Medical Artificial Intelligence Laboratory
At Yonsei University

Image Style Transfer

CAST

ACM SIGGRAPH

2022

Domain Enhanced Arbitrary Image Style Transfer via Contrastive Learning

Yuxin Zhang
NLPR, Institute of Automation, CAS
School of Artificial Intelligence, UCAS
China
zhangyuxin2020@ia.ac.cn

Fan Tang*
School of Artificial Intelligence,
Jilin University
China
tangfan@jlu.edu.cn

Weiming Dong*
NLPR, Institute of Automation, CAS
School of Artificial Intelligence, UCAS
China
weiming.dong@ia.ac.cn

Haibin Huang
Kuaishou Technology
China
huanghaibin03@kuaishou.com

Chongyang Ma
Kuaishou Technology
China
chongyangma@kuaishou.com

Tong-Yee Lee
National Cheng-Kung University
Taiwan
tonylee@ncku.edu.tw

Changsheng Xu
NLPR, Institute of Automation, CAS
csxu@nlpr.ia.ac.cn

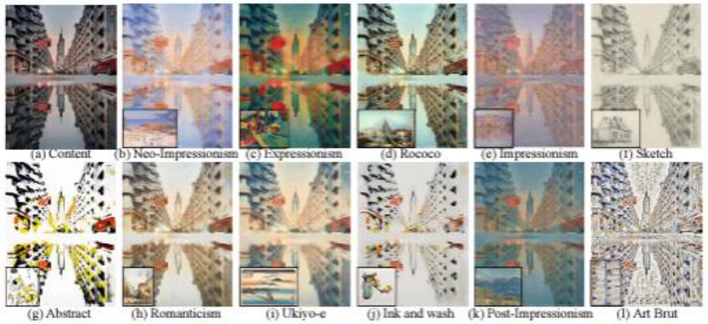


Figure 1: Style transfer results using our method, which can robustly and effectively handle various painting styles. The input content image is in the top-left corner and the style reference is shown as the inset for each result. Our method can faithfully capture the style of each painting and generate a result with a unique artistic visual appearance. Style image credits: [(b) Henri Edmond Cross, (c) Vasily Kandinsky, (d) Michele Marieschi, (e) Claude Monet, (h) Richard Parkes Bonington, (i) Utagawa Hiroshige, (k) Paul Cezanne]/The Art Institute of Chicago (CC0), (f) Vincent van Gogh/National Gallery of Art (CC0).

ABSTRACT

In this work, we tackle the challenging problem of arbitrary image style transfer using a novel style feature representation learning method. A suitable style representation, as a key component in image stylization tasks, is essential to achieve satisfactory results. Existing deep neural network based approaches achieve reasonable results with the guidance from second-order statistics such as

Image Style Transfer

CAST

Background & Goal

Previous research limitation

- Methods that utilize secondary statistics fail to utilize sufficient style information, resulting in artifacts such as l ocal distortions or style inconsistencies

Goal

- To learn style representation directly from image features, by analyzing the similarities and differences between m ultiple styles and considering the style distribution

(a) Content (b) Style (c) Ours (d) AdaAttN (e) IEST

Fig 1. Compared with models which rely on second-order statistic s

Zhang, Y., Tang, F., Dong, W., Huang, H., Ma, C., Lee, T. Y., & Xu, C. (2022, July). Domain enhanced arbitrary image style transfer via contrastive learning. In *ACM SIGGRAPH 2022 conference proceedings* (pp. 1-8).

Image Style Transfer

CAST

Network Architecture

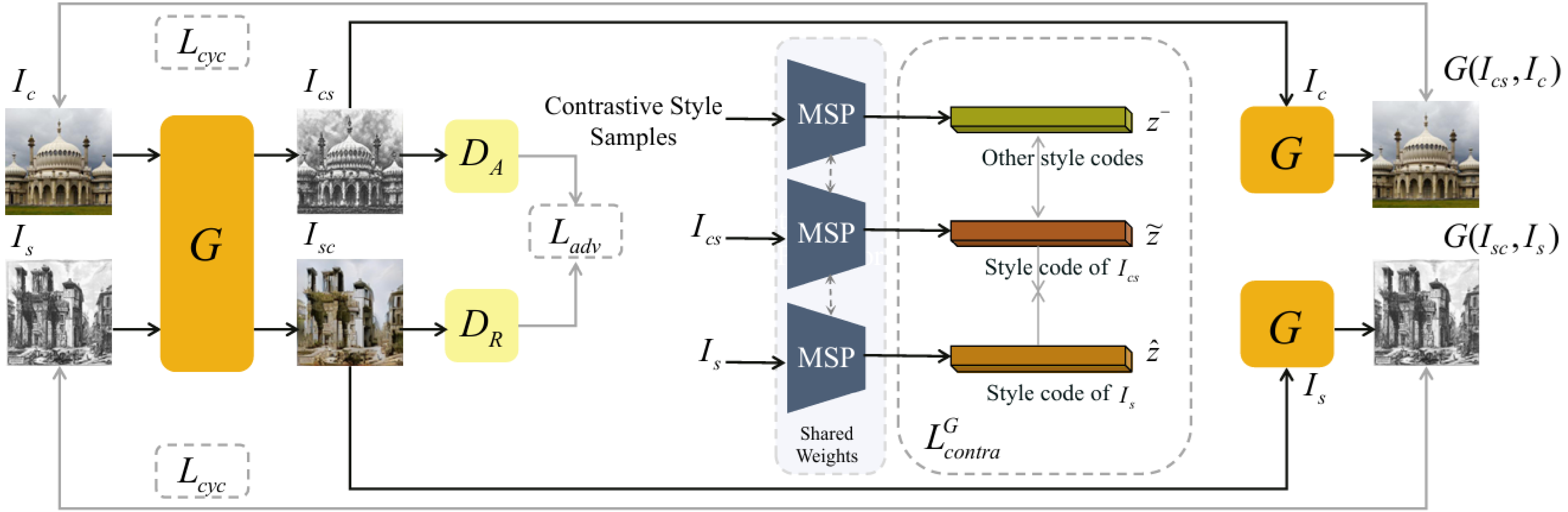


Fig 2. Overview of Network Architecture

Zhang, Y., Tang, F., Dong, W., Huang, H., Ma, C., Lee, T. Y., & Xu, C. (2022, July). Domain enhanced arbitrary image style transfer via contrastive learning. In *ACM SIGGRAPH 2022 conference proceedings* (pp. 1-8).

Image Style Transfer

CAST

Loss functions

Multi-layer style projector

- MSP projects features of different layers into separate latent style spaces to encode local and global style cues

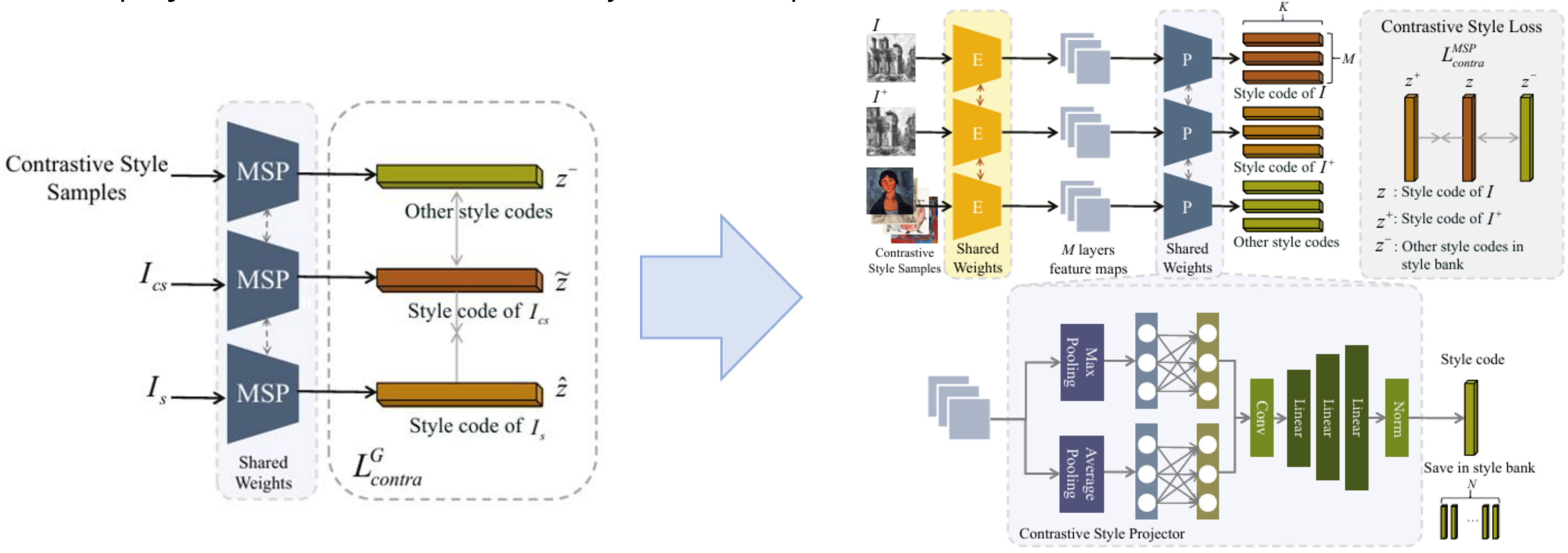


Fig 3. Overview of MSP Module

Zhang, Y., Tang, F., Dong, W., Huang, H., Ma, C., Lee, T. Y., & Xu, C. (2022, July). Domain enhanced arbitrary image style transfer via contrastive learning. In *ACM SIGGRAPH 2022 conference proceedings* (pp. 1-8).

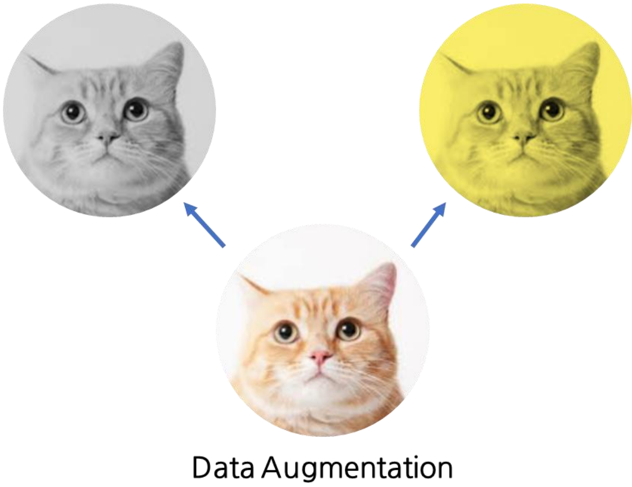
Image Style Transfer

CAST

Loss functions

Contrastive Style Learning

- Contrastive learning learns to encode images in a low-dimensional space so that similar images are close toget her and dissimilar images are far apart



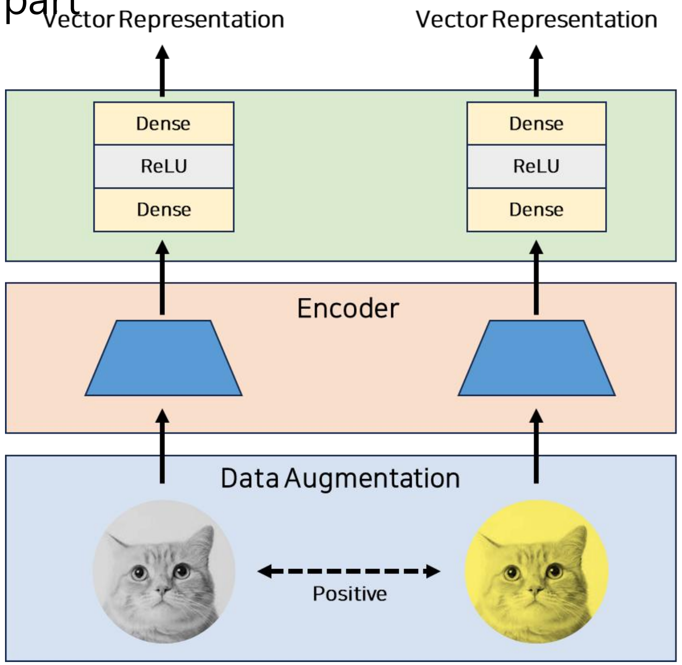


Fig 4. Contrastive learning

$$-\log \frac{\exp(z_0 \cdot z_1)}{\sum_{i=1}^{K} \exp(z_0 \cdot z_i)}$$

Eq 1. Contrastive loss function

Zhang, Y., Tang, F., Dong, W., Huang, H., Ma, C., Lee, T. Y., & Xu, C. (2022, July). Domain enhanced arbitrary image style transfer via contrastive learning. In *ACM SIGGRAPH 2022 conference proceedings* (pp. 1-8).

Image Style Transfer

CAST

Loss functions

Contrastive Style Learning

- Because of the lack of ground truth style code, adopt contrastive learning & design a novel contrastive style lo ss
- Eq 2: Learn how to distinguish different styles well with style codes
- Eq 3: Learn to make the style of the converted image similar to the style of the original style image

$$\mathcal{L}_{contra}^{MSP} = -\sum_{i=1}^{M} \log \frac{\exp(\mathbf{z}_i \cdot \mathbf{z}_i^{+}/\tau)}{\exp(\mathbf{z}_i \cdot \mathbf{z}_i^{+}/\tau) + \sum_{j=1}^{N} \exp(\mathbf{z}_i \cdot \mathbf{z}_{i_j}^{-}/\tau)}$$

Eq 2. Contrastive loss function to train MSP module

$$\mathcal{L}_{contra}^{G} = -\sum_{i=1}^{M} \log \frac{\exp(\tilde{\mathbf{z}}_i \cdot \hat{\mathbf{z}}_i/\tau)}{\exp(\tilde{\mathbf{z}}_i \cdot \hat{\mathbf{z}}_i/\tau) + \sum_{j=1}^{N} \exp(\tilde{\mathbf{z}}_i \cdot \mathbf{z}_{i_j}^{-}/\tau)}$$

Eq 3. Contrast loss between z,̃ which maps the style-transf erred content image, and ẑ, which maps the style image

Image Style Transfer

CAST

Loss functions

Domain Enhancement

- Domain Enhancement with adversarial loss to enable the network to learn the style distribution
- To maintain the content information of the content image, Add a Cycle consistency loss

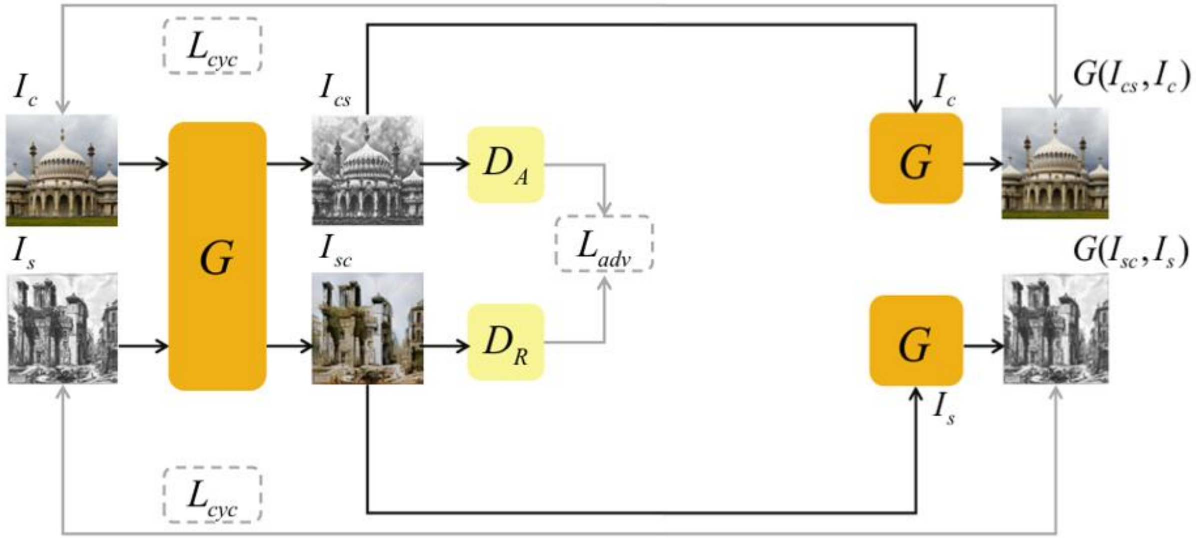


Fig 5. Overview of Domain Enhancement

$$\mathcal{L}_{adv} = \mathbb{E}[\log D_R(I_c)] + \mathbb{E}[\log(1 - D_R(I_{cs}))] + \mathbb{E}[\log D_A(I_s)] + \mathbb{E}[\log(1 - D_A(I_{sc}))]$$

Eq 4. Adversarial loss

$$\mathcal{L}_{cyc} = \mathbb{E}[\|I_c - G(I_{cs}, I_c)\|_1] + \mathbb{E}[\|I_s - G(I_{sc}, I_s)\|_1]$$

Eq 5. Cycle consistency loss

Image Style Transfer

CAST

Experiments

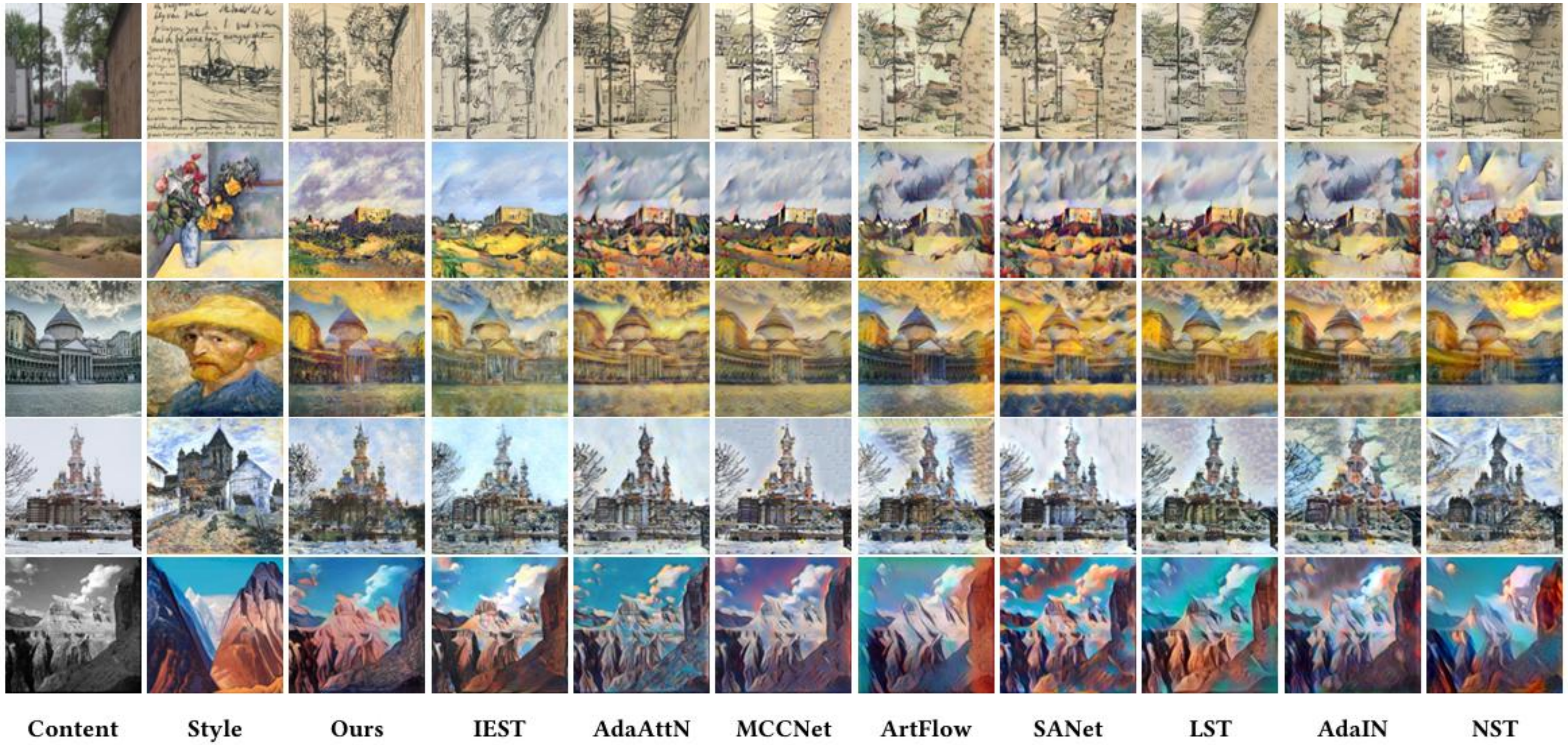


Fig 6. Qualitative comparisons with several SOTA style transfer methods

Zhang, Y., Tang, F., Dong, W., Huang, H., Ma, C., Lee, T. Y., & Xu, C. (2022, July). Domain enhanced arbitrary image style transfer via contrastive learning. In *ACM SIGGRAPH 2022 conference proceedings* (pp. 1-8).

Image Style Transfer

CAST

Experiments

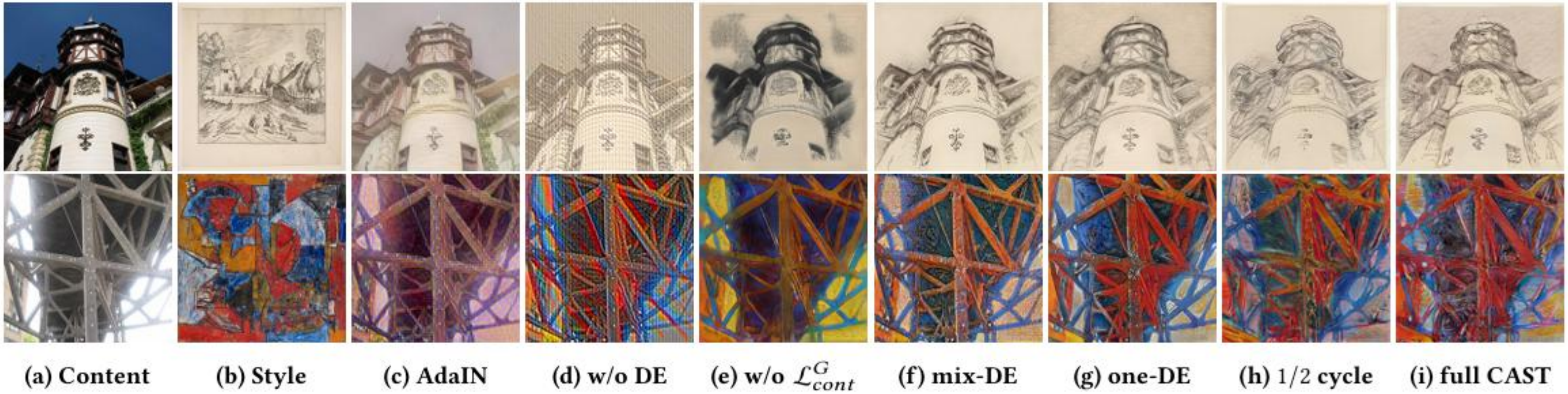


Fig 7. Ablation study results

Zhang, Y., Tang, F., Dong, W., Huang, H., Ma, C., Lee, T. Y., & Xu, C. (2022, July). Domain enhanced arbitrary image style transfer via contrastive learning. In *ACM SIGGRAPH 2022 conference proceedings* (pp. 1-8).

Image Style Transfer

CAST

Implications & Limitations

Implications

- Use image features directly by introducing an MSP module for style encoding
- Propose a DE scheme to effectively model the distribution of realistic and artistic image domains

Limitations

- The patch-based discriminator has limitations, leading to inconsistencies in style uniformity.