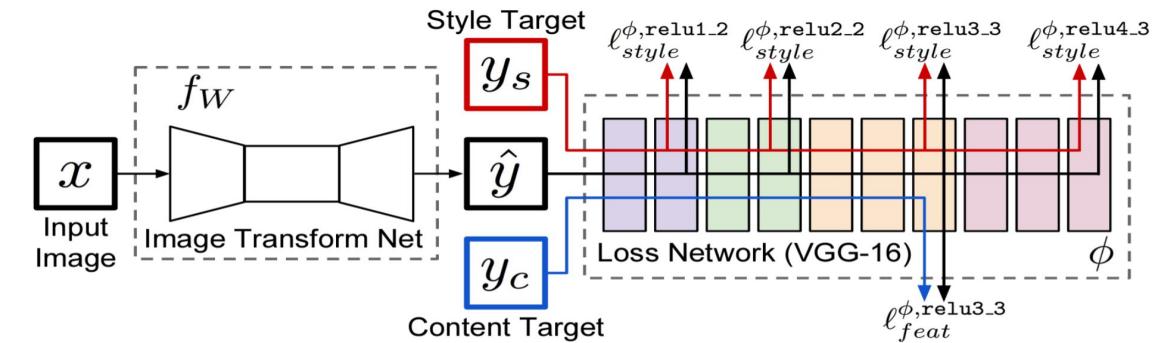
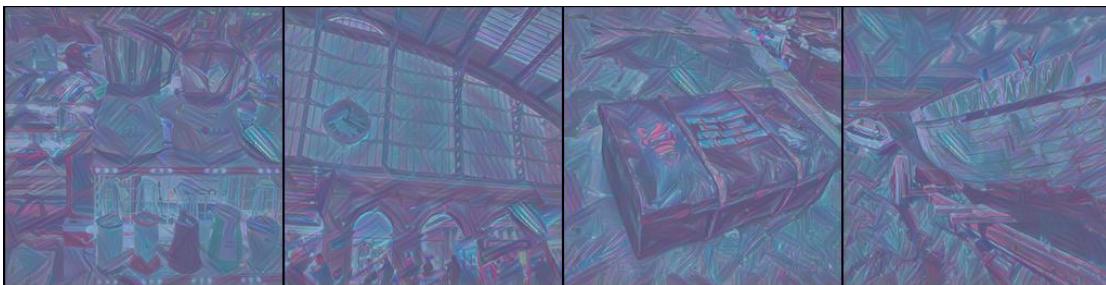
y_c



y_s

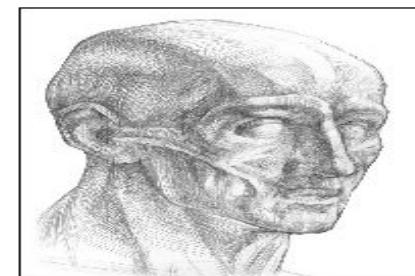
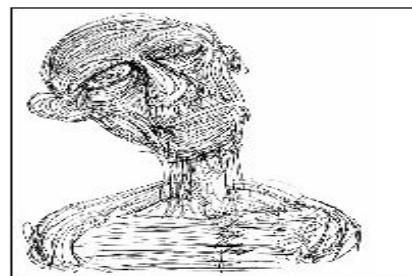
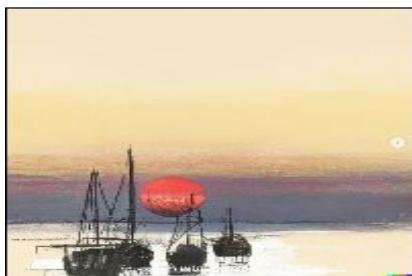
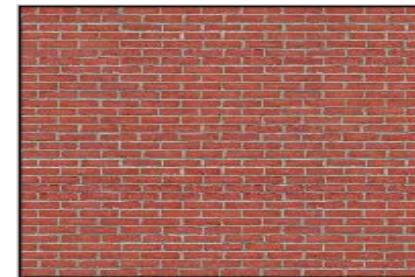


\hat{y}



Implementação

Nossos modelos



Implementação



Implementação

Compensa?



2 épocas



44 épocas

Implementação

Compensa?



2 épocas



44 épocas



Diferença

Implementação

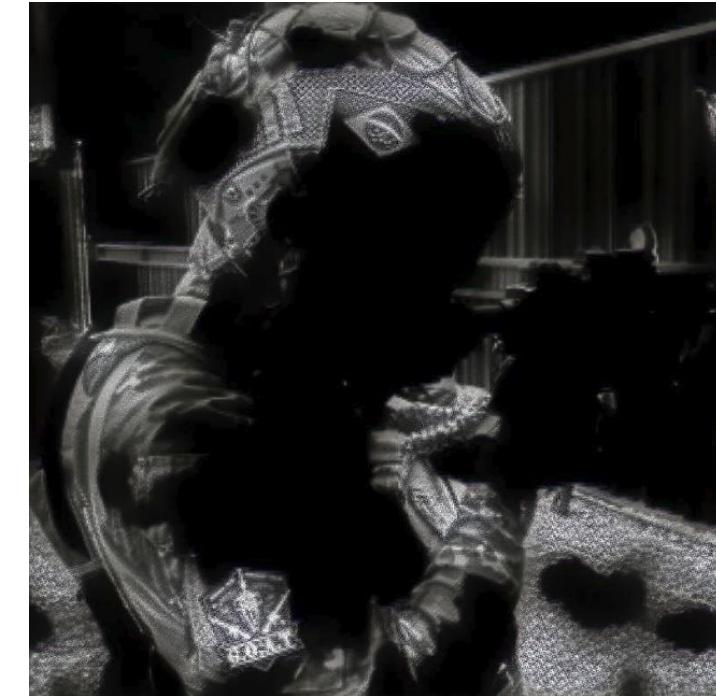
1ª Peculiaridade, **treino incorreto**



+



Conteúdo

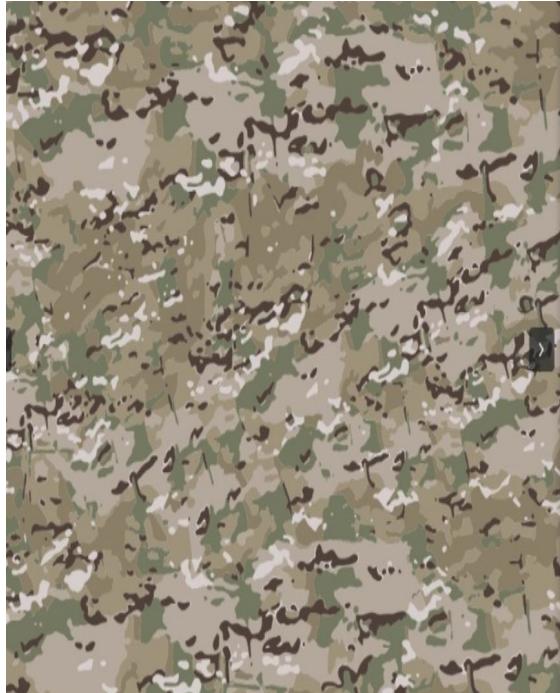


Resultado

Estilo

Implementação

1ª Peculiaridade, solução



+



Conteúdo

=

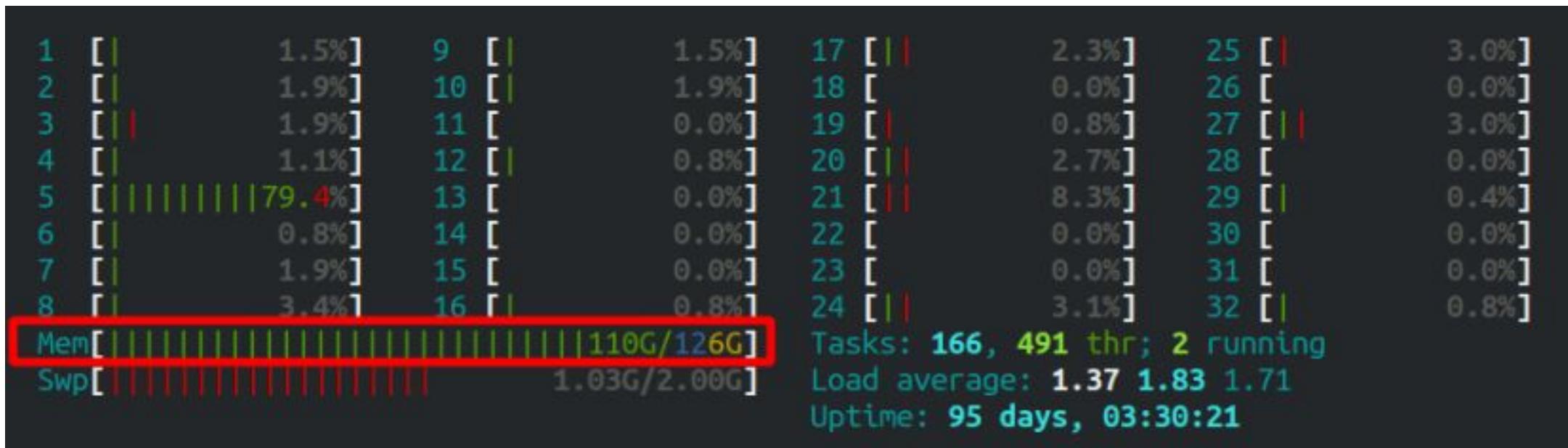


Resultado

Estilo

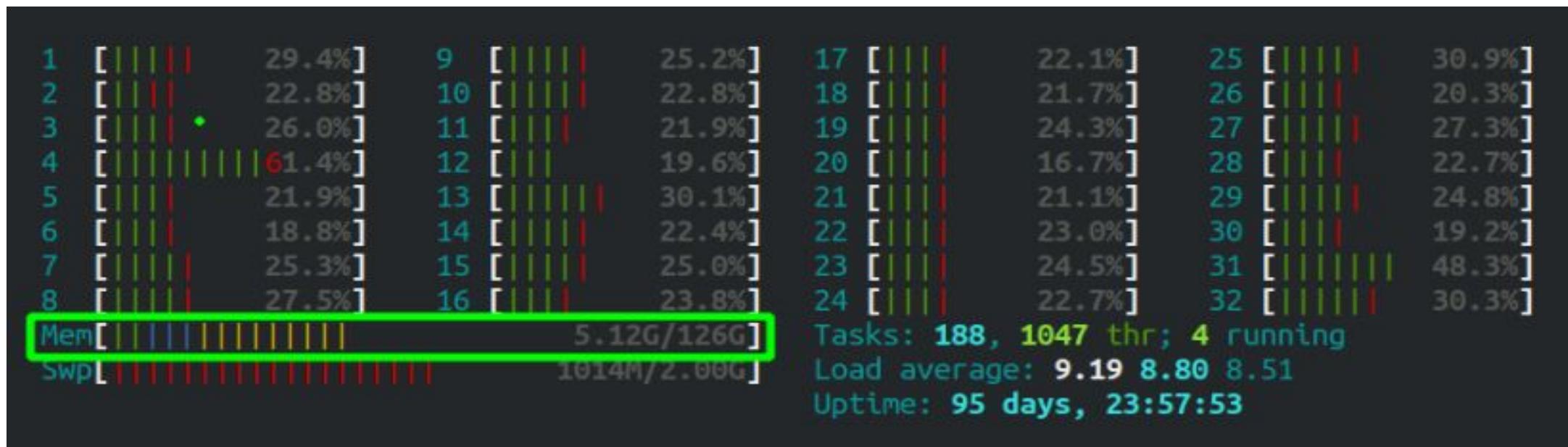
Implementação

2ª Peculiaridade, *memory leak*



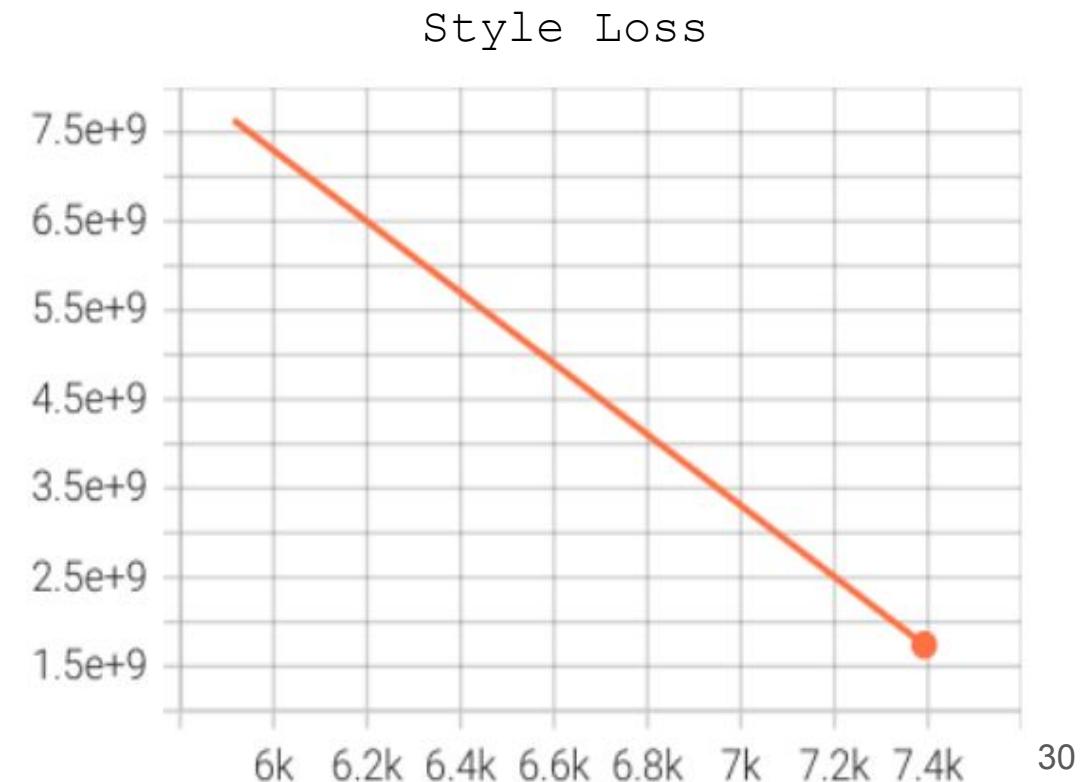
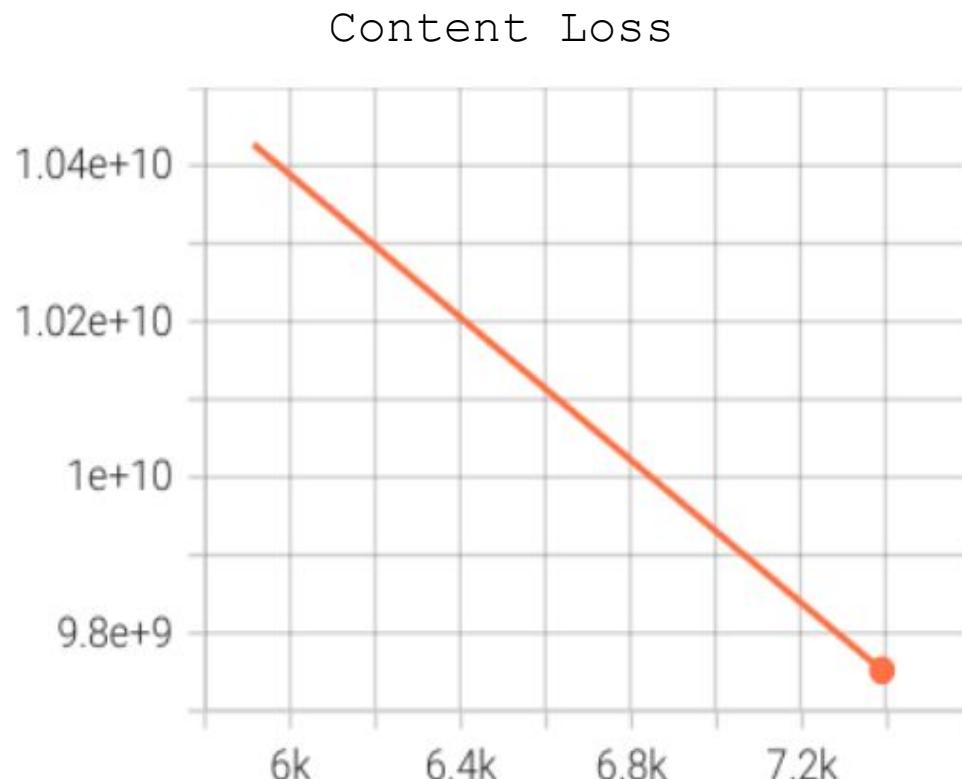
Implementação

2ª Peculiaridade, solução



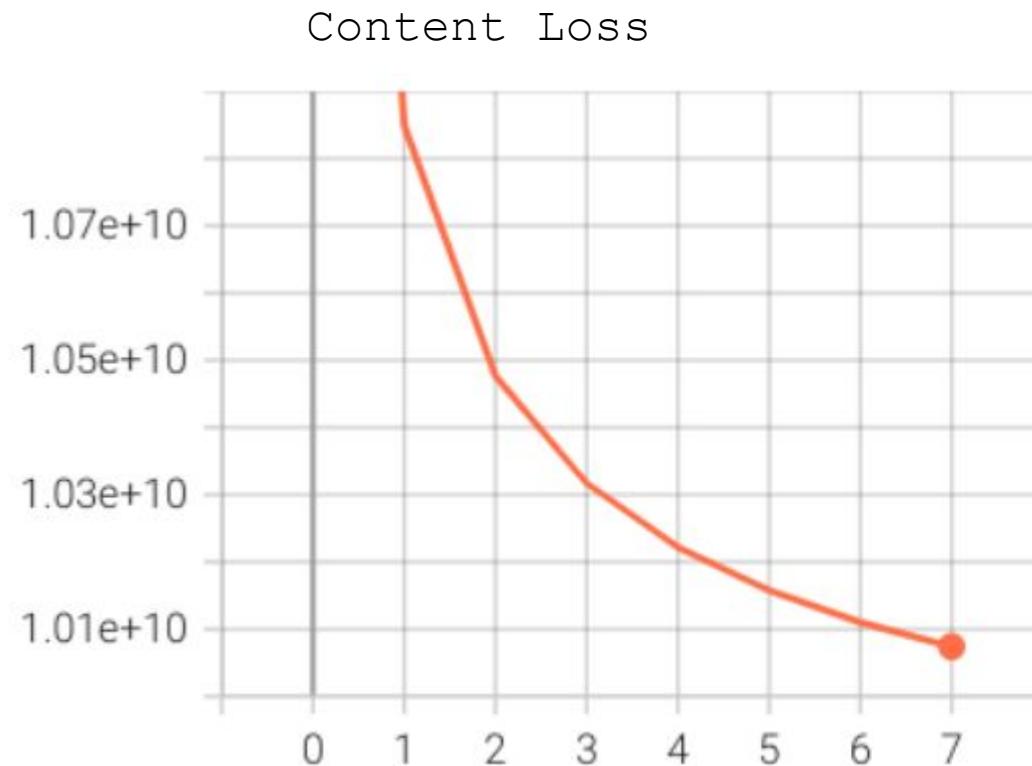
Implementação

3ª Peculiaridade, *loss esquisita!*



Implementação

3ª Peculiaridade, solução



Implementação

4ª Peculiaridade, *vram error*

Testando um *batch size* = 8 ...

```
RuntimeError: CUDA out of memory. Tried to allocate 9.14 GiB (GPU 0; 11.91 GiB total
capacity; 10.03 GiB already allocated; 1.00 GiB free; 10.05 GiB reserved in total by PyTorch)
If reserved memory is >> allocated memory try setting max_split_size_mb to avoid
fragmentation. See documentation for Memory Management and PYTORCH_CUDA_ALLOC_CONF (t2-ml)
adilson@motoz-pin:/A/adilson/git/faculdade/style-transfer/neural_style$
```

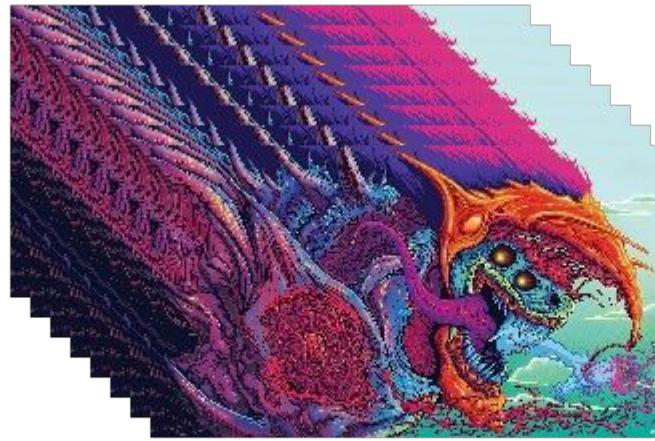
Implementação

4ª Peculiaridade, *vram error*

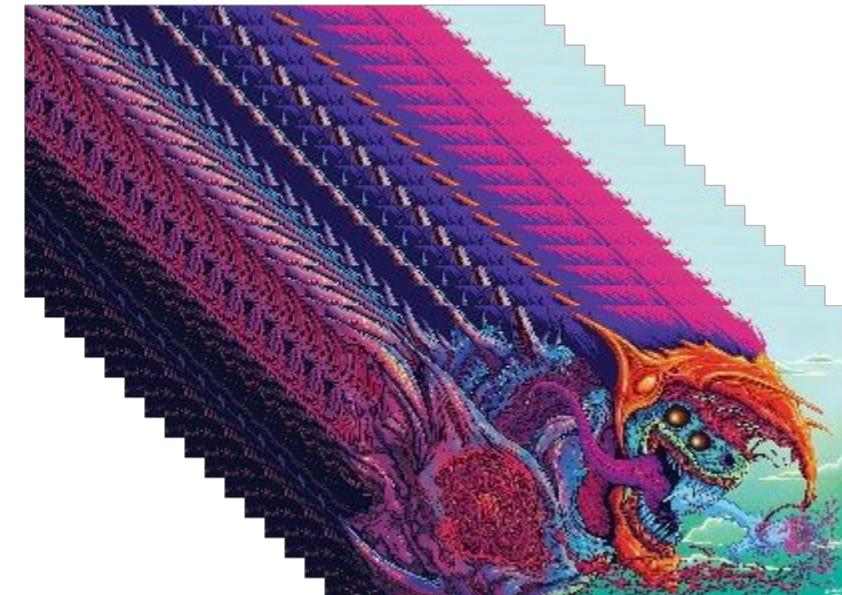
$$1461 \times 820 \times 3 = 3594060$$



$$3594060 \times 8 = 28752480$$



$$3594060 \times 16 = 57504960$$



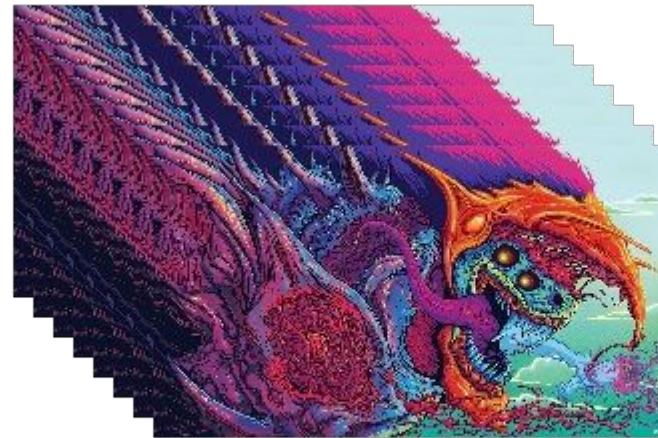
Implementação

4ª Peculiaridade, solução (*style resized*)

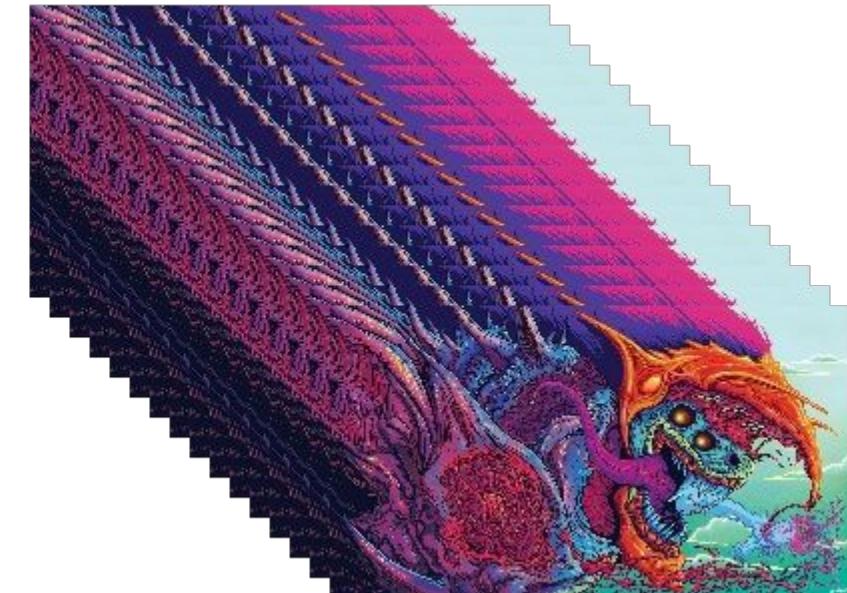
$256 \times 256 \times 3 = 196608$



$196608 \times 8 = 1572864$

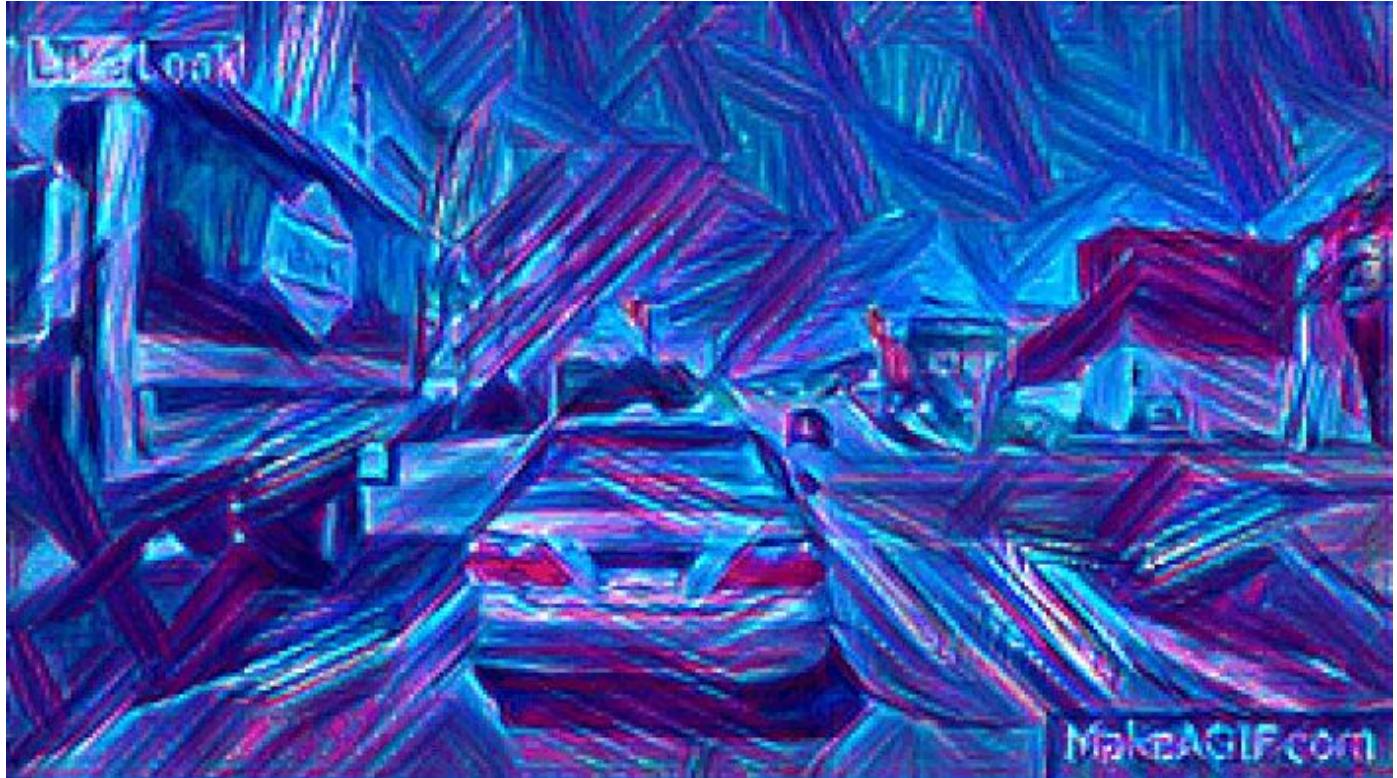


$196608 \times 16 = 3145728$



Implementação

Limitação do Método



Demonstração

Selecionar a imagem

tree.png

Selecionar o tipo

starry_night

Transferência de Estilo

Imagem:



Estilizar

Imagen Estilizada:



Referências

- Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. A neural algorithm of artistic style. *CoRR*, abs/1508.06576, 2015.
- Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. *CoRR*, abs/1603.08155, 2016.
- Yongcheng Jing, Yezhou Yang, Zunlei Feng, Jingwen Ye, and Mingli Song. Neural style transfer: A review. *CoRR*, abs/1705.04058, 2017.

Obrigado! Perguntas?

MALTA

Machine Learning Theory
and Applications Lab

