Style Transfer for Videos

KU-BIG

Contents

1. Paper Review
2. Code Review
3. Results so far & Further plans

KU-BIG

Paper Review

KU-BIG

Style Transfer

***P**: content image*

+

***A**: style image*

=

***X**: synthesized image*

"with pre-trained model"

Main points

- 어떤 CNN모델을 사용할 것인가
- CNN모델의 어느 레이어에서 feature를 추출할 것인가
- Loss는 어떻게 관리할 것인가

Which CNN model to use

ConvNet Configuration						NAME:
A	A-LRN	B	C	D	E	
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers	
input (224 × 224 RGB image)						
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv1_1 conv1_2
maxpool						pool1
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv2_1 conv2_2
maxpool						pool2
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**	conv3_1 conv3_2 conv3_3 conv3_4
maxpool						pool3
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**	conv4_1 conv4_2 conv4_3 conv4_4
maxpool						pool4
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**	conv5_1 conv5_2 conv5_3 conv5_4
maxpool						pool5
FC-4096						fc6
FC-4096						fc7
FC-1000						fc8
soft-max						

VGG 19

16 convolution layers, 5 pooling layers,

3 connected layers

⇩

Feature map

⇩

Computing Style & Content loss

How to extract feature from various layers

Style Recon.

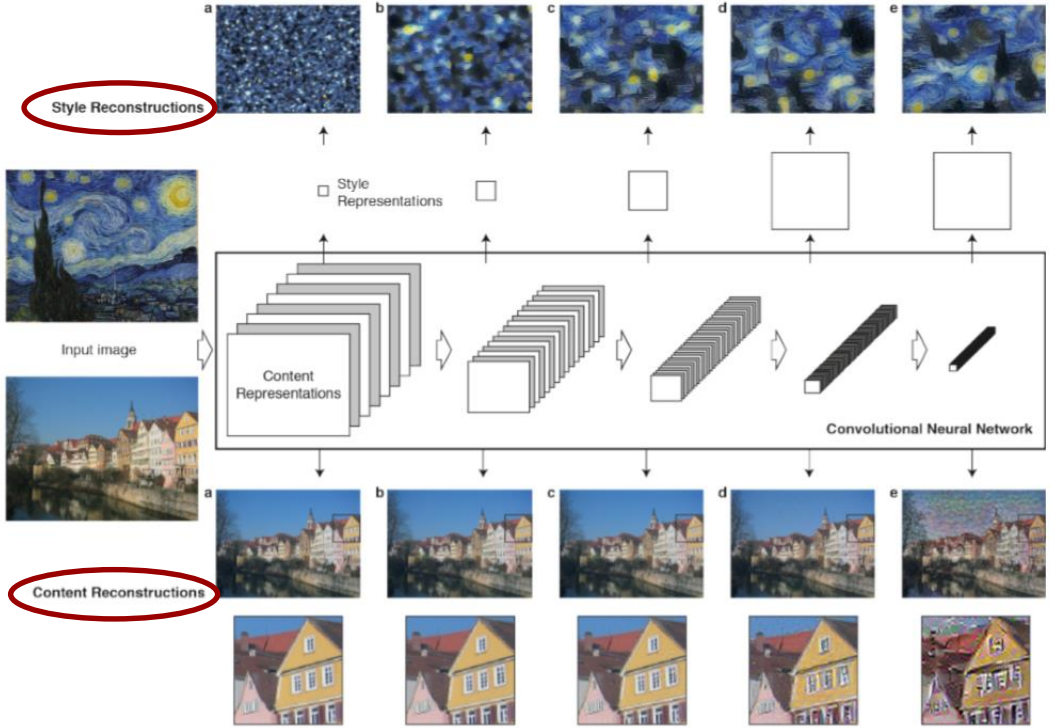
Lower layer : Content ↓, Texture ↑

Higher layer : Content ↑, Texture ↓

Content Recon.

Lower layer : Contain Original image

Higher layer : Remain High level content



How to deal with Loss

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

Lower $\frac{\alpha}{\beta}$

More focused on Style, rather than Content

Higher $\frac{\alpha}{\beta}$

More focused on Content, rather than Style

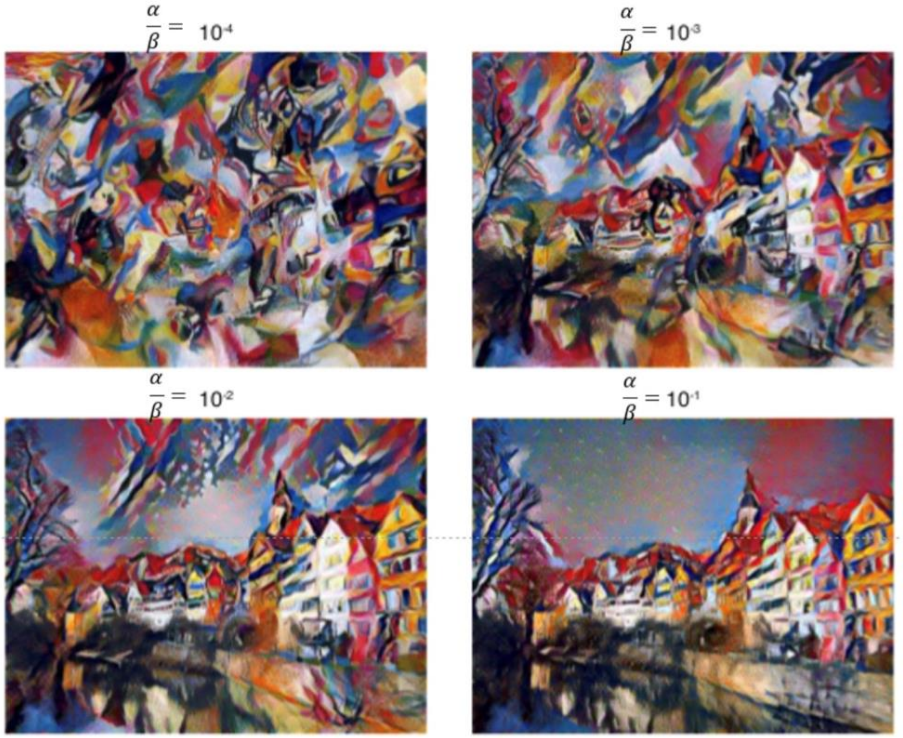


그림. 가중치 변경에 따른 합성 이미지 x의 변화

Code Review

Code Review

Load pre -trained model → Load Content Image → Load Style Image

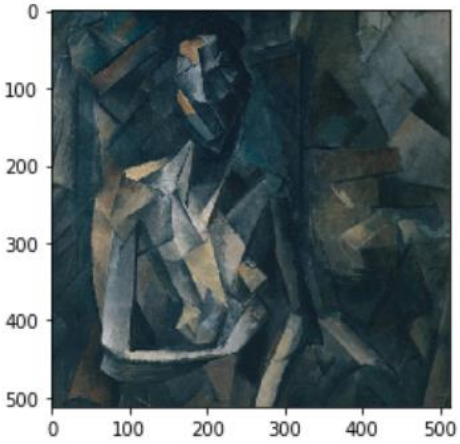
Define Content Loss → Define Style Loss → Extract Content/Style Feature

Define Synthesized Image → Proceed Learning

Results so far & Further plans

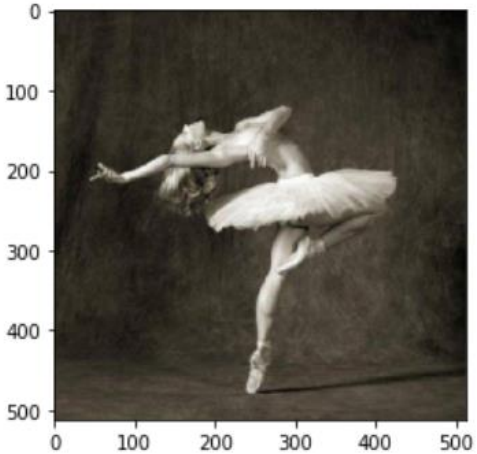
Results so far

Style Image



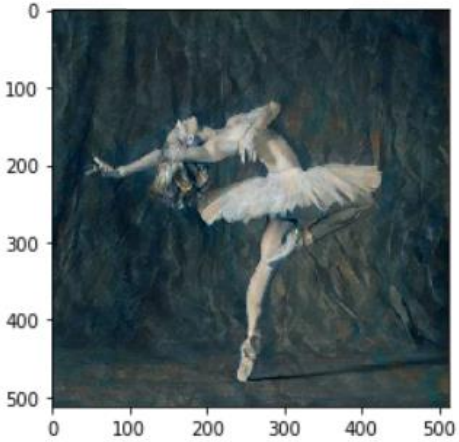
+

Content Image

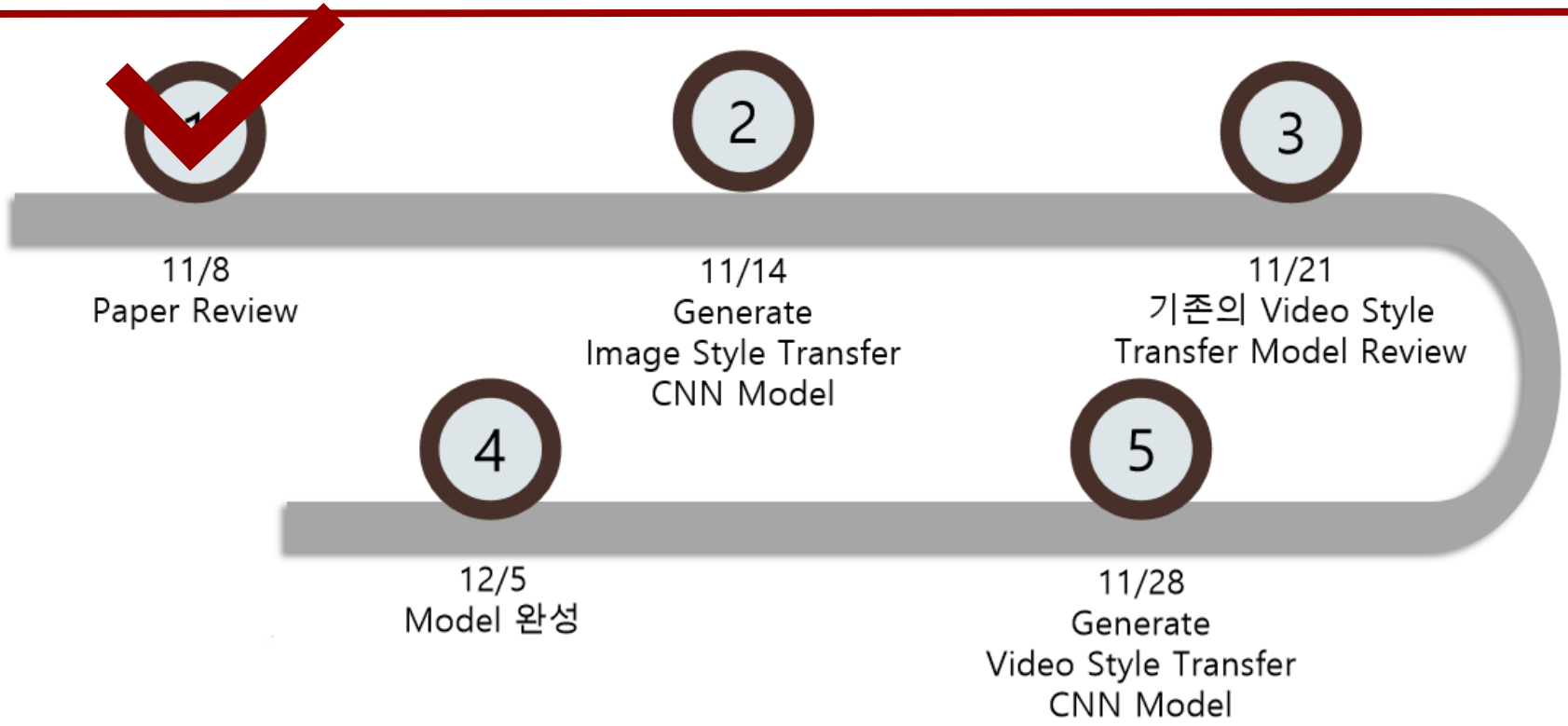


=

Output Image



Further plans



Thanks a lot