IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Style Transfer by Relaxed Optimal Transport and Self-Similarity

Group 5:
Houda GHALLAB
Renzo MORALES
Carlos ARGUILAR

SUMMARY

1. Method recap
2. Hyperparameters
3. User control
4. Examples and results
5. Conclusion

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

METHOD RECAP

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Content

+

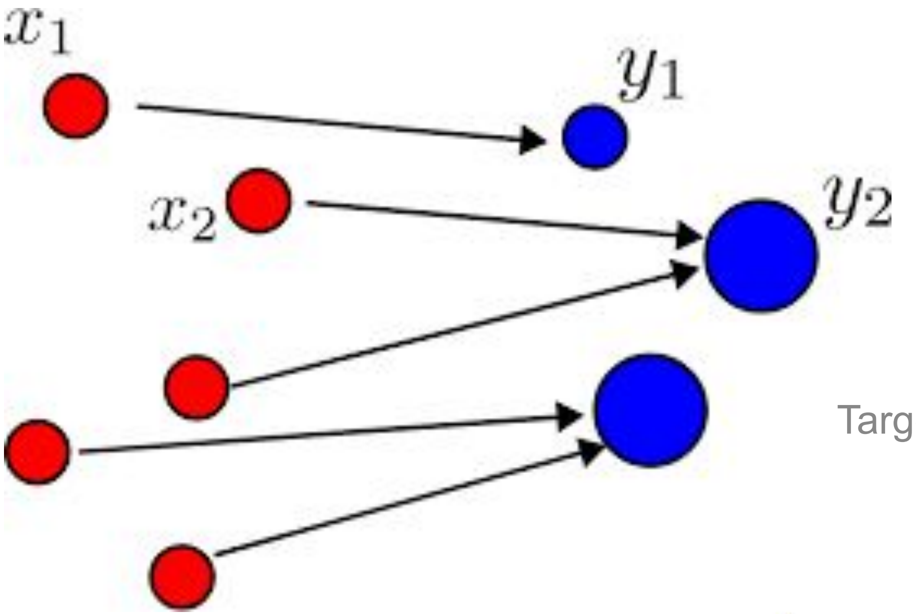
Style

=

Combination

The optimal transport is the problem that tries to find the transport map

Source Space



Target Space

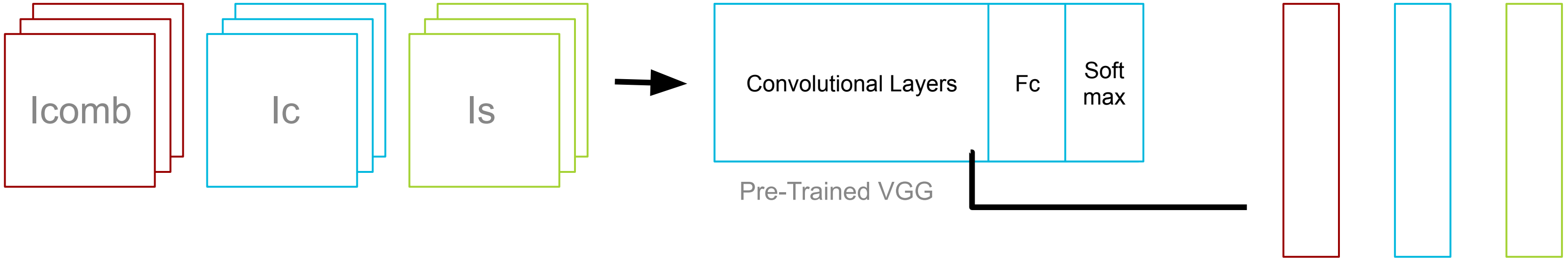
$$\ell_r = \max\left(\frac{1}{n}\sum_i \min_j C_{ij}, \frac{1}{m}\sum_j \min_i C_{ij}\right)$$

CHAPITRE 1 : METHOD RECAP

Iterative process: Gradient descent

Feature map

1



2

$$L(X, I_C, I_S) = \frac{\alpha \ell_C + \ell_m + \ell_r + \frac{1}{\alpha}\ell_p}{2 + \alpha + \frac{1}{\alpha}}$$

3 Backpropagate, reconstruct

Palette loss

Alpha pondered content

Style Loss

$$L(X, I_C, I_S) = \frac{\alpha \ell_C + \underbrace{\ell_m + \ell_r + \frac{1}{\alpha}\ell_p}_{\text{Style Loss}}}{2 + \alpha + \frac{1}{\alpha}}$$

Content loss

Moment loss

REMD Loss

HYPERPARAMETERS

CONTENT WEIGHT (α)

Default content weight : 1.0

x16

Adjust its influence effectively

BALANCE

CONTENT IMAGE

STYLE IMAGE

HIGHER content weight

LOWER content weight

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CHAPITRE 2 : HYPERPARAMETERS

CONTENT WEIGHT (α)

The iterative process described previously is made for 4 different scales using the (upscaled) output of the previous scale as input, halving the content weight (α) for the next scale:

long side 64px

16α

long side 128px

4α

long side 256px

2α

long side 512px

α

→ Optimization followed by upscaling

LEARNING RATE (lr)

Learning rate influences the **CONVERGENCE SPEED** and **STABILITY OF THE OPTIMIZATION PROCESS**

Default lr = 0.002 → Final lr = 2e-06

HIGHER → **LOWER**

The algorithm strikes a balance between **rapid convergence** and **fine-tuning.**

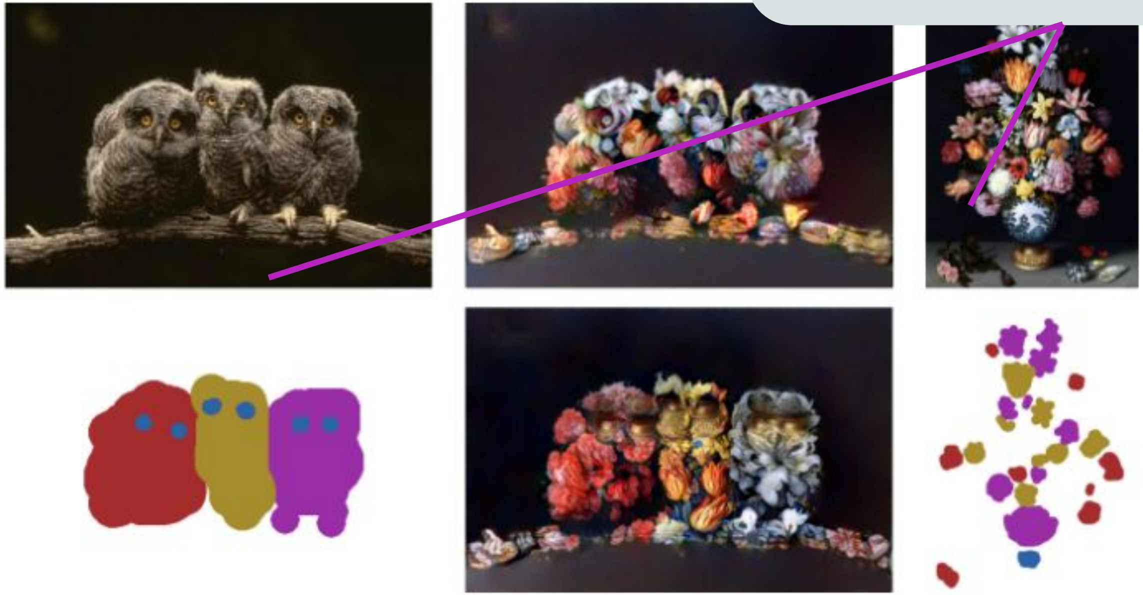
By a higher lr, the algorithm can explore solution space more extensively.

USER CONTROL

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Control is enforced by making the pairs of points in the same region have higher weight in the loss calculation

Mask specific areas to have the same style:



Content mask

Style mask

EXAMPLES AND RESULTS

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CHAPTER 4: EXAMPLE AND RESULTS
Unguided Style Transfer

$\alpha = 1$

CHAPTER 4: EXAMPLE AND RESULTS

Varying content weight


α = 0.2


α = 0.5


α = 1

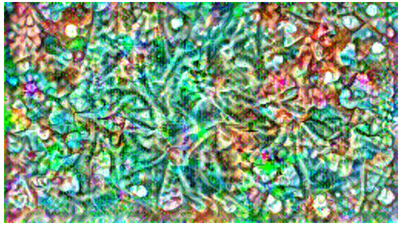

α = 2


α = 4

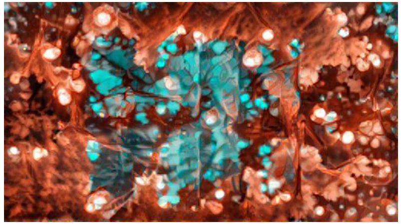
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Varying learning rate

α = 1



lr = 0,2 76.451s

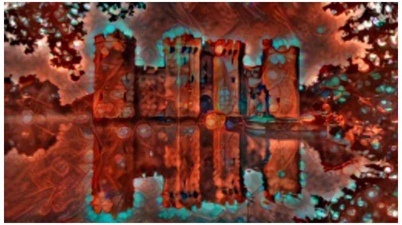


lr = 0.02 77.522s

lr = 2e-3 76.912s

lr = 2e-4 76.556s

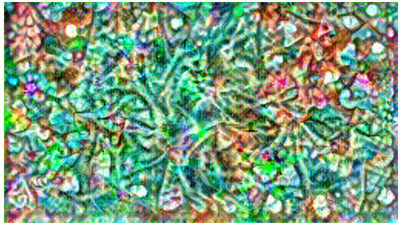
lr = 2e-5 76.691s



lr = 2e-6 79.937s

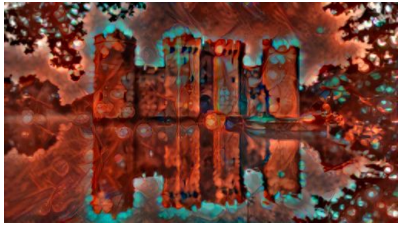
Varying learning rate

α = 1

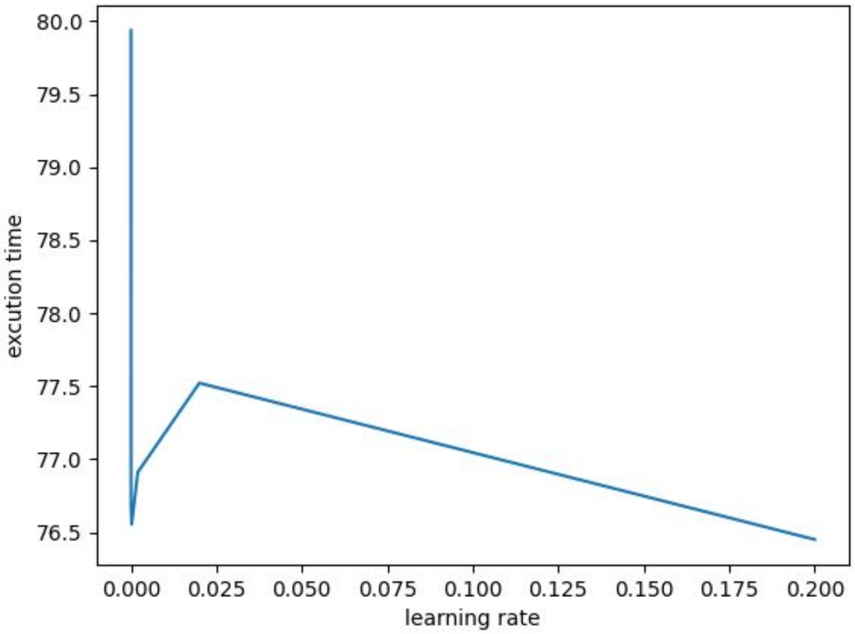


lr = 0,2   76.451s

lr = 2e-3   76.912s



lr = 2e-6   79.937s



CHAPTER 4: EXAMPLE AND RESULTS

Losses

$\alpha = 1$

lr = 2e-3

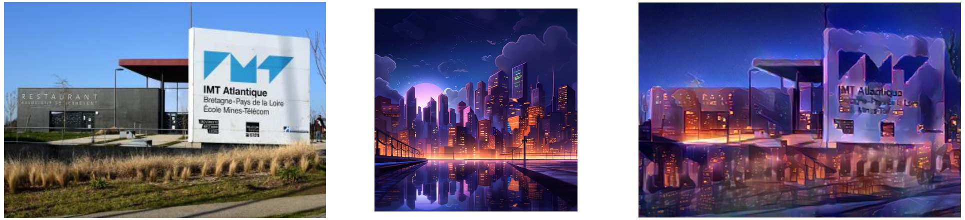
No REMD

No moment loss

Guided Style Transfer (masks)

Unguided ST (runtime: ~85s)

Guided ST (runtime: ~273s)



Unguided ST (runtime: ~85s)

Guided ST (runtime: ~273s)

*runtime relative to V100 GPU

CHAPTER 5: CONCLUSION

- Algorithm implementation give good aesthetic results but takes a fairly big amount of time to run.

- User control with guided style can increase the control of the output by the user, but may not always generate better results for all applications.

- Runtime running with masks increases too much, making a excessive long runtime even in fairly powerful GPU's.

Impact of content weight alpha:

lr = 2e-3

average running time on Colab’s gpu: 74s

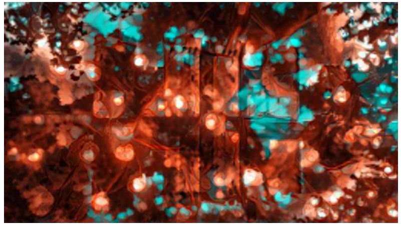
alpha = 1

alpha = 0.8

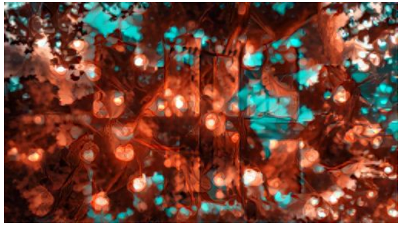
alpha = 0.6

alpha = 0.01

alpha = 0.4



alpha = 0.2



alpha = 0.1

THANK YOU FOR YOUR ATTENTION.
ANY QUESTIONS?

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom