2021 8th NAFOSTED Conference on Information and Computer Science (NICS)

Deep Feature Rotation for Multimodal Image Style Transfer

Son Truong Nguyen, Nguyen Quang Tuyen, Nguyen Hong Phuc

Outline

1. Introduction
2. Motivation
3. Proposed Method
4. Experimental results
5. Future works

Input

Output

1.
Introduction

*What is **Image style transfer?***

Image Style Transfer

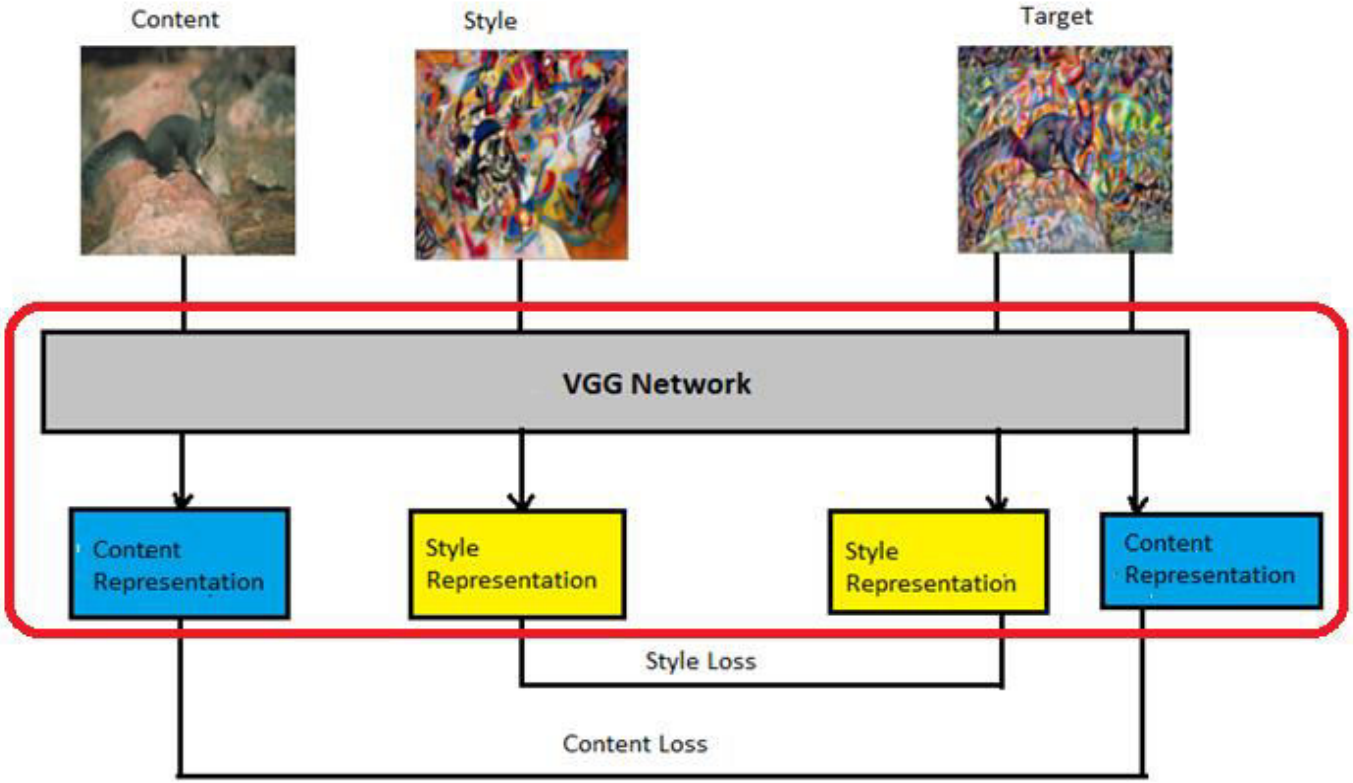
Content image

Style Image

Output image

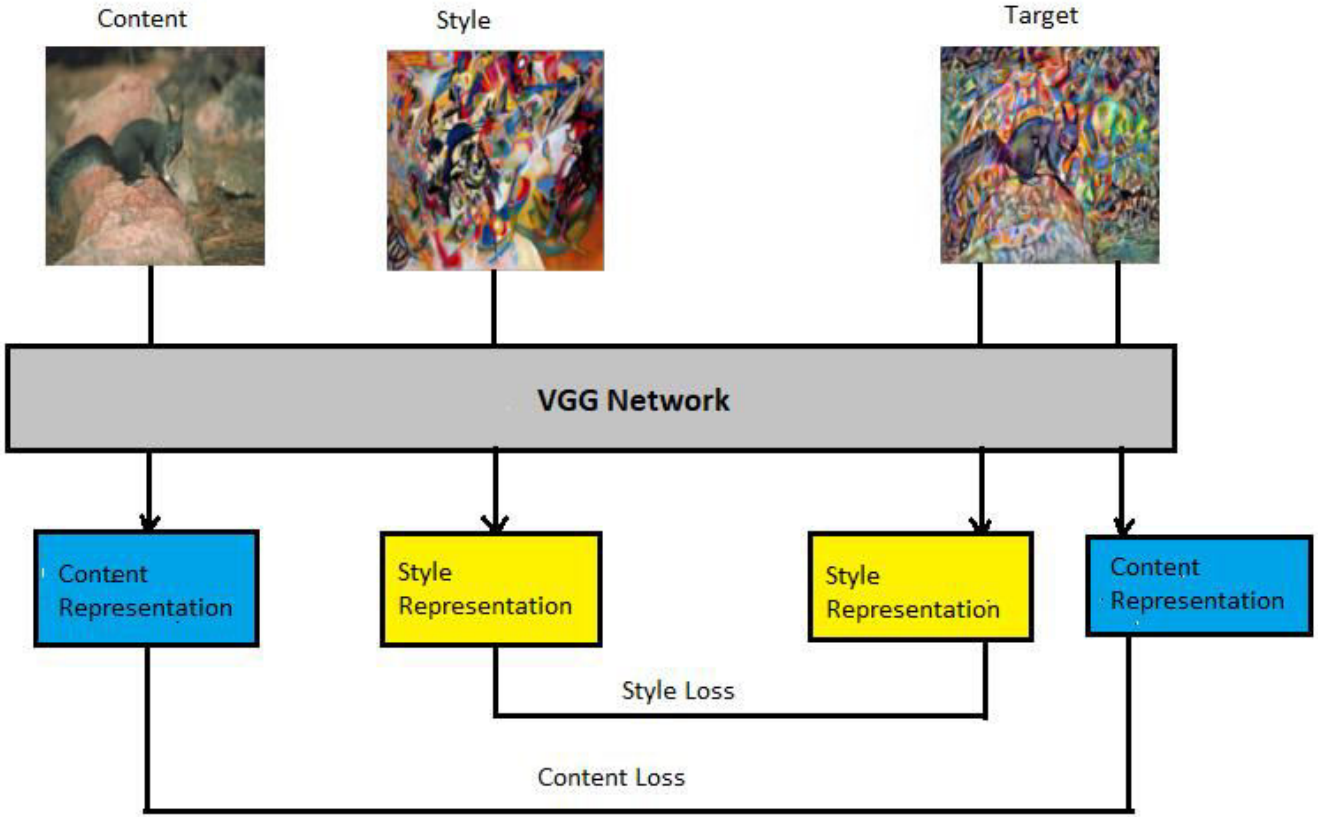
2.
Motivation

Previous Approaches



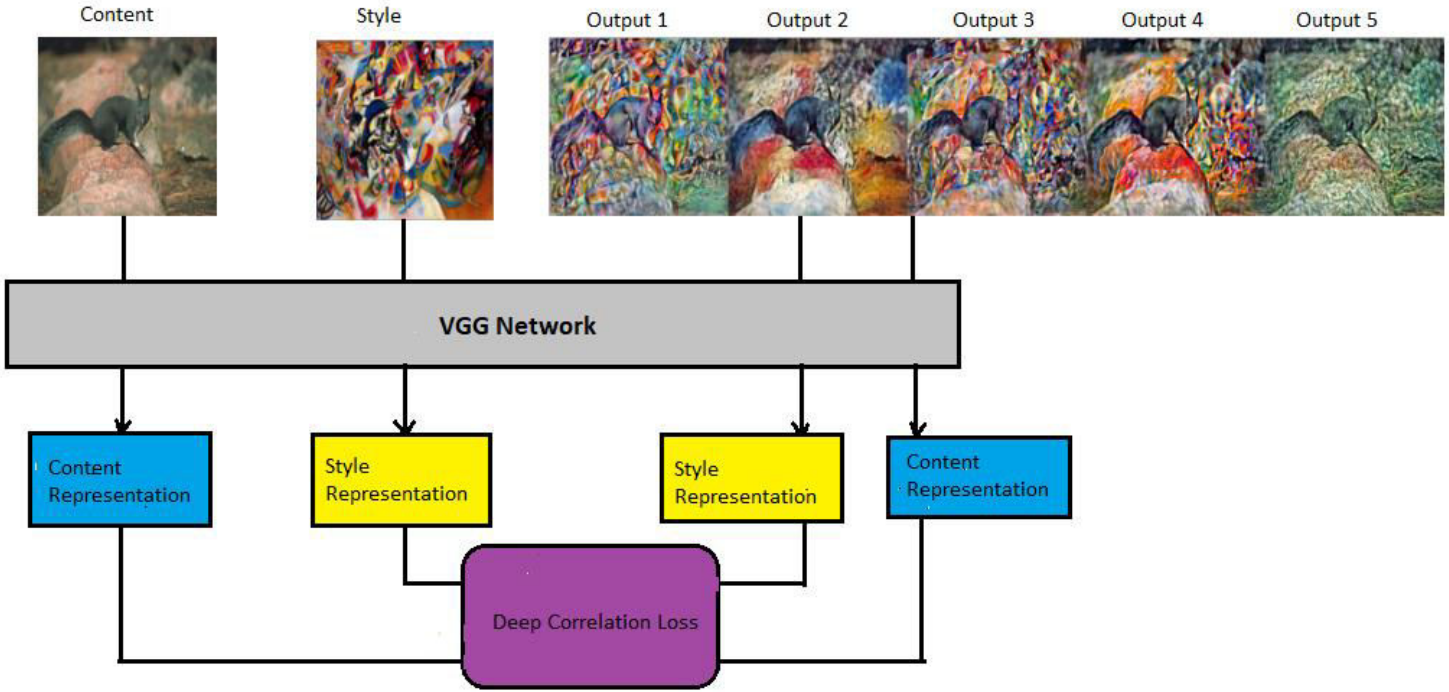
[1] Image Style Transfer using convolutional neural network, Gatys et al., CVPR 2016

Previous Approaches



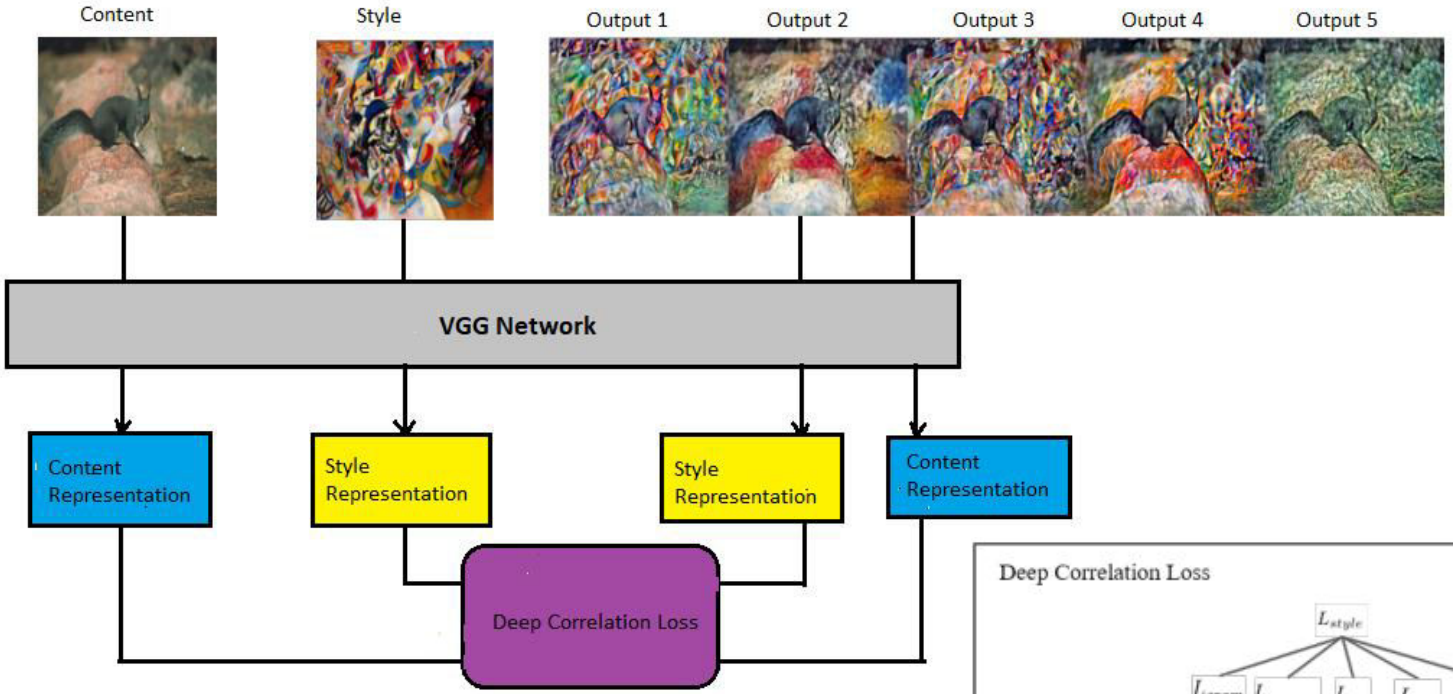
Only one output can be generated for each content and style image pair

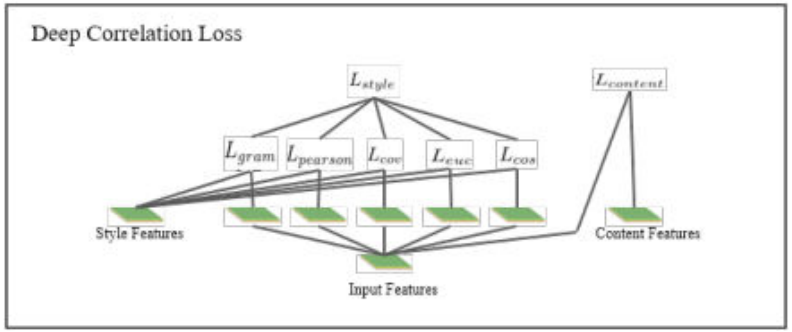
Previous Approaches



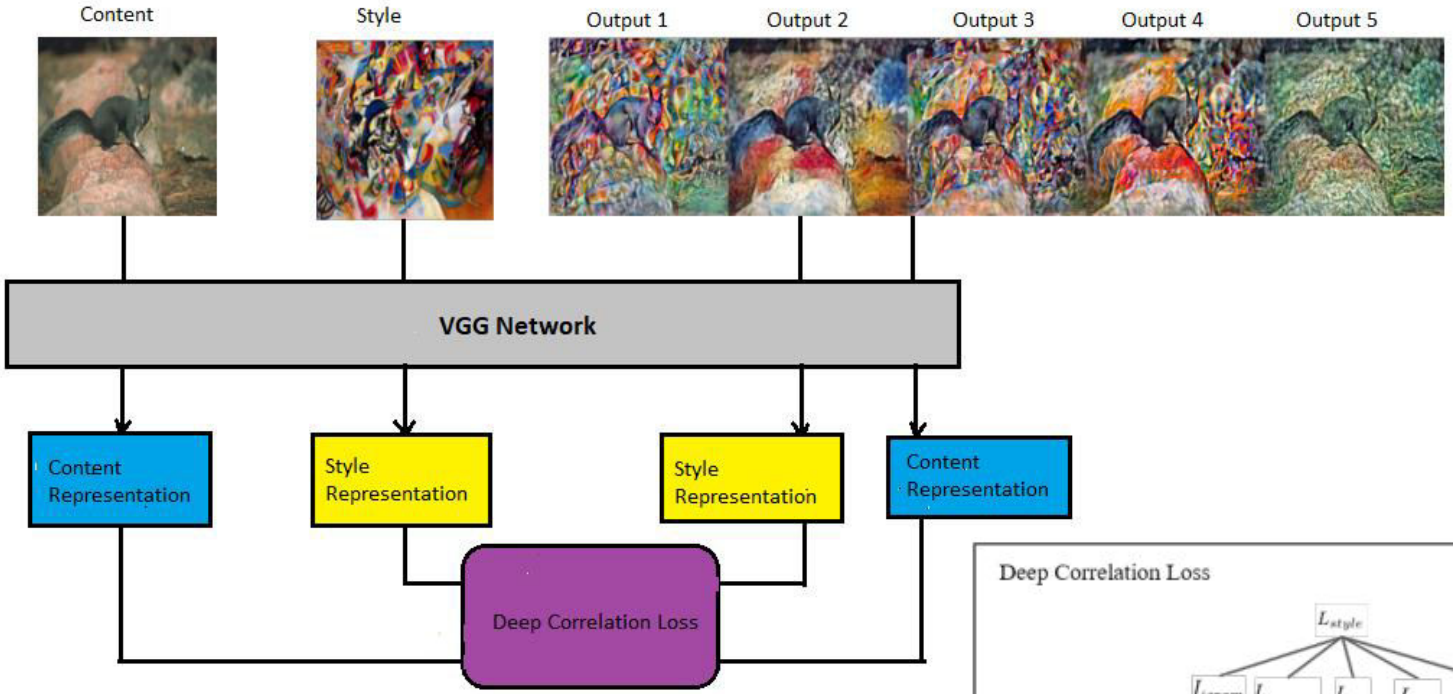
[1] Deep Correlation Multimodal Neural Style Transfer, Tuyen et al., IEEE Access

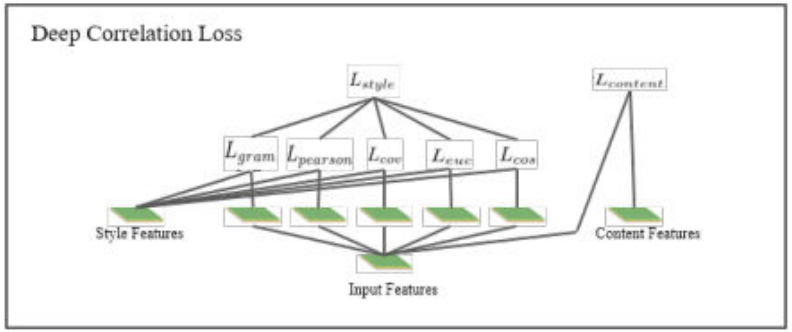
Previous Approaches





Previous Approaches





Slow down the back-propagation process, making it difficult to use in reality.

3.
Proposed Method

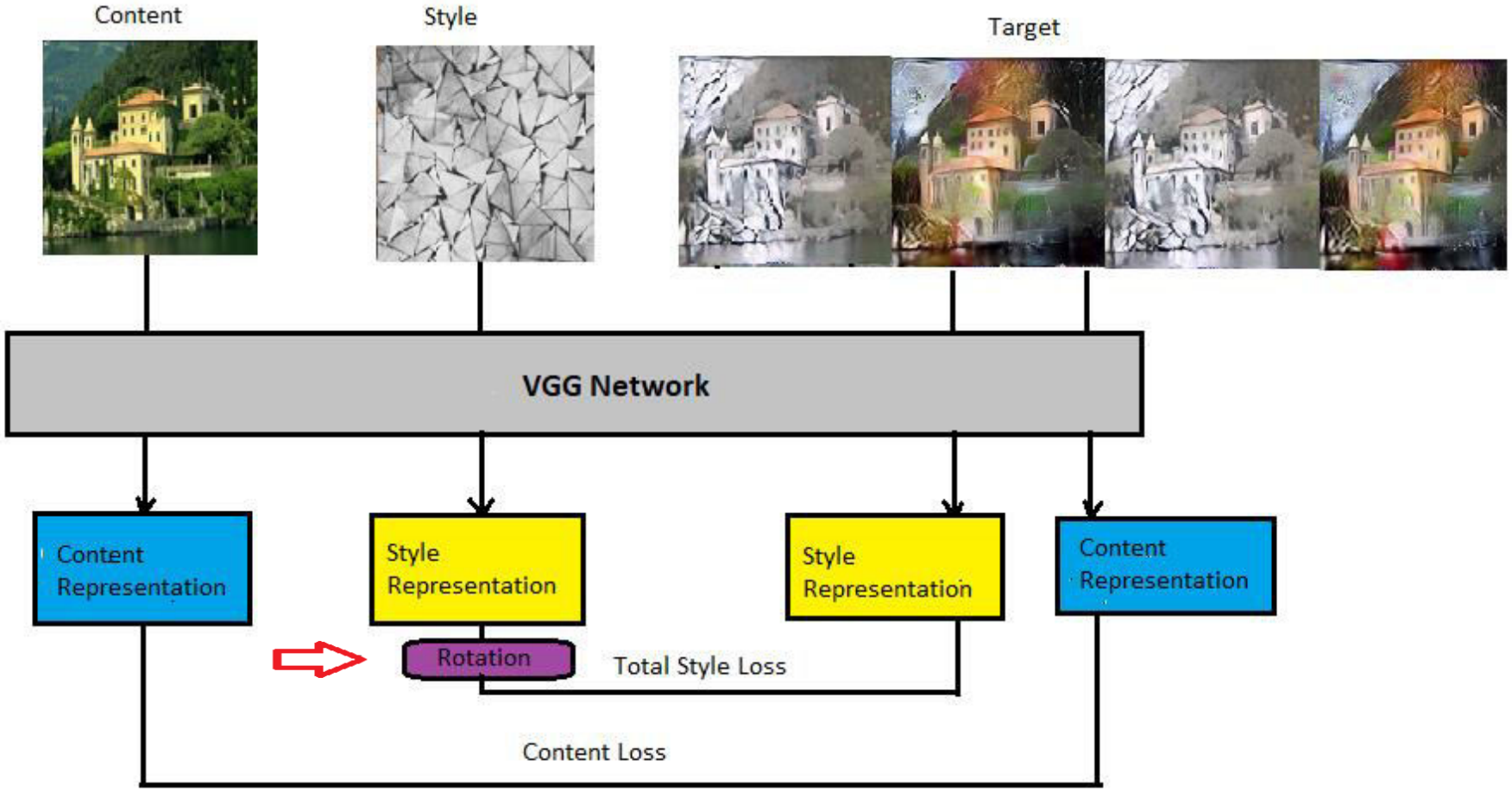
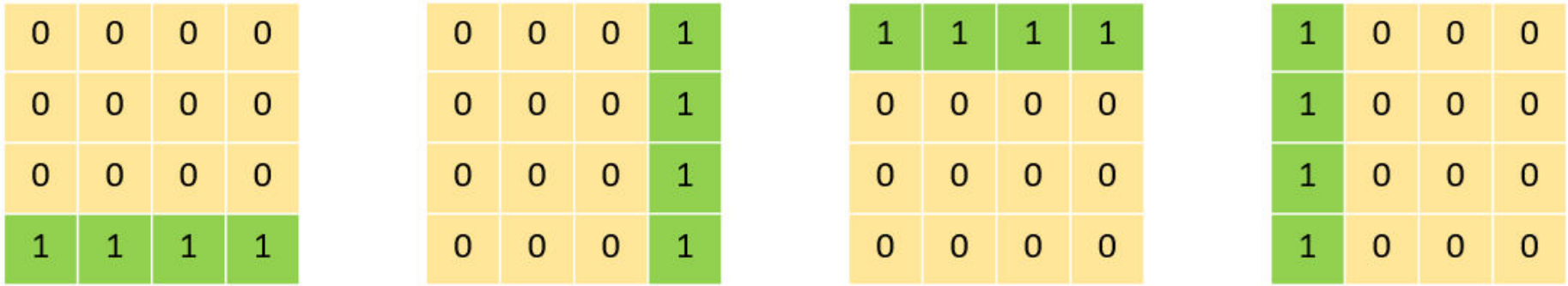

Content
Style
Target
VGG Network
Content Representation
Style Representation
Style Representation
Content Representation
Rotation
Total Style Loss
Content Loss

Illustration of Rotation Mechanism



(a) 0 (b) 90 (c) 180 (d) 270

- With 90 degrees:

$$W_r = W_{i,j,q,k}, \forall\ 0 \le i \le w, \\ 0 \le j \le h, 0 \le q \le c, 0 \le k \le n$$

- With 180 degrees:

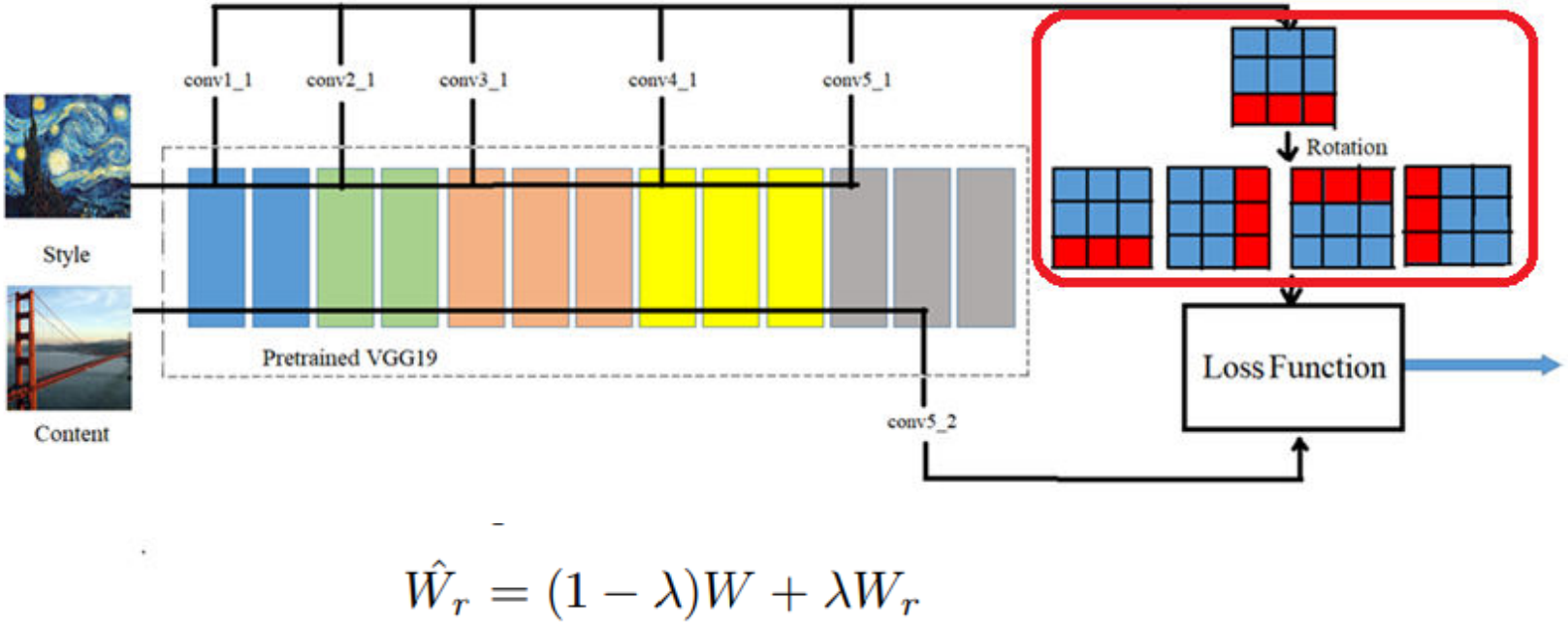
$$W_r = W_{i,j,q::-1,k}, \forall\ 0 \le i \le n, \\ 0 \le j \le c, 0 \le q \le h, 0 \le k \le w$$

- With 270 degrees:

$$W_r = W_{i,j,q::-1,k}, \forall\ 0 \le i \le w, \\ 0 \le j \le h, 0 \le q \le c, 0 \le k \le n$$

No need to care about statistical values or complicated calculation steps.

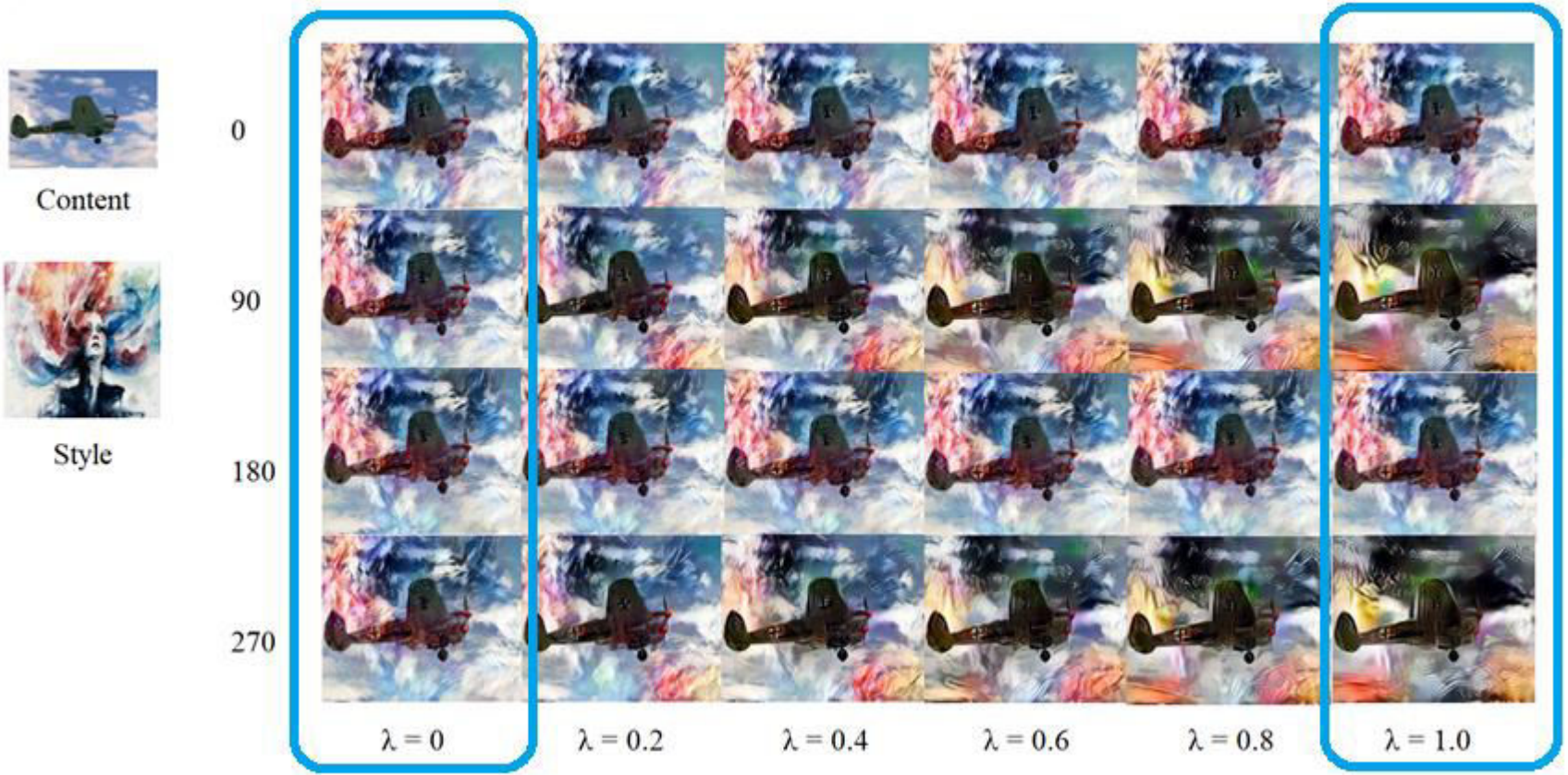
End-to-end style transfer model



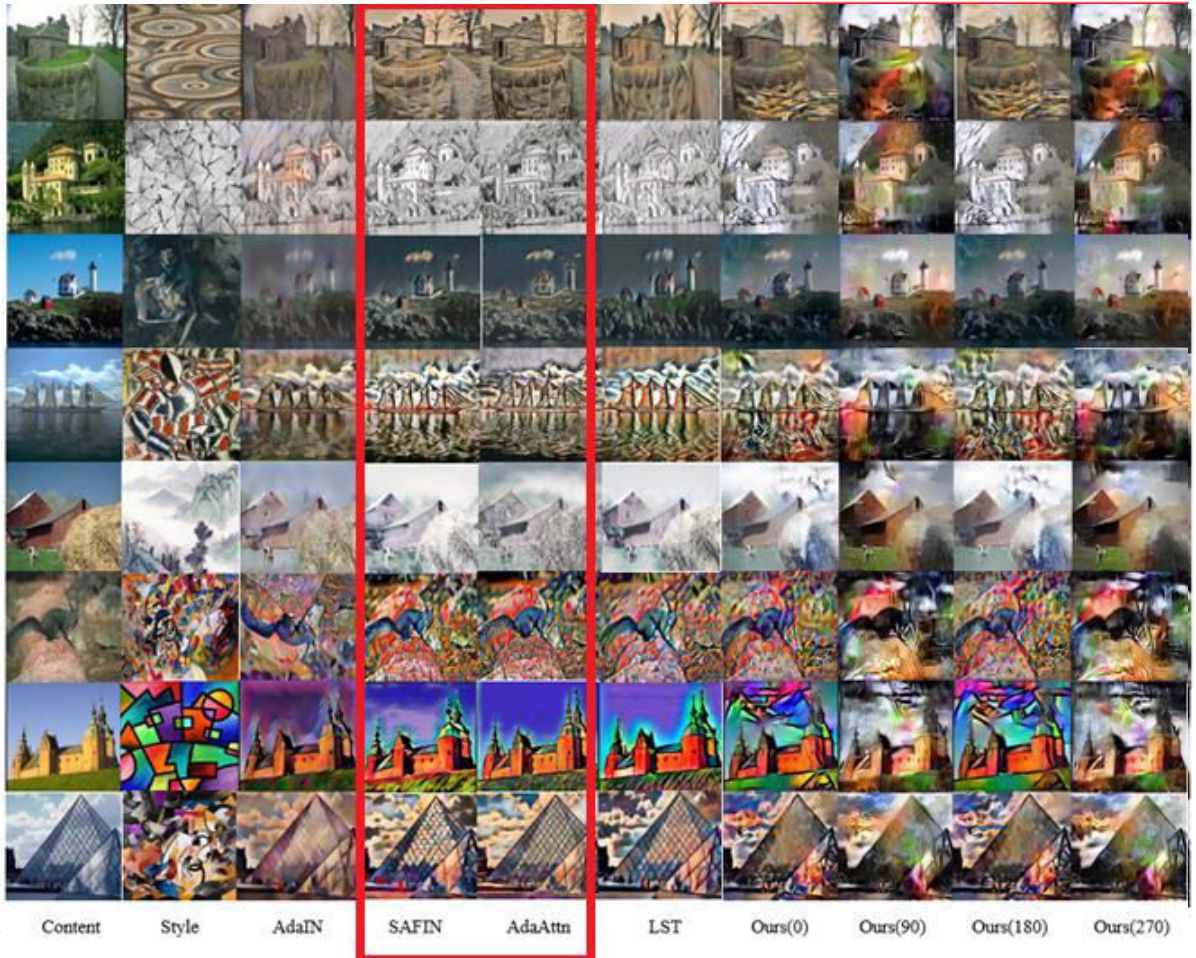
$$\hat{W}_r = (1-\lambda)W + \lambda W_r$$

4.
Experimental Results

Qualitative Result

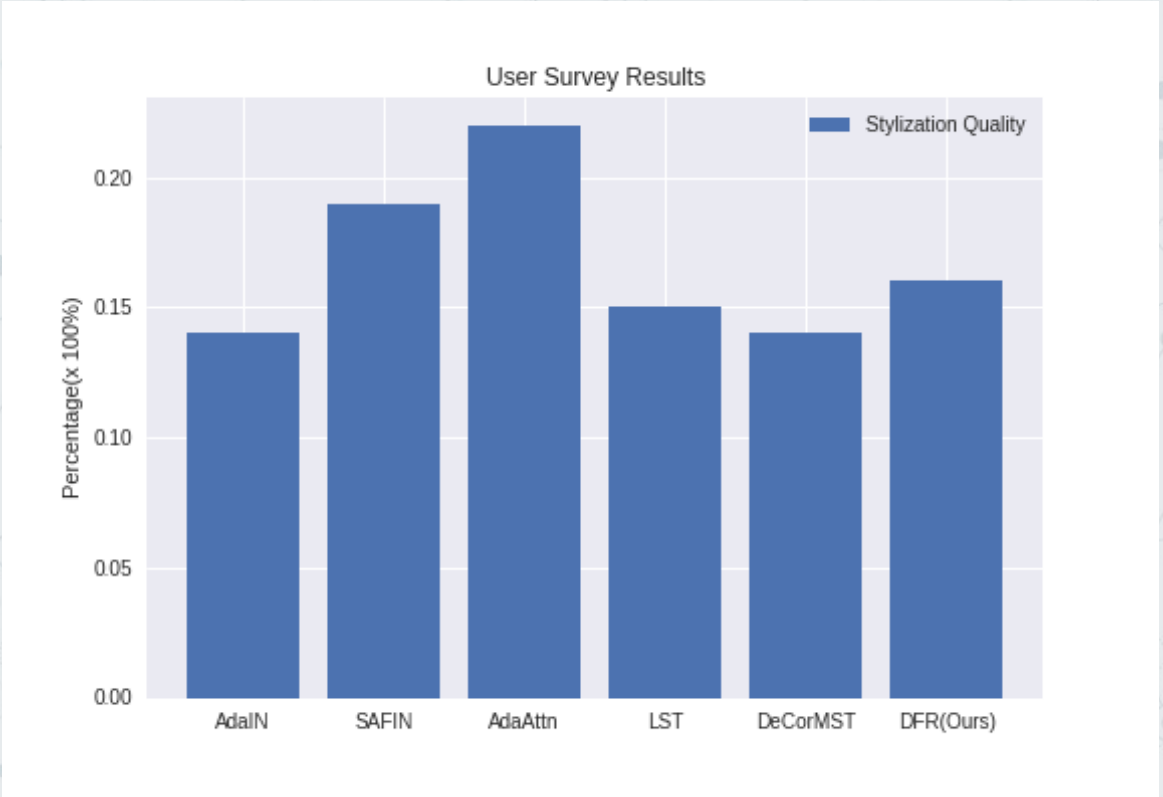


Comparison with State-of-the-Art Methods



User Study

- **15** content images from MSCOCO dataset, and **20** style images from WikiArt dataset.
- **600** votes from **30** people



Efficiency

Method	DeCorMST	DFR-1	DFR-2	DFR-3	DFR-4
Time	3 days	361s	377s	396s	417s

RUNNING TIME COMPARISON BETWEEN MULTIMODAL METHODS.

DFR-X equivalent to the number of outputs generated

.

5.
Future Work

Generate more diverse images

Produces better quality images

Reduce inference time

Thanks!

Source code:

https://github.com/sonnguyen129/deep-feature-rotation

Any questions?