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

Computational Imaging Project

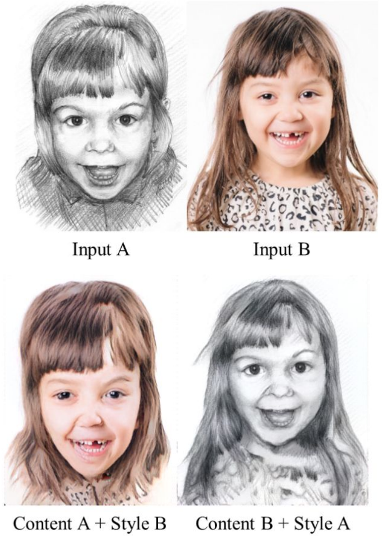
Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis

Team 5-a

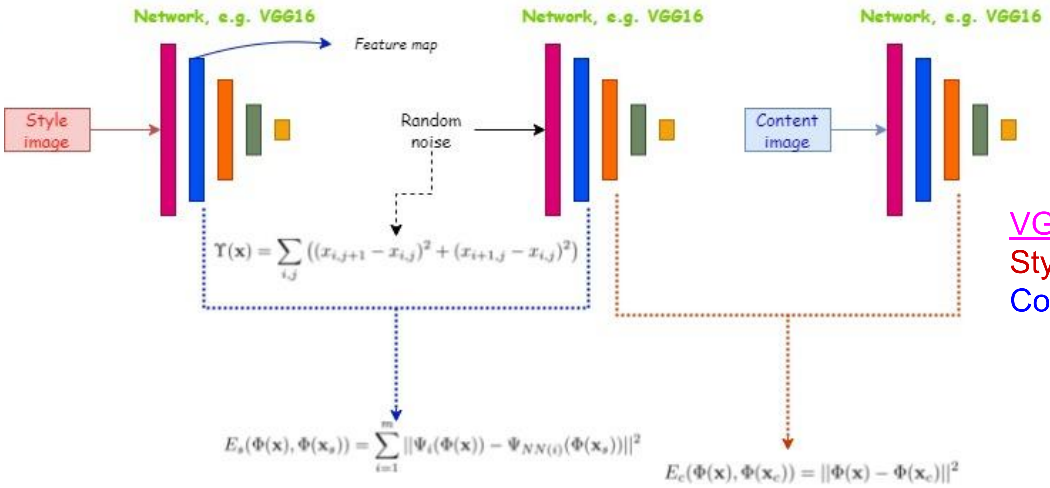
Content

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

Objective:
Synthesize an image x with its template guided by a content image x_c and having the textures from a style image x_s → A **minimization problem**:



$$\mathbf{x} = \arg\min_{x} E_s(\Phi(\mathbf{x}), \Phi(\mathbf{x}_s)) + \alpha_1 E_c(\Phi(\mathbf{x}), \Phi(\mathbf{x}_c)) + \alpha_2 \Upsilon(\mathbf{x})$$



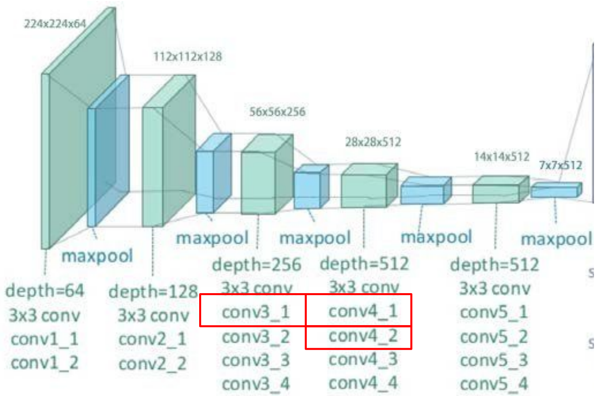
VGG-19
Style: relu3_1, relu4_1
Content: relu4_2

Example

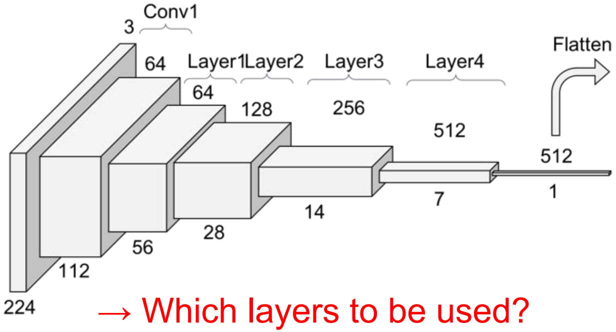
Visualization for better understanding of the objective function.

Idea:
Change pretrained VGG-19 to another pretrained network.

VGG-19, source: ***Research Gate***



Resnet-34, source: ***Toward Data Science***



→ Which layers to be used?

Layer name	Resnet-34	
conv1	7x7, 64, stride 2	
pool1	3x3, max pool, stride 2	
conv2_x	$3 \times 3, 64$ $3 \times 3, 64$	$\times 3$
conv3_x	$3 \times 3, 128$ $3 \times 3, 128$	$\times 4$
conv4_x	$3 \times 3, 256$ $3 \times 3, 256$	$\times 6$
conv5_x	$3 \times 3, 512$ $3 \times 3, 512$	$\times 3$
fc1	4x1, 512, stride 1	
pool time	1x10, avg pool, stride 1	
fc2	1x1, 50	

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

Idea:

Choose some relevant metrics for **quantitative** evaluation of synthesized images (*This is a bonus section beside the qualitatively perceptual metric).*

→ Wang, Zhizhong & Zhao, Lei & Chen, Haibo & Zuo, Zhiwen & Li, Ailin & Xing, Wei & Lu, Dongming. (2021). ***Evaluate and improve the quality of neural style transfer.*** *Computer Vision and Image Understanding. 207. 103203. 10.1016/j.cviu.2021.103203.*

3 metrics: Content fidelity (CF), Global effect (GE) and Local pattern (LP) → *Use simplified version*

Content fidelity

$$\boxed{CF(\vec{x},\vec{c}) = \frac{1}{N}\sum_{l=1}^{N}\frac{f_l(\vec{x})\cdot f_l(\vec{c})}{\|f_l(\vec{x})\|\cdot\|f_l(\vec{c})\|}}$$

Global effect

$$\boxed{GC(\vec{x},\vec{s}) = \frac{1}{3}\sum_{c=1}^{3}\frac{hist_c(\vec{x})\cdot hist_c(\vec{s})}{\|hist_c(\vec{x})\|\cdot\|hist_c(\vec{s})\|}}$$

$$HT(\vec{x},\vec{s}) = \frac{1}{N}\sum_{l=1}^{N}\frac{\mathcal{G}(f_l(\vec{x}))\cdot\mathcal{G}(f_l(\vec{s}))}{\|\mathcal{G}(f_l(\vec{x}))\|\cdot\|\mathcal{G}(f_l(\vec{s}))\|}$$

$$GE(\vec{x},\vec{s}) = \frac{1}{2}(GC(\vec{x},\vec{s}) + HT(\vec{x},\vec{s}))$$

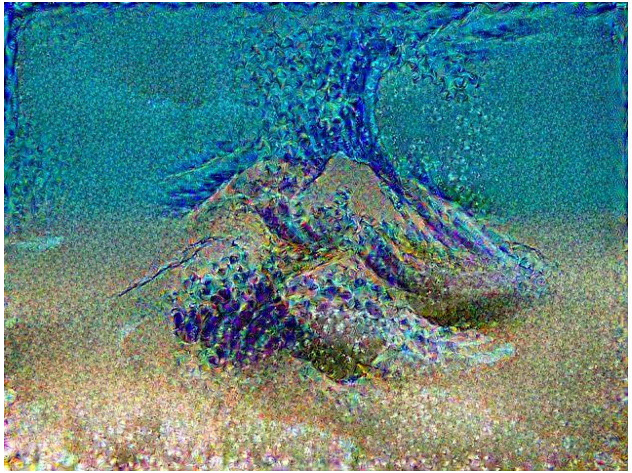
Local pattern

$$\boxed{LP_1(\vec{x},\vec{s}) = \frac{1}{Z}\sum_{l=1}^{N}\sum_{i=1}^{n_x}\frac{\Phi_i^l(\vec{x})\cdot\Phi_{CM(i)}^l(\vec{s})}{\|\Phi_i^l(\vec{x})\|\cdot\|\Phi_{CM(i)}^l(\vec{s})\|}}$$

$$LP_2(\vec{x},\vec{s}) = \frac{1}{N}\sum_{l=1}^{N}\frac{t_{cm}^l}{t_s^l}$$

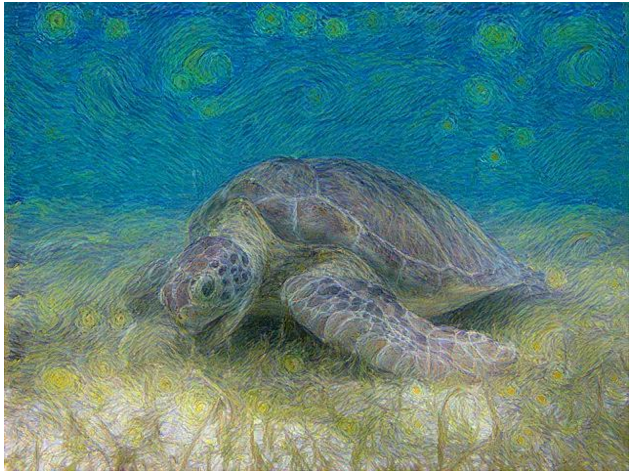
$$LP(\vec{x},\vec{s}) = \frac{1}{2}(LP_1(\vec{x},\vec{s}) + LP_2(\vec{x},\vec{s}))$$

Content + Style 1
CF = 0.78
GE = 0.76
LP = 0.0147



Content + Style 2
CF = 0.82
GE = 0.55
LP = 0.0149

Credit: *Code implementation was adapted and modified from the one of git user jonzhaocn*
https://github.com/jonzhaocn/cnnmrf-pytorch (and he/she used a different initialization)



Content + Style 1

CF = 0.78

GE = 0.76

LP = 0.0148

Content + Style 2

CF = 0.53

GE = 0.55

LP = 0.0143

Credit: *Code implementation was adapted and modified from the one of git user jonzhaocn* https://github.com/jonzhaocn/cnnmrf-pytorch (and he/she used a different initialization)

Content

Style

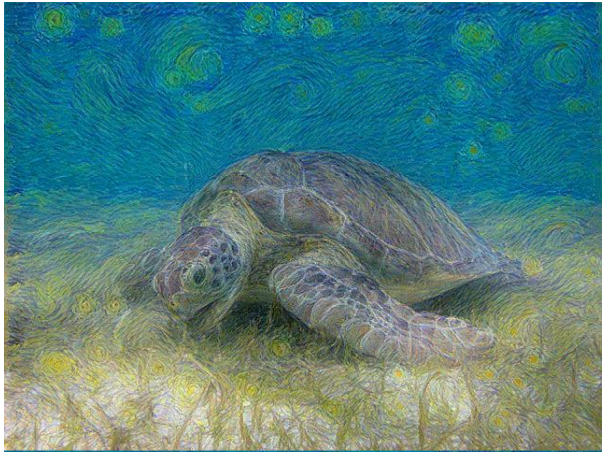
Style 1

Style 2

VGG-19 results

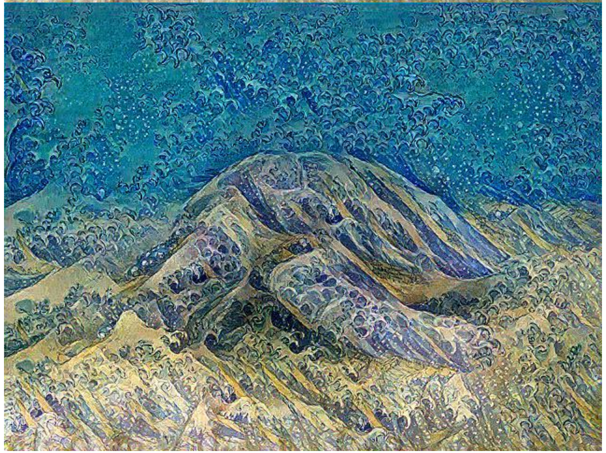
Content + Style 1

CF = 0.781866
GE = 0.7619
LP = 0.014755



Content + Style 2

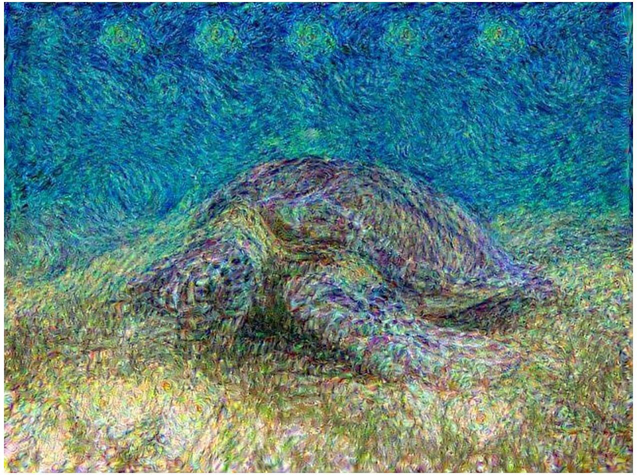
CF = 0.531313
GE = 0.55236
LP = 0.014287



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

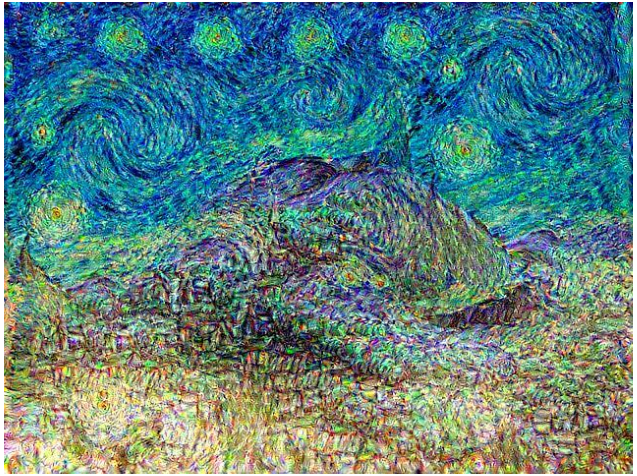
Content + Style 1

Results on different layers, **style_weight = 0.5**, **number of iterations = 60**, **number of resolutions = 3, style patch size = 3, stride = 1**



Content = conv4_5
content_weight = 0.3
Style = conv3_1
CF = 0.83
GE = 0.77
LP = 0.0151

Content = conv4_5
content_weight = 0.3
Style = conv4_1
CF = 0.88
GE = 0.60
LP = 0.0084

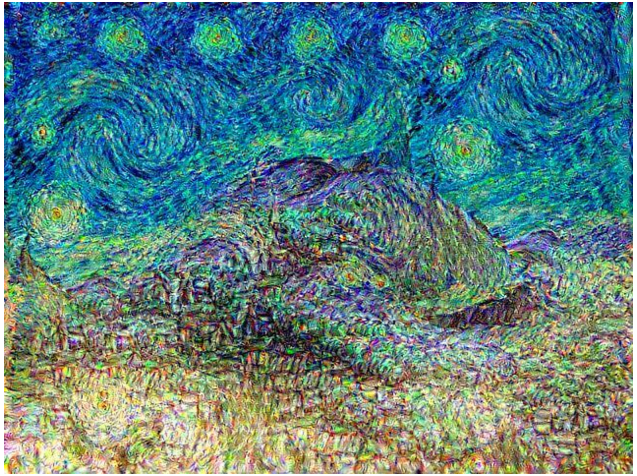


Content = conv4_5
content_weight = 0.3
Style = conv3_1, conv4_1
CF = 0.68
GE = 0.80
LP = 0.0161

Content + Style 1

Results on different layers, **style_weight = 0.5**, **number of iterations = 60**, **number of resolutions = 3, style patch size = 3, stride = 1**

Style weight variation



Content = conv4_5
content_weight = 0.3
Style = conv3_1, conv4_1
CF = 0.68
GE = 0.80
LP = 0.0161

Content = conv4_5
content_weight = 0.4
Style = conv3_1, conv4_1
CF = 0.78
GE = 0.76
LP = 0.0147

Content = conv4_5
content_weight = 0.5
Style = conv3_1, conv4_1
CF = 0.83
GE = 0.68
LP = 0.0137

Content + Style 1

Results on different layers, **style_weight = 0.5**, **number of iterations = 60**, **number of resolutions = 3, style patch size = 3, stride = 1**

Style patch size + Stride variation

CF = 0.95
GE = 0.58
LP = 0.0126

Content = conv5_3
Style = conv4_2, conv4_5

CF = 0.95
GE = 0.58
LP = 0.0035

Content = conv5_3
Style = conv4_1, conv4_2
stride = 3

Content + Style 1

Results on different layers, **style_weight = 0.1**, **number of iterations = 60**, **number of resolutions = 3, style patch size = 3, stride = 2**

Style patch size + Stride variation

CF = 0.95
GE = 0.57
LP = 0.0021

Content = conv5_3
Style = conv4_2, conv4_5

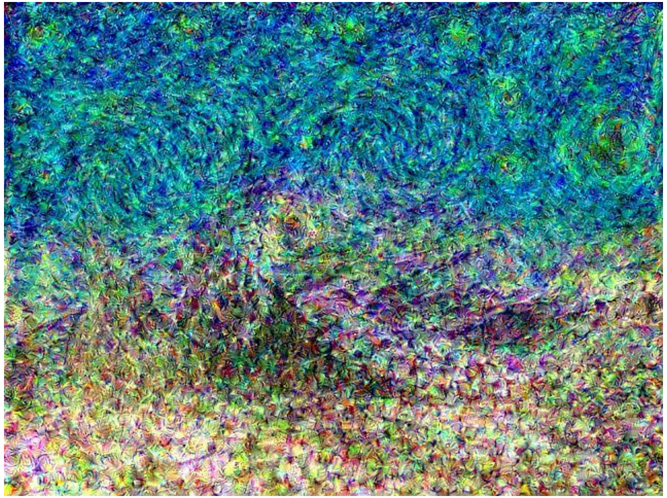
CF = 0.95
GE = 0.57
LP = 0.0022

Content = conv5_3
Style = conv4_1, conv4_2
style patch size = 5
stride = 1

Content + Style 1

Some other extreme cases (some parameters are abnormal)

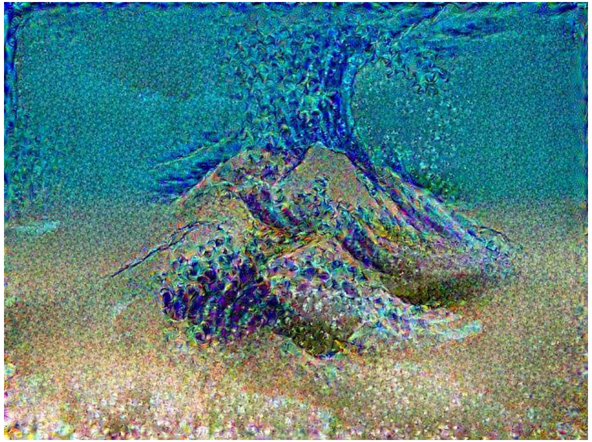
Content = conv4_5
Style = conv4_1, conv4_2
number of iterations =250
CF = 0.67
GE = 0.59
LP = 0.0242



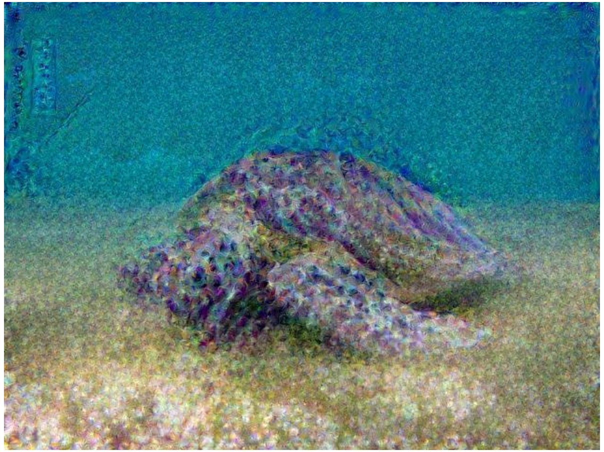
Content = conv4_5
Style = conv4_1, conv4_2
number of resolution = 4
CF = 0.68
GE = 0.68
LP = 0.0221

Content + Style 2

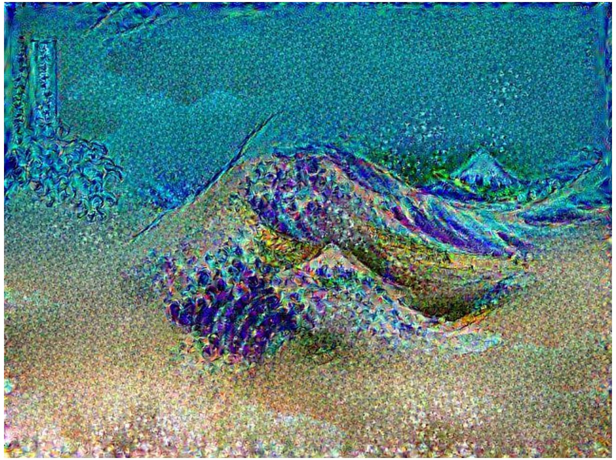
Results on different layers, α_1 **= 0.5 number of iterations = 150**, **number of resolutions = 3, style patch size = 3, stride = 1**



Content = conv4_5
Style = conv3_1, conv4_1
CF = 0.82
GE = 0.55
LP = 0.0149

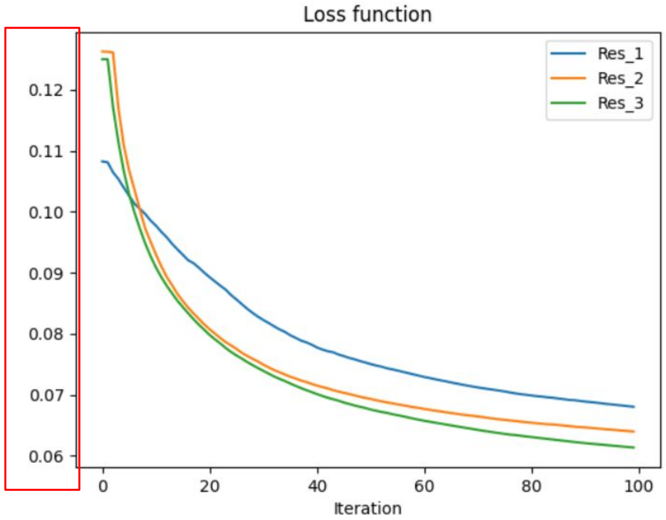


Content = conv4_5
Style = conv3_2, conv4_3
CF = 0.84
GE = 0.48
LP = 0.0125

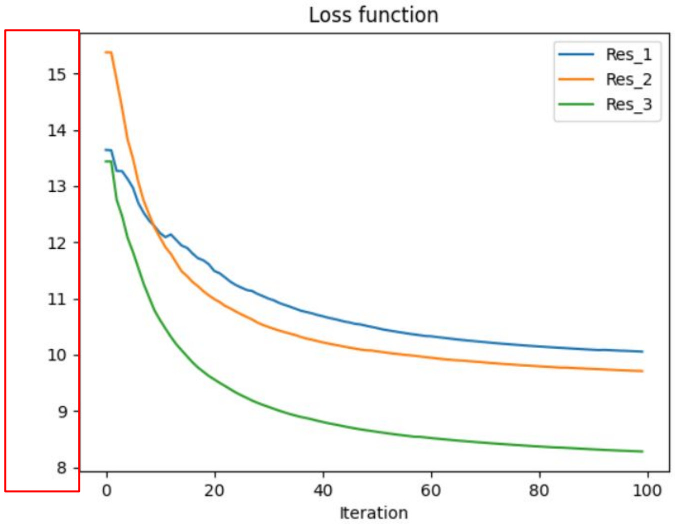


Content = conv4_5
iterations = 100
Style = conv3_1, conv3_2, conv4_4
CF = 0.73
GE = 0.60
LP = 0.0137

LOSS FUNCTION EVOLUTION



Resnet34



VGG19

$$\arg\min_x E_s(\Phi(\mathbf{x}), \Phi(\mathbf{x}_s)) + \alpha_1 E_c(\Phi(\mathbf{x}), \Phi(\mathbf{x}_c)) + \alpha_2 \Upsilon(\mathbf{x})$$

Thank you for your attention

Team 5-a