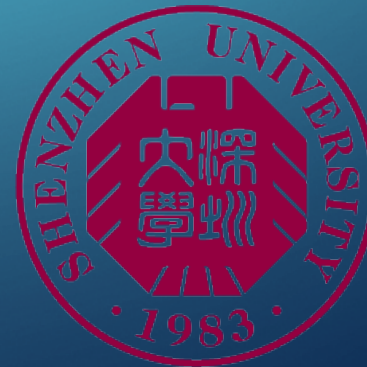


NOVEL NETWORK ARCHITECTURES FOR ARBITRARY IMAGE STYLE TRANSFER

MINGLUN GONG, SCHOOL OF COMPUTER SCIENCE, UNIVERSITY OF GUELPH

COLLABORATORS: ZHIJIE WU, CHUNJIN SONG, YANG ZHOU, AND HUI HUANG

SHENZHEN UNIVERSITY



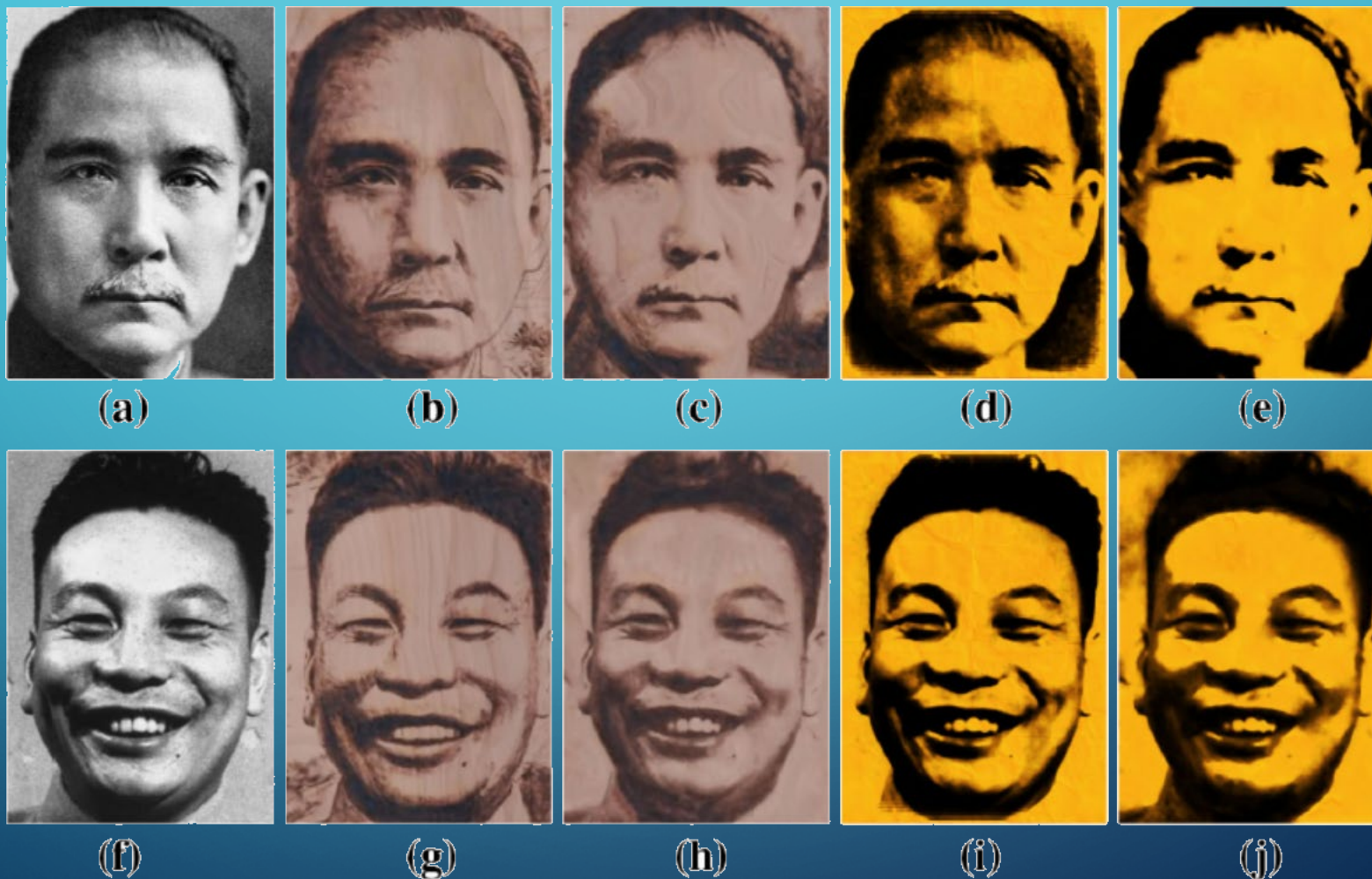
BACKGROUND

- **October 2017:**
 - My PhD student Zili Yi presented his work on image-to-image translation at ICCV
 - Zili Yi, Hao Zhang, Ping Tan, & Minglun Gong: DualGAN: Unsupervised dual learning for image-to-image translation. IEEE International Conference on Computer Vision. Venice, Italy, October 22-29, 2017
- **September 2018:**
 - I visited Visual Computing Center, Shenzhen University
 - Gave a talk titled “Two routes for image-to-image translation: Rule-based vs. learning-based”

DUANGAN FOR UNSUPERVISED IMAGE TRANSLATION



STYLIZATION THROUGH IMAGE TRANSLATION



BACKGROUND (CONT'D)

- After the talk, Zijie Wu explained his novel ideas for image stylization
- November 2018:
 - Submitted the first algorithm to CVPR 2019, but was later rejected.
- March 2019:
 - Resubmitted the first algorithm to ICCV 2019, but was rejected again.
- May 2019:
 - Submitted the second algorithm to NIPS 2019 and was accepted.
- September 2019:
 - Resubmitted the first algorithm to AAAI 2020 and was accepted.

WHAT IS ARBITRARY STYLE TRANSFER?



Content



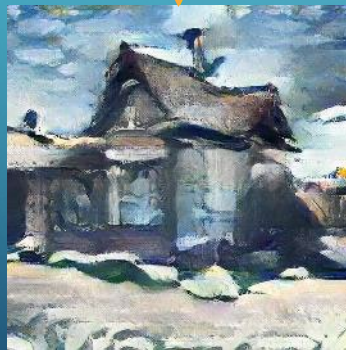
Style



...



synthesize



...



PIONEER WORK

Content



Style

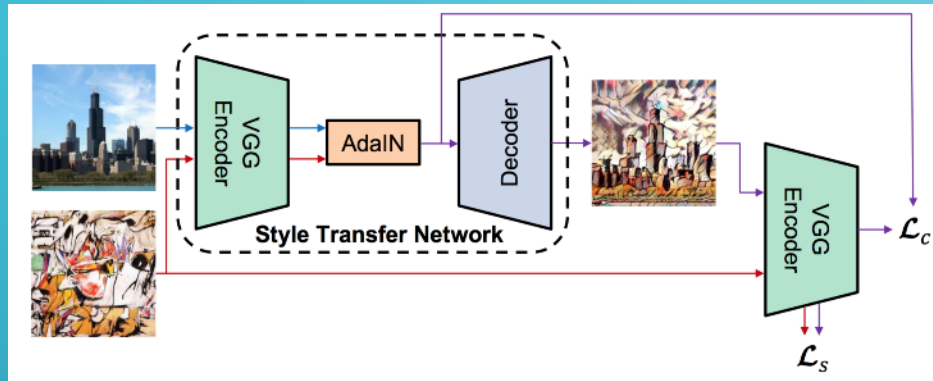


$Loss$ — $\begin{cases} L_{content} \\ L_{style} \end{cases}$

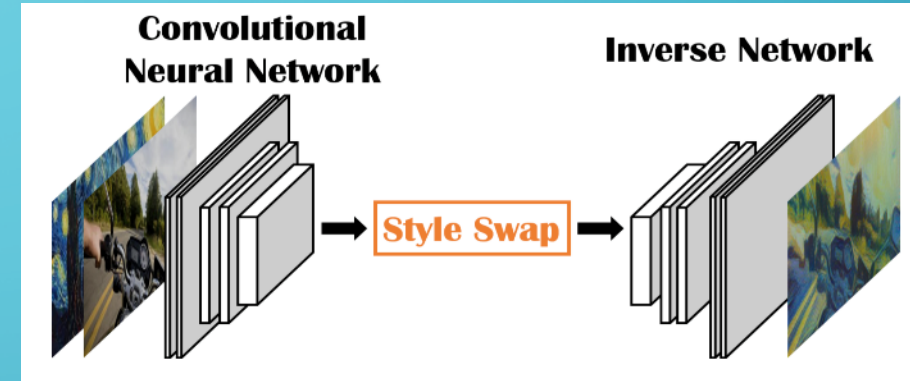


[Gatys et al., CVPR 2016]

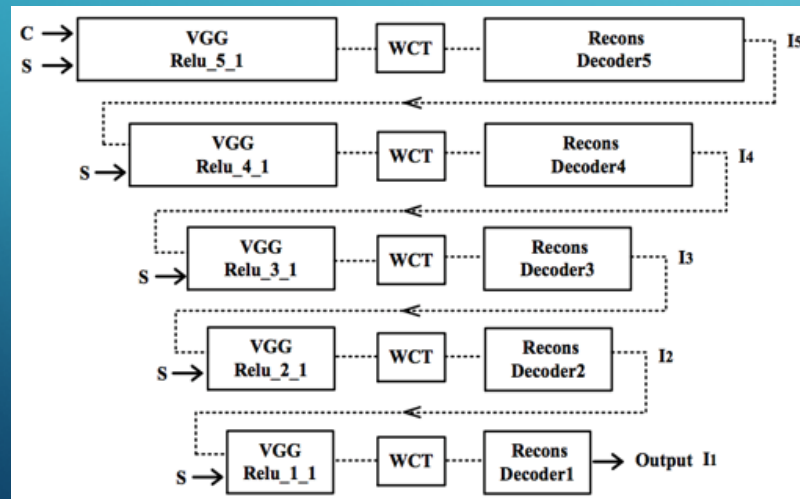
OTHER EXISTING WORK



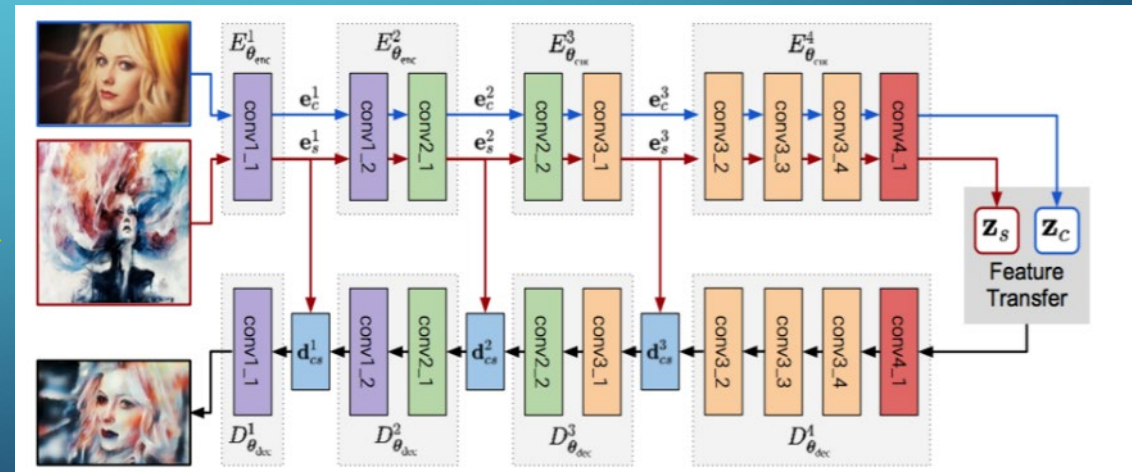
[AdaIN ICCV 2017]



[StyleSwap NeurIPS 2016]



[WCT NeurIPS 2017]



[AvatarNet CVPR 2018]

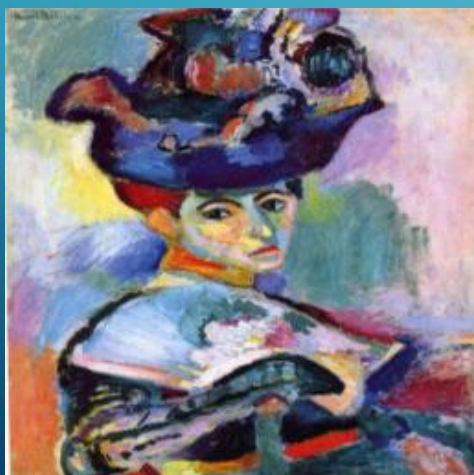
EXCHANGEABLE FEATURE ALIGNMENT NETWORK

AAAI 2020 – “EFANet: Exchangeable Feature Alignment Network for Arbitrary Style Transfer”

OBSERVATION AND MOTIVATION



Content



Style



[AdaIN 2017]



[WCT 2017]



[AvatarNet 2018]

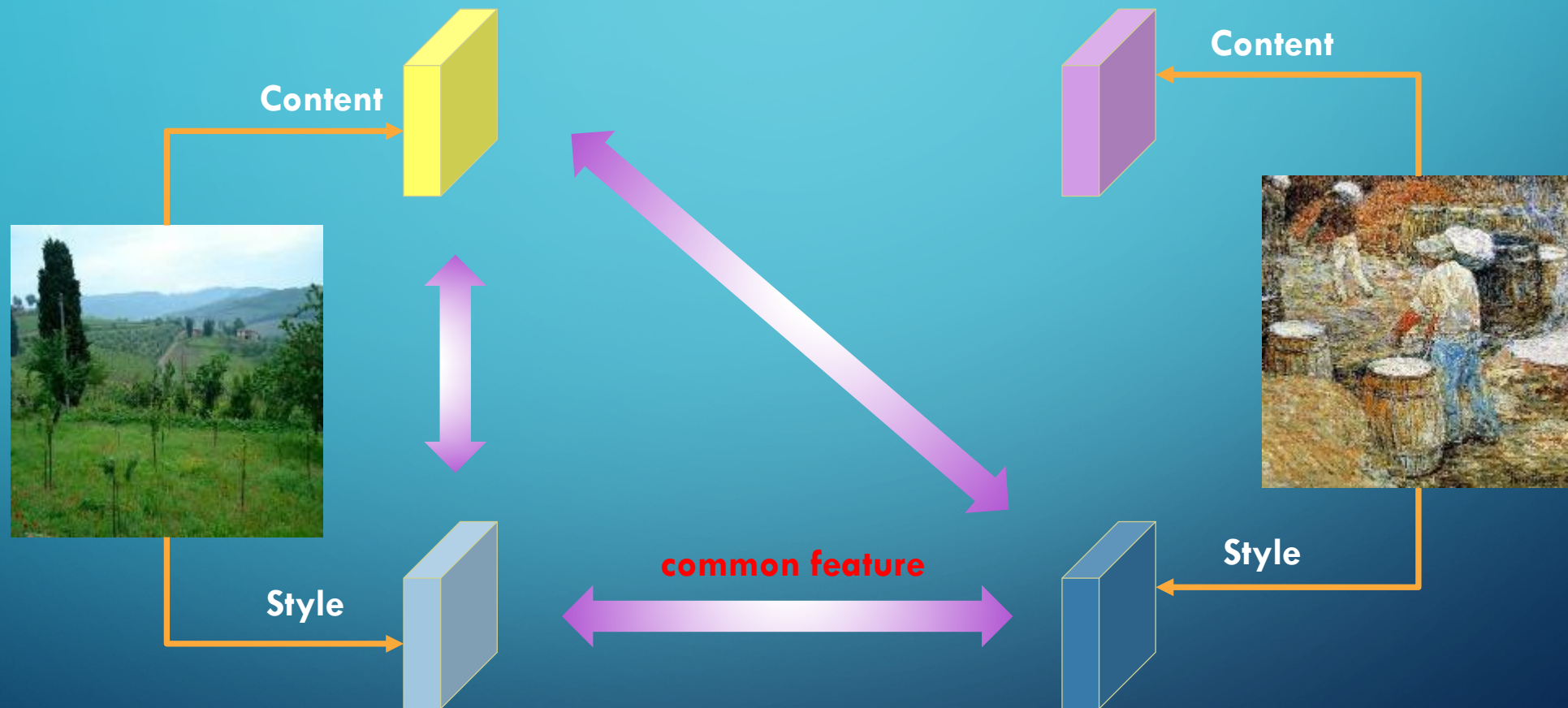


[SANet 2019]

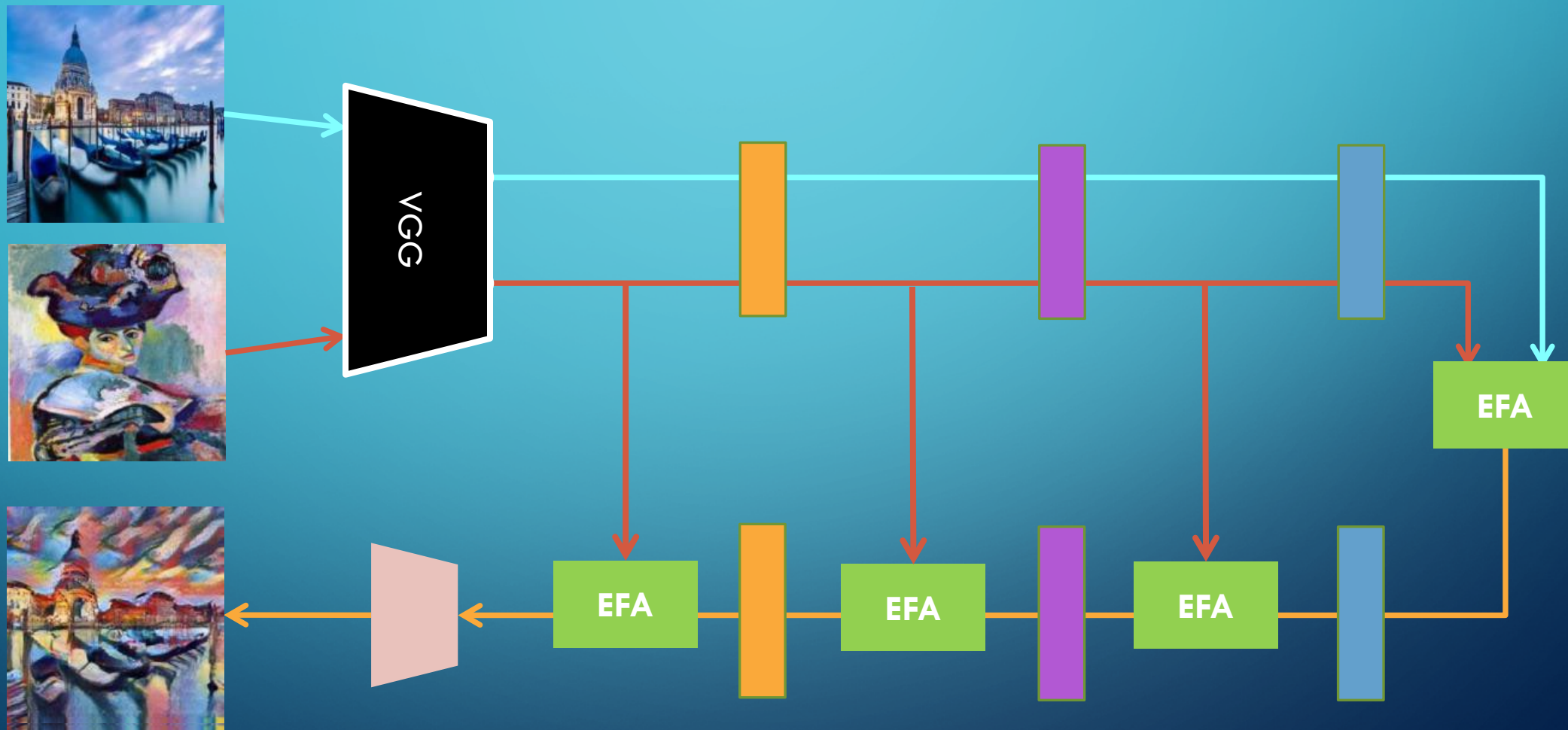


[EFTNet]

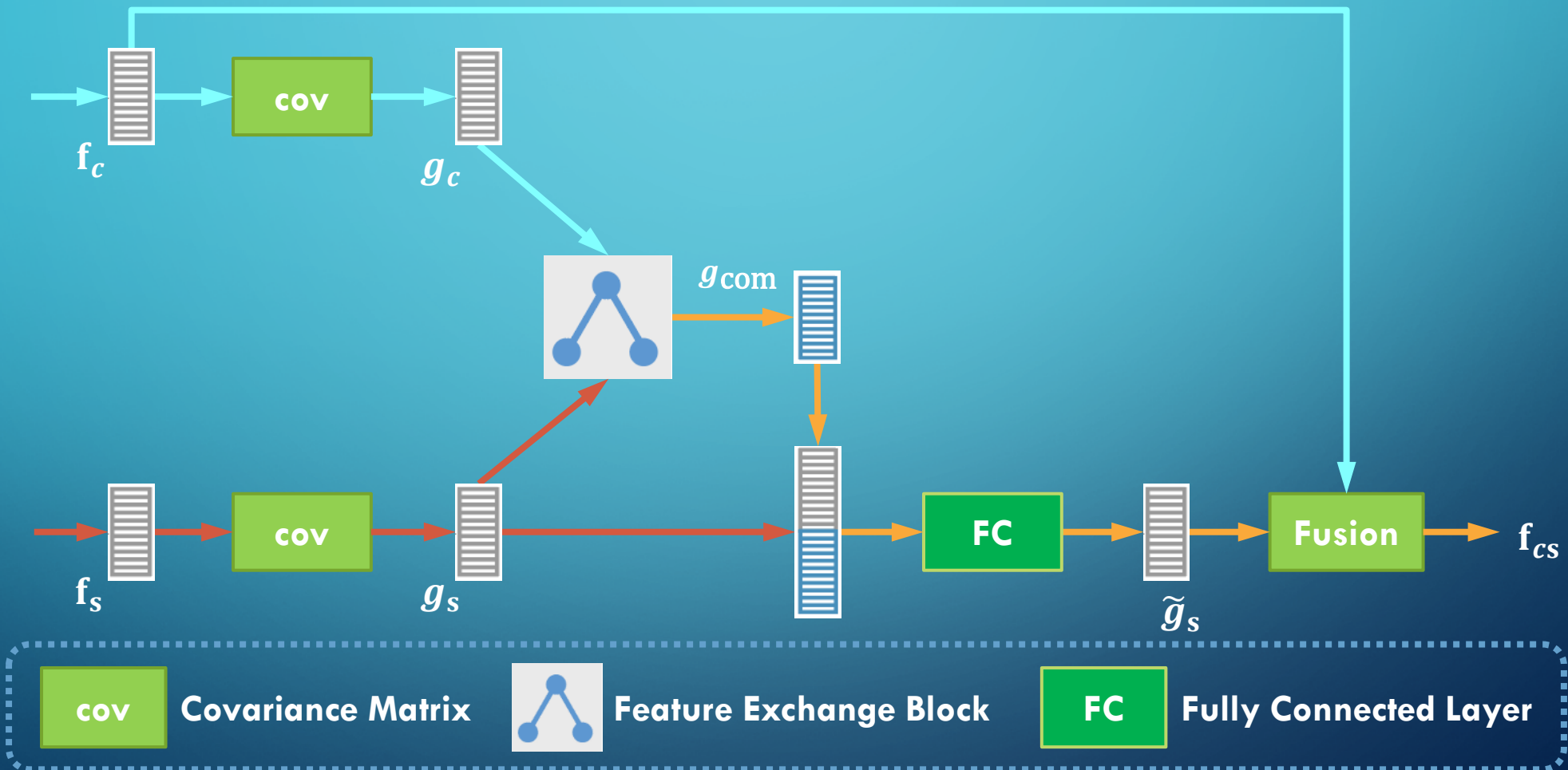
JOINTLY ANALYZE CONTENT-STYLE IMAGE PAIR



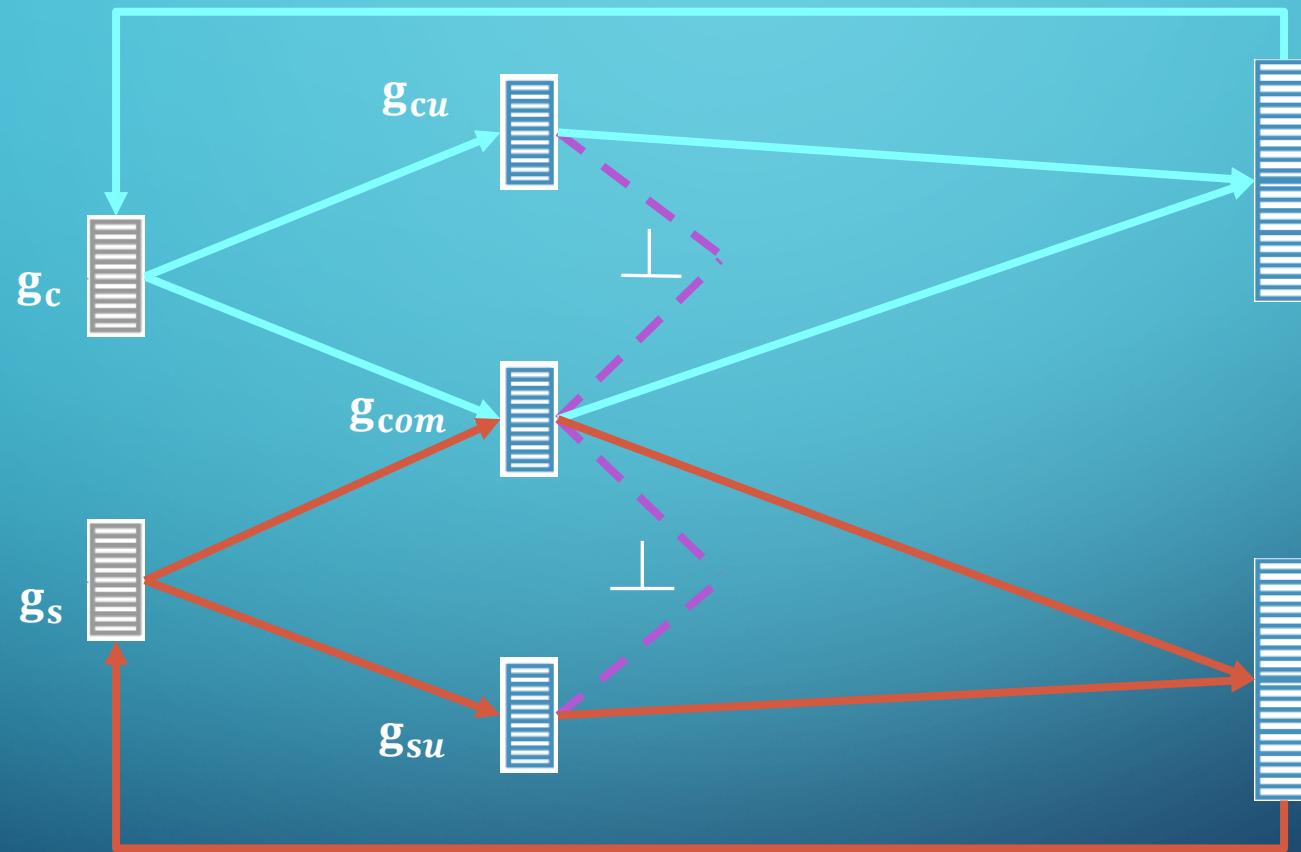
ARCHITECTURE OVERVIEW



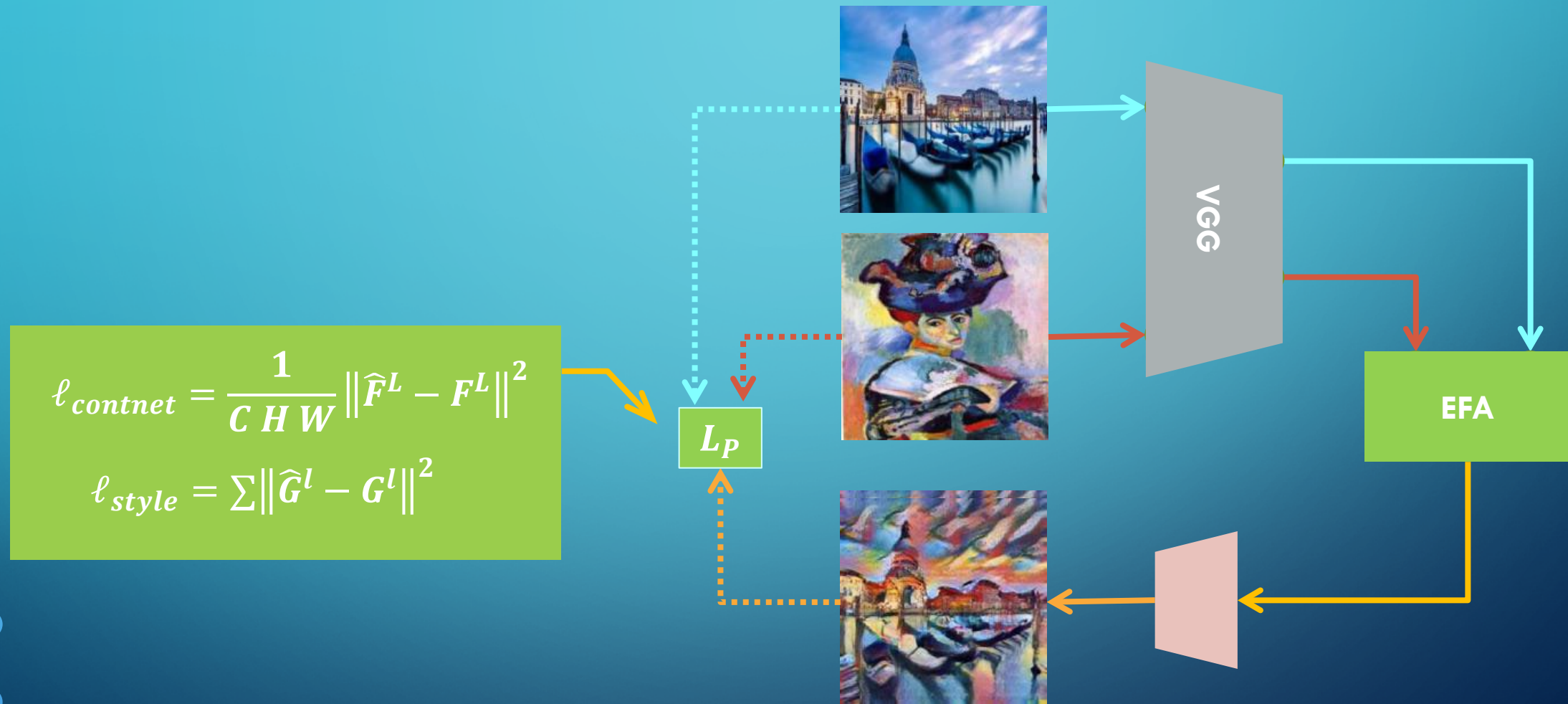
EXCHANGEABLE FEATURE ALIGNMENT (EFA) MODULE



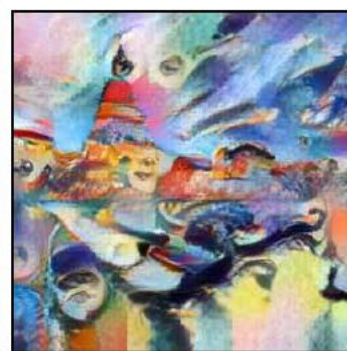
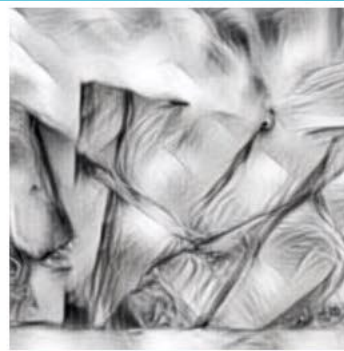
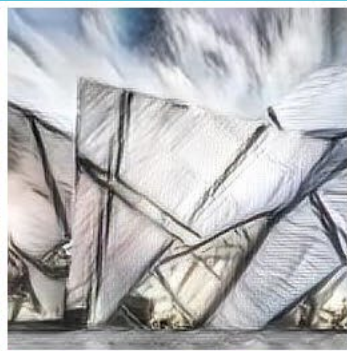
FEATURE EXCHANGE BLOCK



LOSS FUNCTION



EXPERIMENTAL RESULTS



Style

Content

[AdaIN 2017]

[WCT 2017]

[AvatarNet 2018]

[SANet 2019]

EFANet

EXPERIMENTAL RESULTS (CONT'D)



Style

Content

[AdaIN 2017]

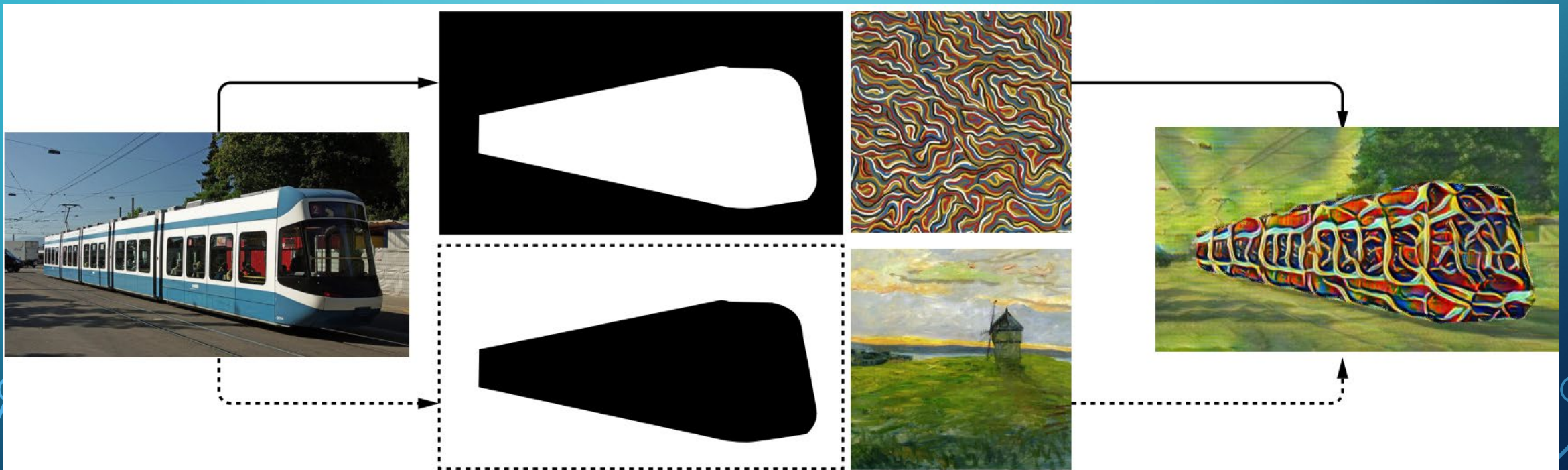
[WCT 2017]

[AvatarNet 2018]

[SAnet 2019]

EFANet

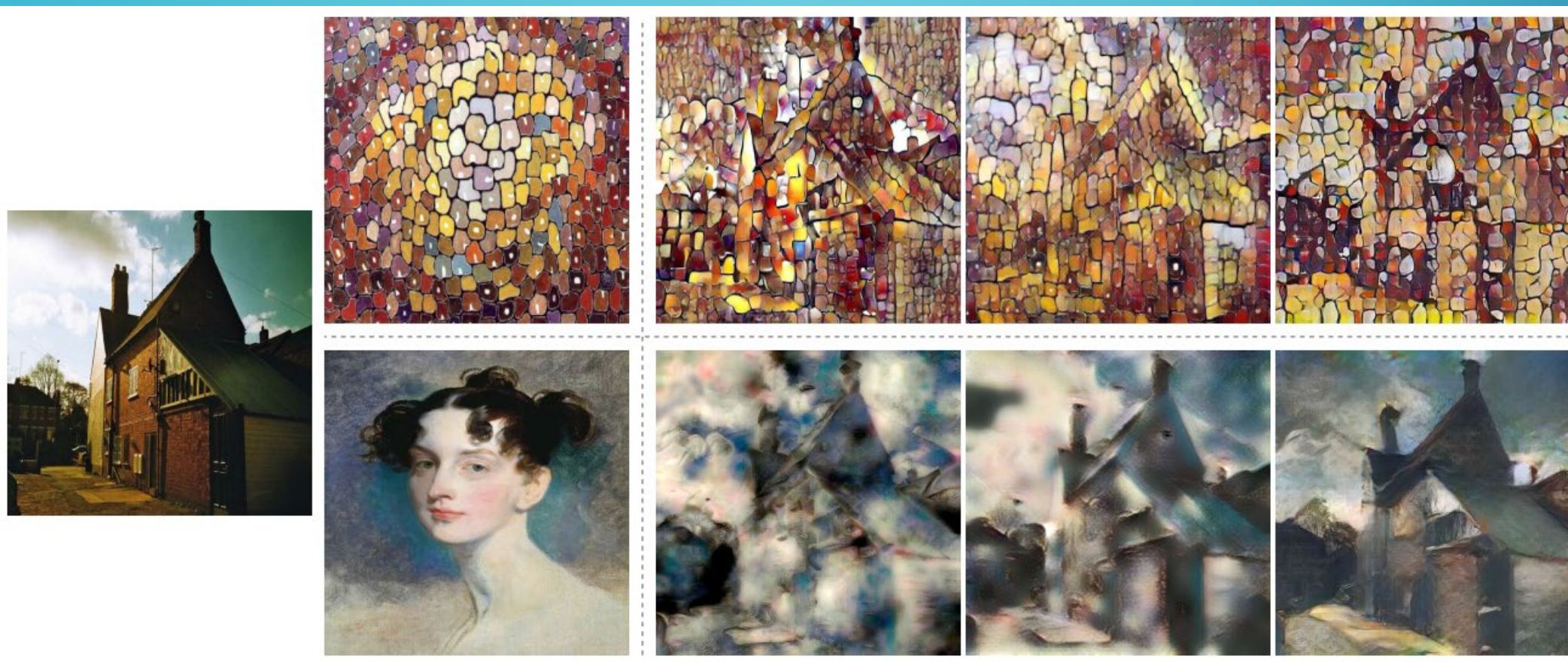
APPLICATION ON SPATIAL CONTROL



ERROR TRANSITION NETWORK

NeurIPS 2019 – “ETNet: Error Transition Network for Arbitrary Style Transfer”

OBSERVATION AND MOTIVATION



Content

Style

WCT

AvatarNet

ETNet

MULTI-STEP REFINEMENT



Content

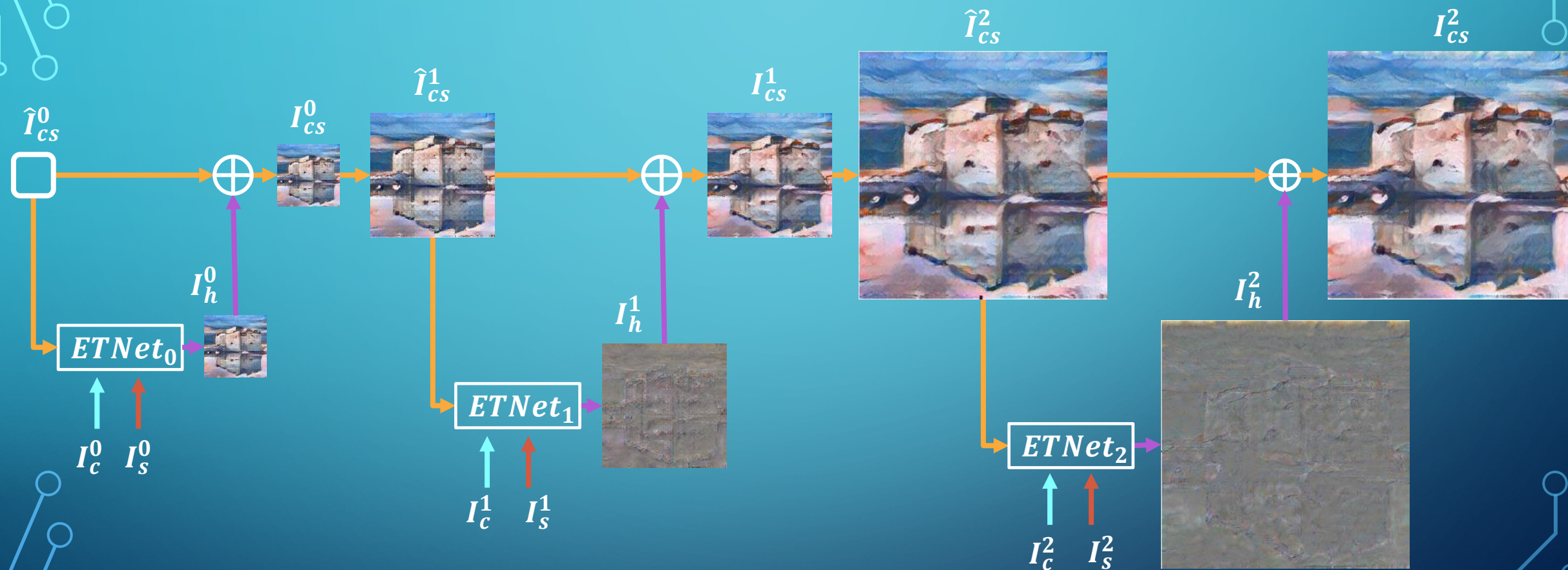
K=1

K=2

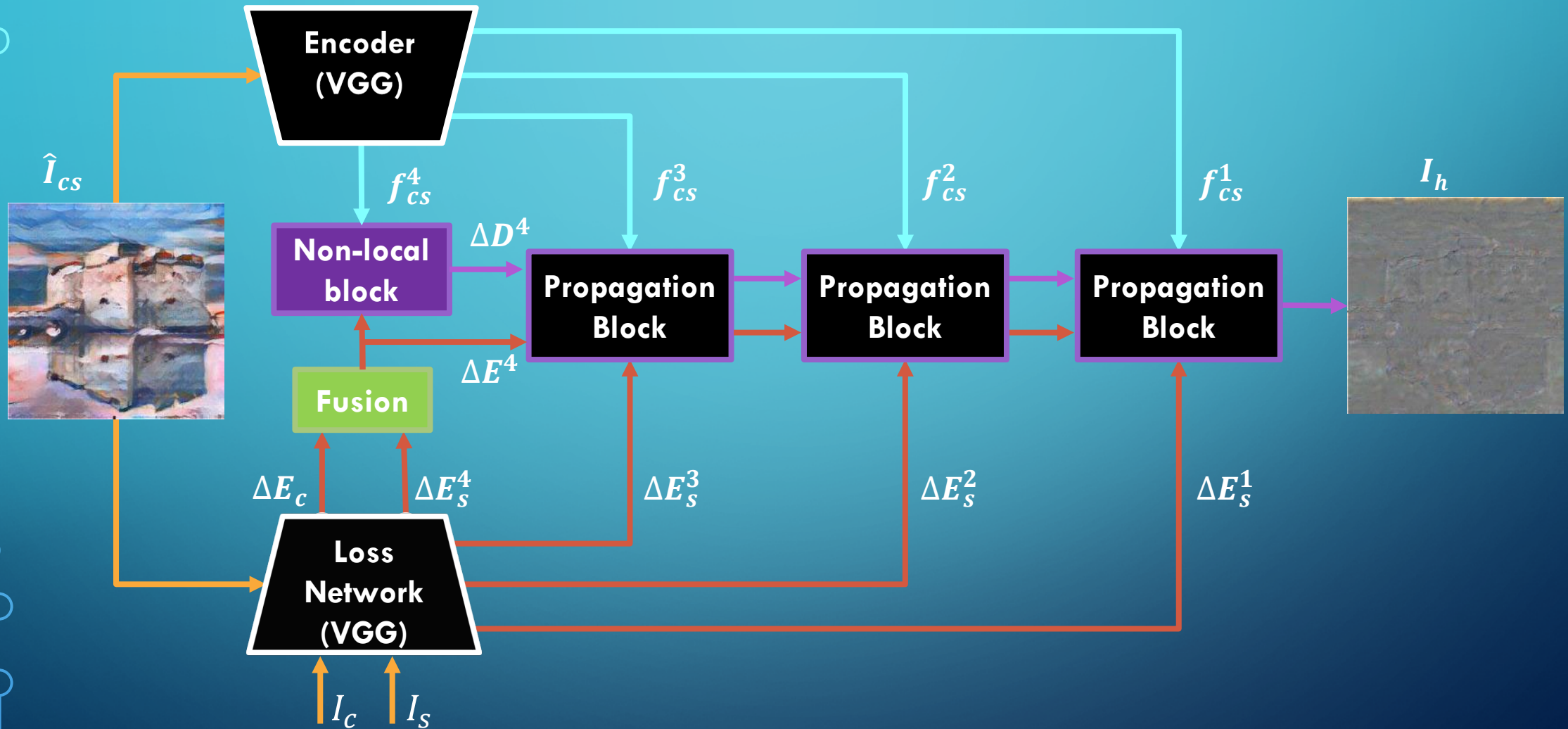
K=3

Style

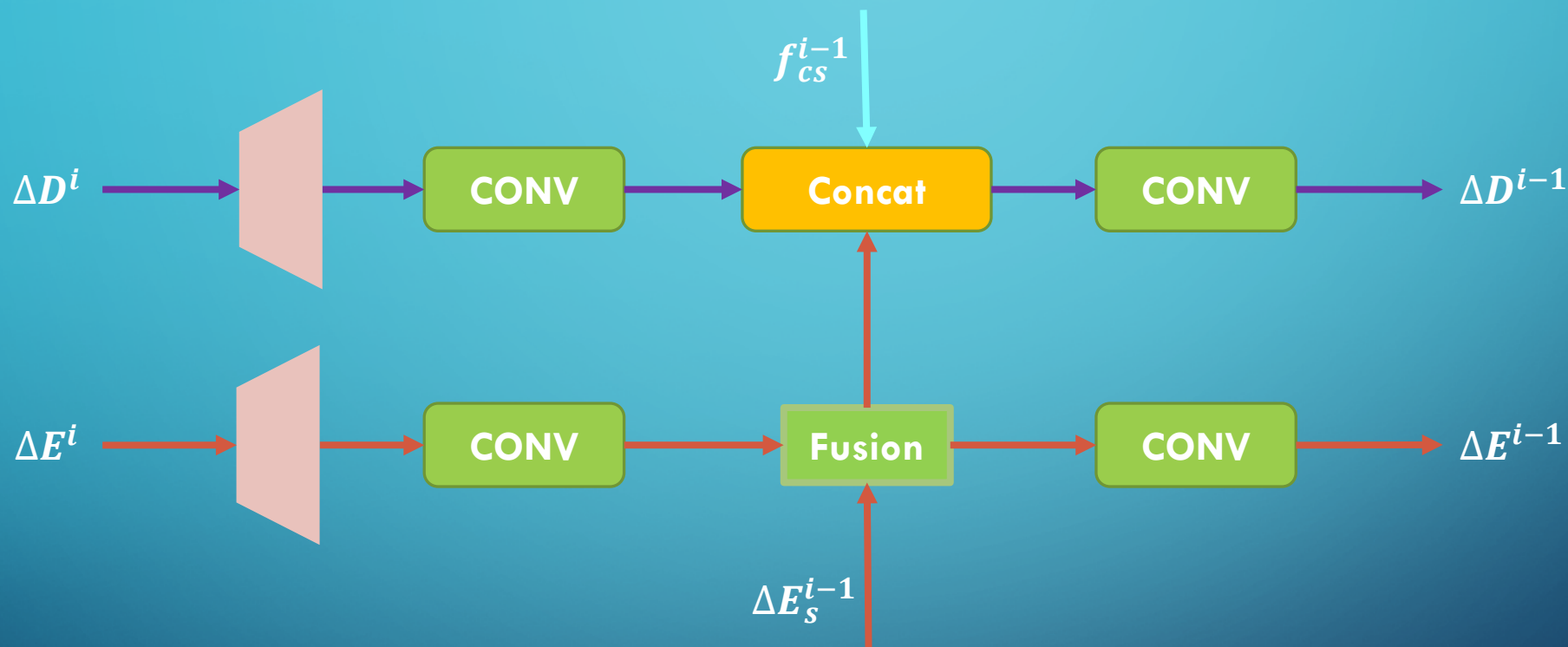
ARCHITECTURE OVERVIEW



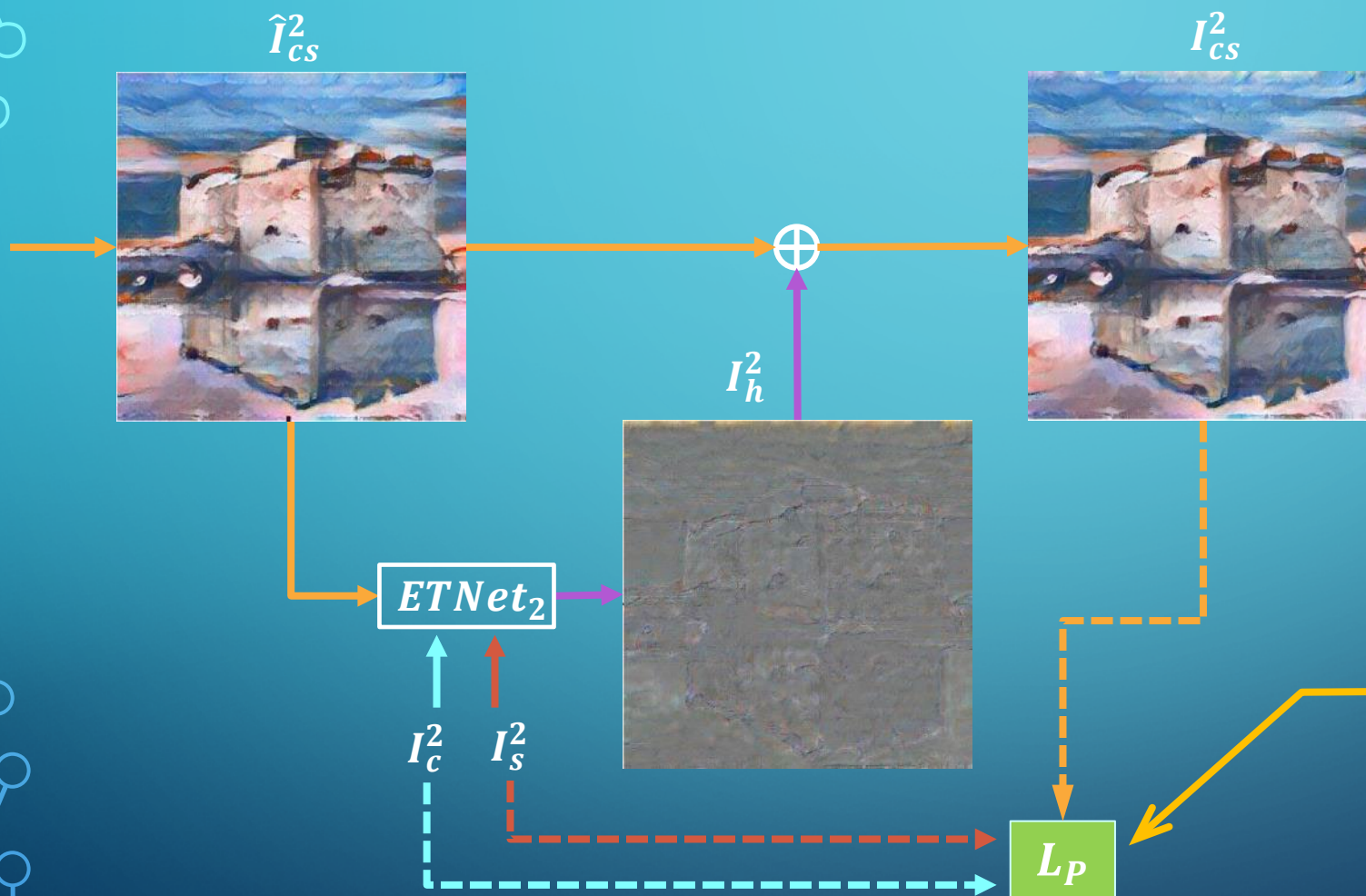
ERROR TRANSITION NETWORK



RESIDUAL PROPAGATION BLOCK



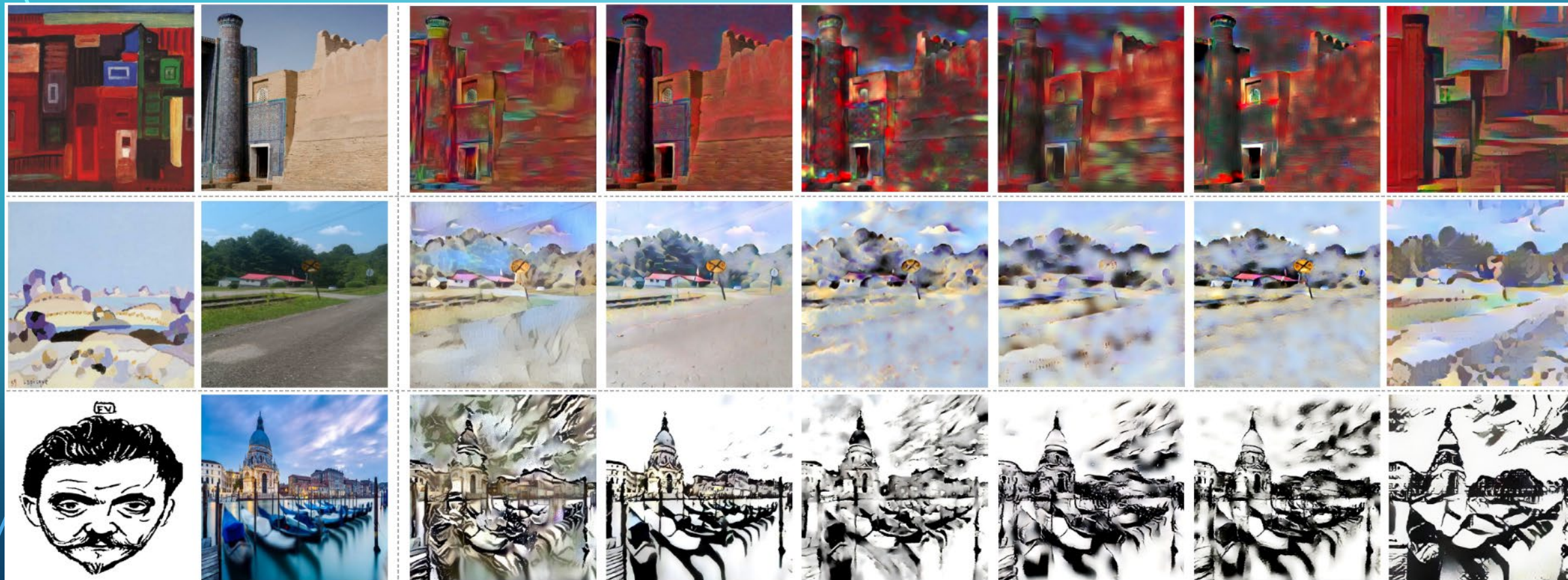
LOSS FUNCTION



$$\ell_{contnet} = \frac{1}{C H W} \|\hat{F}^L - F^L\|^2$$

$$\ell_{style} = \sum \|\hat{G}^l - G^l\|^2$$

EXPERIMENTAL RESULTS



Style

Content

Gatys

AdaIN

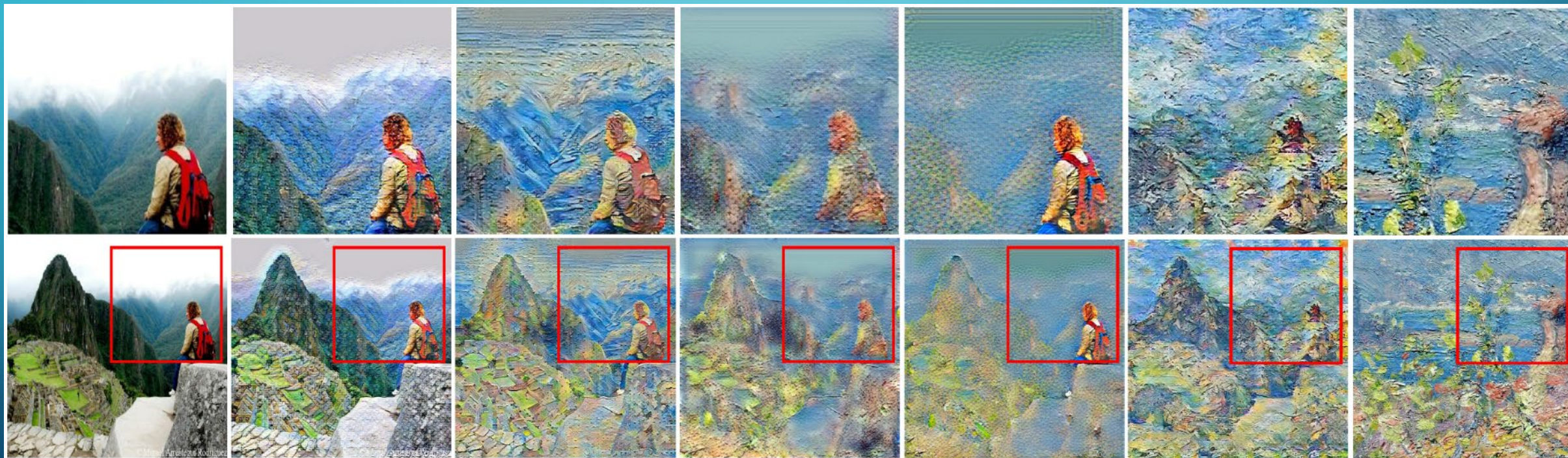
WCT

AvatarNet

AAMS

ETNet

COMPARISON ON DETAILS



Content

AdaIn

WCT

AvatarNet

AAMS

ETNet

Style

APPLICATION ON SPATIAL CONTROL



COMPARISON BETWEEN EFANet AND ETNet



Style

Content

AdaIN

WCT

AvatarNet

EFANet

ETNet

CONCLUSIONS

- **Two novel network architectures are presented for arbitrary image style transfer, each based on one simple idea**
- **Exchangeable Feature Alignment Network:**
 - The style features extracted from the style image may not be compatible with the content features from the content image, preventing effective fusion
 - Jointly analyze both style and content images to extract exchangeable features
- **Error Transition Network:**
 - Try to generate stylized images in a single shot may not produce optimal results
 - Use self-correcting model to predict errors in the current stylization and refine it accordingly in an iterative manner



ABSTRACT

- Style transfer has been an important topic in both Computer Vision and Graphics. Since the pioneer work of Gatys et al. demonstrated the power of stylization through optimization in deep feature space, a number of approaches have been developed for real-time arbitrary style transfer. However, even the state-of-the-art approaches may generate insufficiently stylized results under challenging cases. Two novel network architectures are discussed in this talk for addressing the issues and delivering better performances.
- We first observe that only considering features in the input style image for the global deep feature statistic matching or local patch swap may not always ensure a satisfactory style transfer. Hence, we propose a novel transfer framework that aims to jointly analyze and better align exchangeable features extracted from the content and style image pair. This allows the style features used for transfer to be more compatible with content information in the content image, leading to more structured stylization results.
- Another observation is that existing methods try to generate stylized result in a single shot, making it difficult to satisfy constraints on semantic structures in the content images and style patterns in the style images. Inspired by the works on error-correction, we propose a self-correcting model to predict what is wrong with the current stylization and refine it accordingly in an iterative manner. For each refinement, we transit the error features across both the spatial and scale domains and invert the processed features into a residual image.