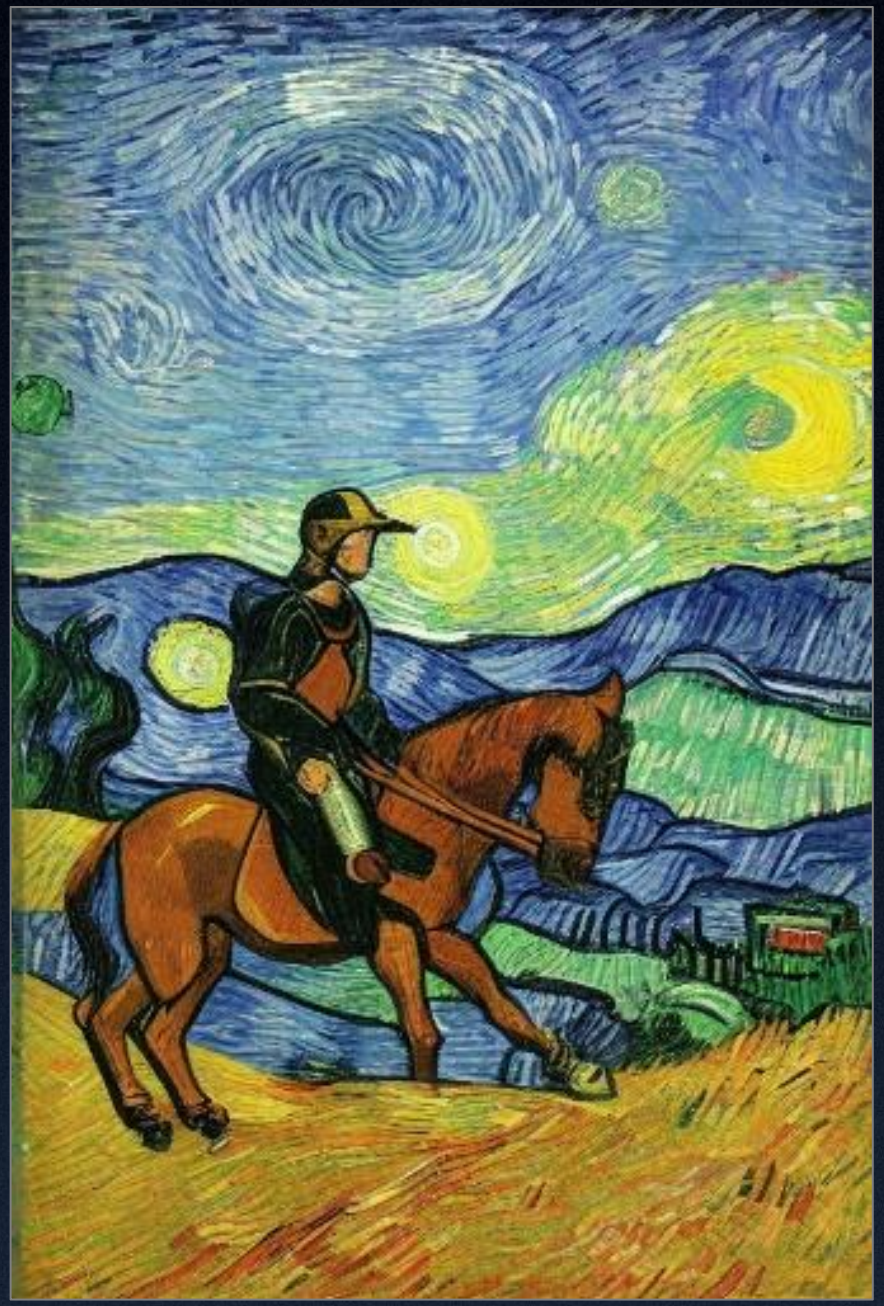


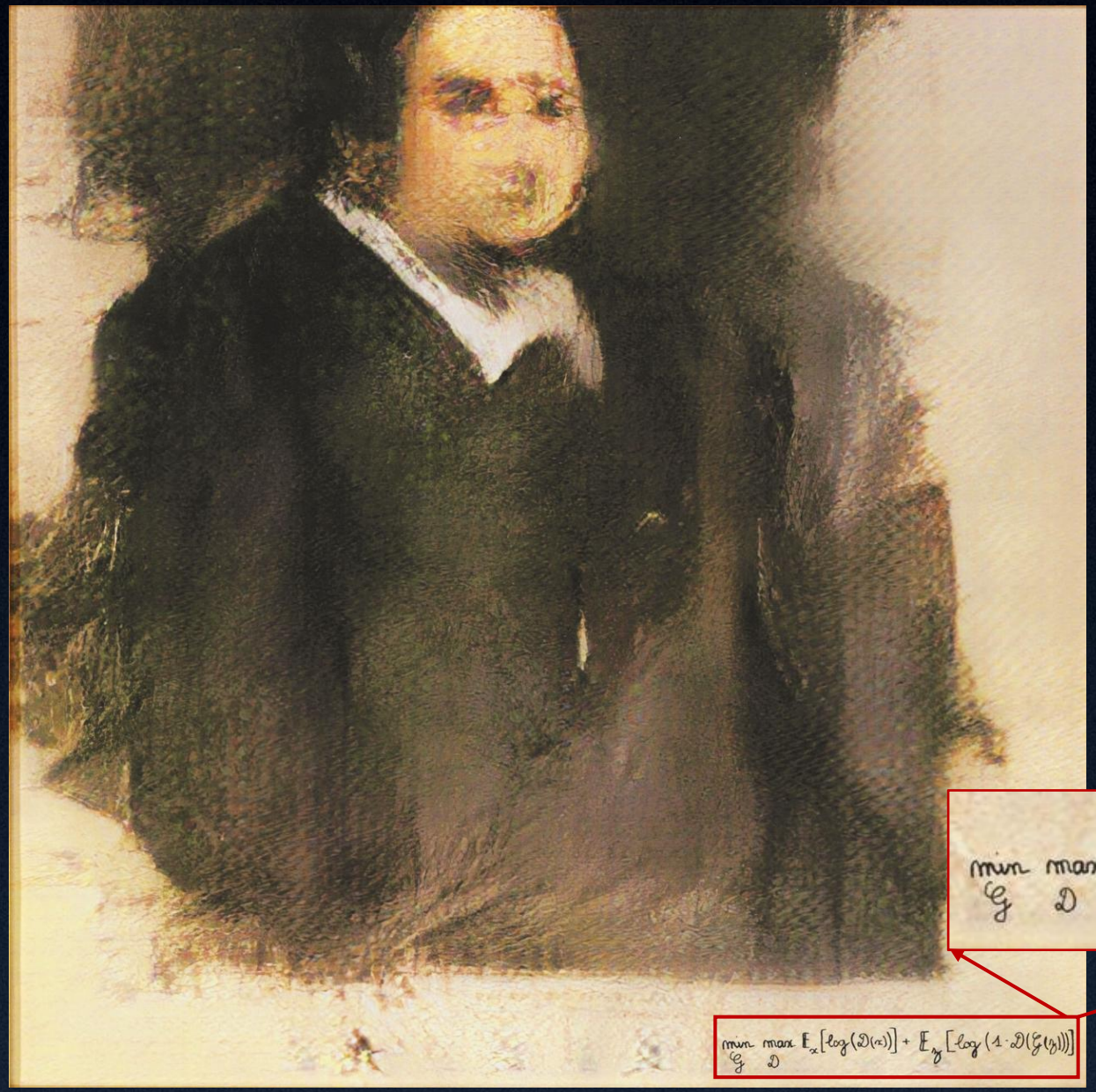
风格迁移技术：原理、方法与展望

唐帆

中国科学院计算技术研究所

数字内容合成与伪造检测实验室

2018年佳士得纽约Prints & Multiples专场拍卖，最终以43.25万美元成交。



Portrait of Edmond Belamy
埃德蒙·贝拉米的肖像

Théâtre D'opéra Spatial Jason Allen

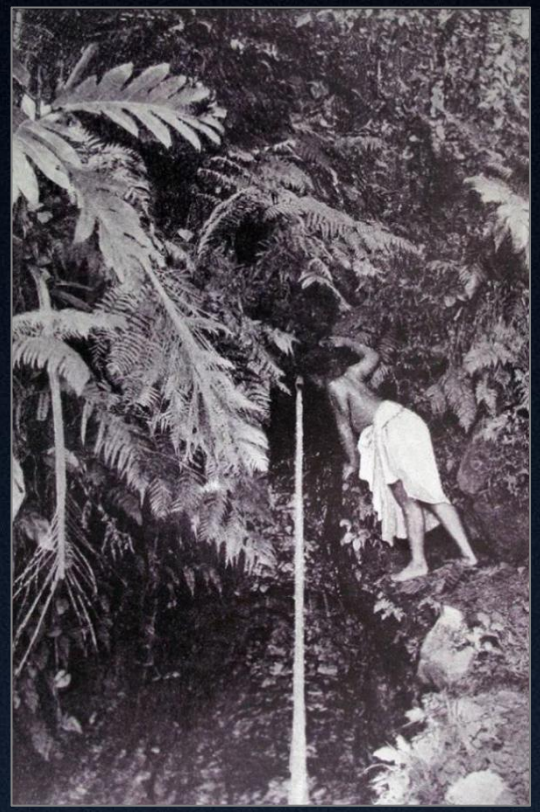
在AI绘画工具中输入
光源、构图、氛围等关键词

通过PS进行了约80个小时修饰

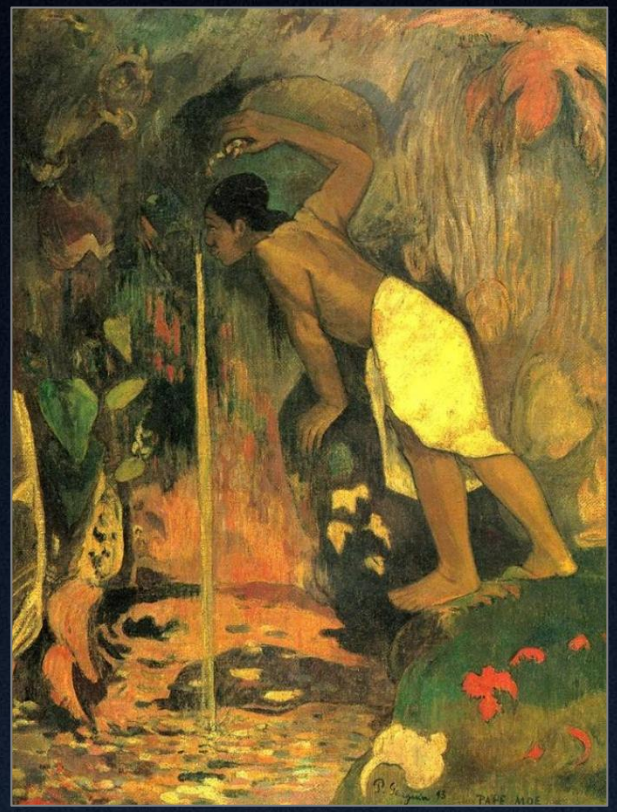
将图像输出并打印到画布上


图像创作的人工智能时代拉开序幕

AI绘画创作模式之：图→图



照片
Charles Spitz 1863



绘画：Pape Moe
保罗·高更 1893

图像风格迁移/风格化

“The painting is directly based on a photograph by Charles Georges Spitz (1857-1894) that he had taken ten years earlier on an excursion to Mount Aorai.”

图像风格迁移

自然图像（内容图）
美国旧金山艺术宫
1962年由德裔建筑师梅贝克重修

+

真实绘画（风格图）
威特尼景观
1901年由莫奈绘制

上个世纪末，图像风格化滤镜

源图像

迭代中间结果1

迭代中间结果2

最终风格化结果

Aaron Hertzmann: Painterly Rendering with Curved Brush Strokes of Multiple Sizes. **SIGGRAPH** 1998: 453-460

上个世纪末，图像风格化滤镜

源图像

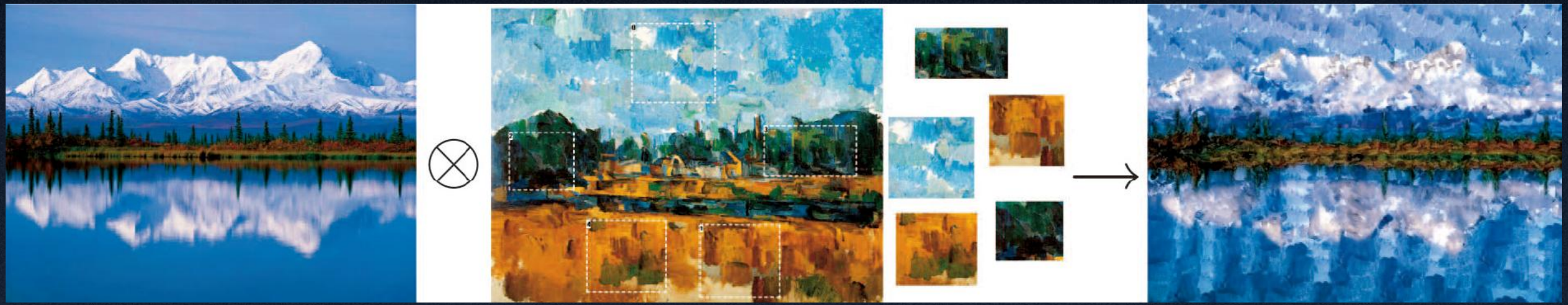
迭代中间结果1

迭代中间结果2

最终风格化结果

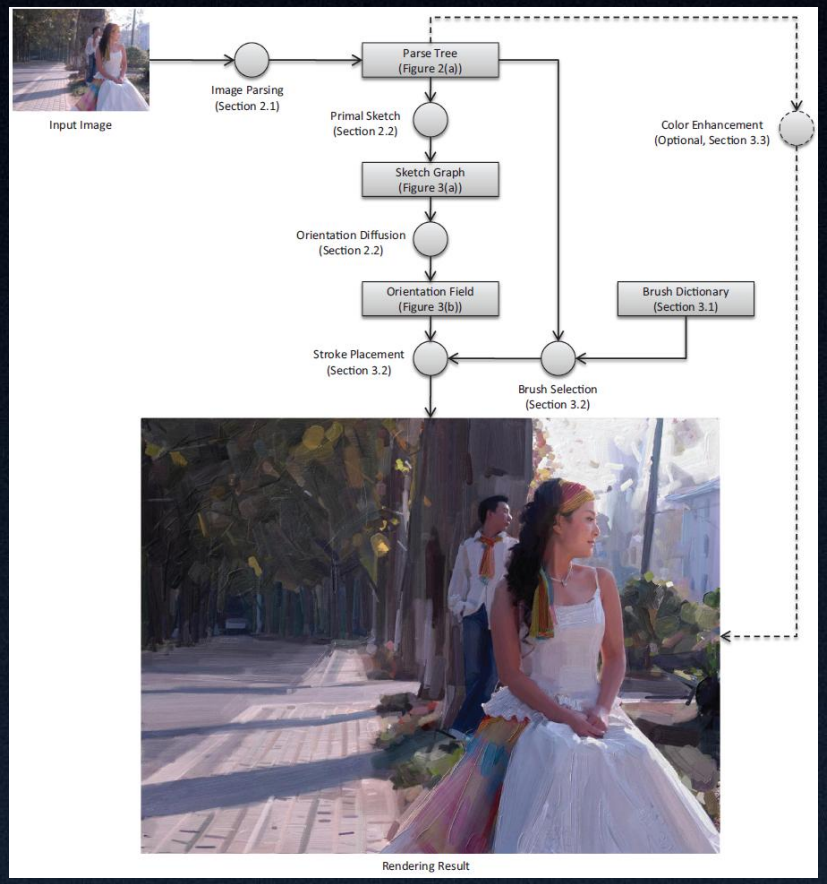
Aaron Hertzmann: Painterly Rendering with Curved Brush Strokes of Multiple Sizes. **SIGGRAPH** 1998: 453-460

20年前，纹理合成



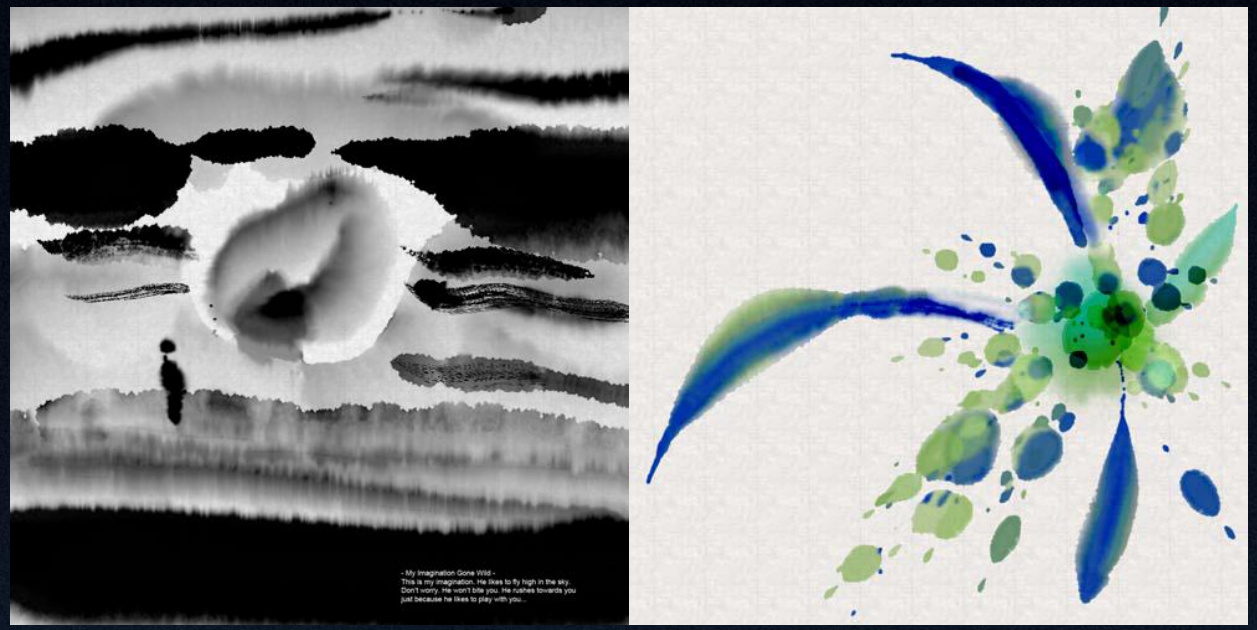
Bin Wang, Wenping Wang, Huaiping Yang, Jia-Guang Sun: Efficient Example-Based Painting and Synthesis of 2D Directional Texture. **IEEE Transactions on Visualization and Computer Graphics** 10(3): 266-277 (2004)

15年前，笔触建模



基于场景解析+笔触放置

Kun Zeng, Mingtian Zhao, Caiming Xiong, Song Chun Zhu: From image parsing to painterly rendering.
ACM Transactions on Graphics 29(1): 2:1-2:11 (2009)



Nelson Siu-Hang Chu, Chiew-Lan Tai: MoXi: real-time ink dispersion in absorbent paper. **ACM Transactions on Graphics** 24(3): 504-511 (2005)

基于物理模拟：水墨画

基于物理模拟：水彩画

Miaoyi Wang, Bin Wang, Yun Fei, Kang-Lai Qian, Wenping Wang, Jiating Chen, Jun-Hai Yong: Towards Photo Watercolorization with Artistic Verisimilitude. **IEEE Transactions on Visualization Computer Graphics** 20(10): 1451-1460 (2014)

Figure 3. Images that combine the content of a photograph with the style of several well-known artworks. The images were created by finding an image that simultaneously matches the content representation of the photograph and the style representation of the artwork. The original photograph depicting the Neckarfront in Tübingen, Germany, is shown in **A** (Photo: Andreas Praefcke). The painting that provided the style for the respective generated image is shown in the bottom left corner of each panel. **B** *The Shipwreck of the Minotaur* by J.M.W. Turner, 1805. **C** *The Starry Night* by Vincent van Gogh, 1889. **D** *Der Schrei* by Edvard Munch, 1893. **E** *Femme nue assise* by Pablo Picasso, 1910. **F** *Composition VII* by Wassily Kandinsky, 1913.

Image Style Transfer Using Convolutional Neural Networks

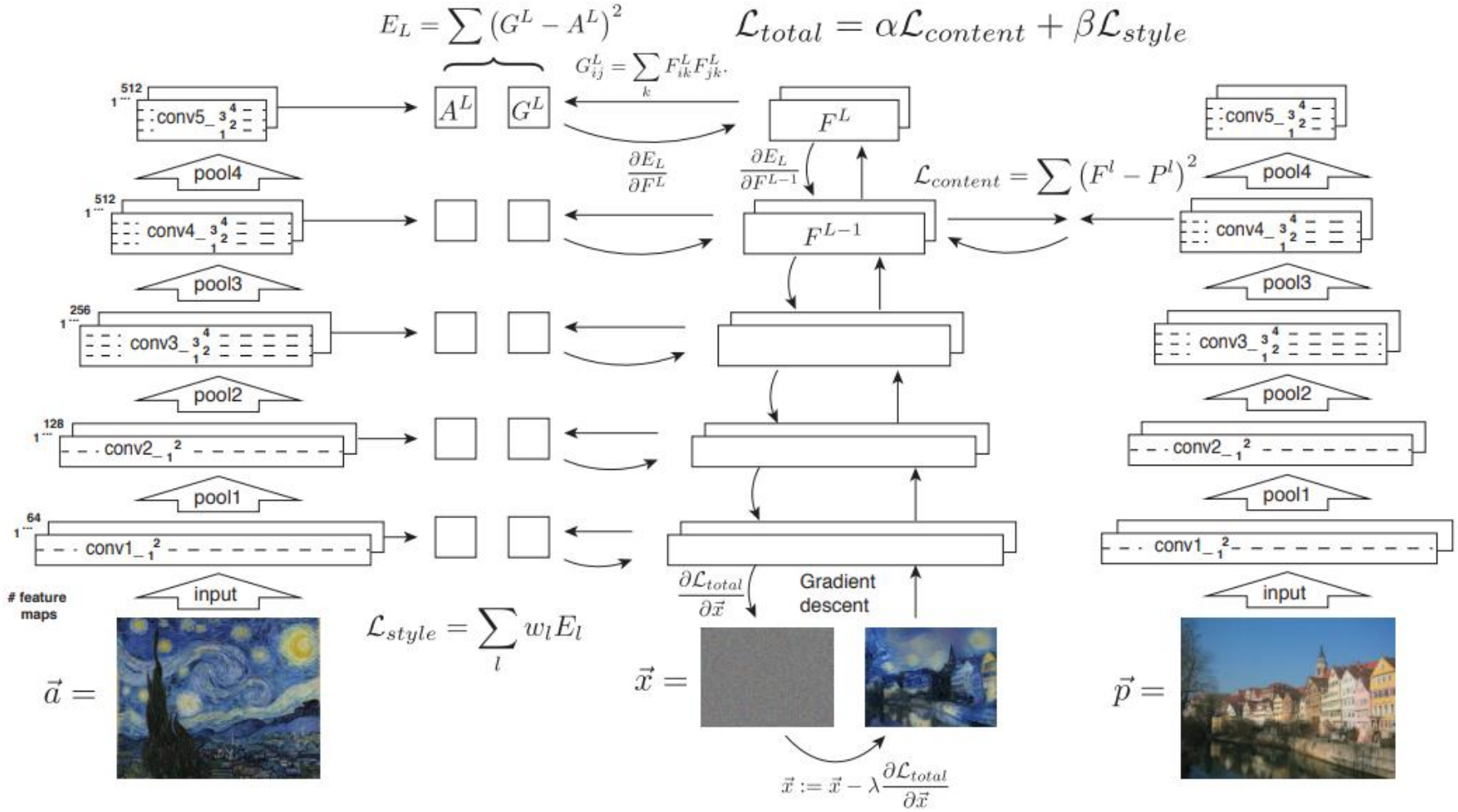
Leon A. Gatys
Alexander S. Ecker
Matthias Bethge

深度神经网络与神经风格迁移

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge: Image Style Transfer Using Convolutional Neural Networks. **CVPR** 2016: 2414-2423

神经风格迁移，Nerual Style Transfer

基本框架：基于预训练神经网络的特征比对与目标图像“寻找”

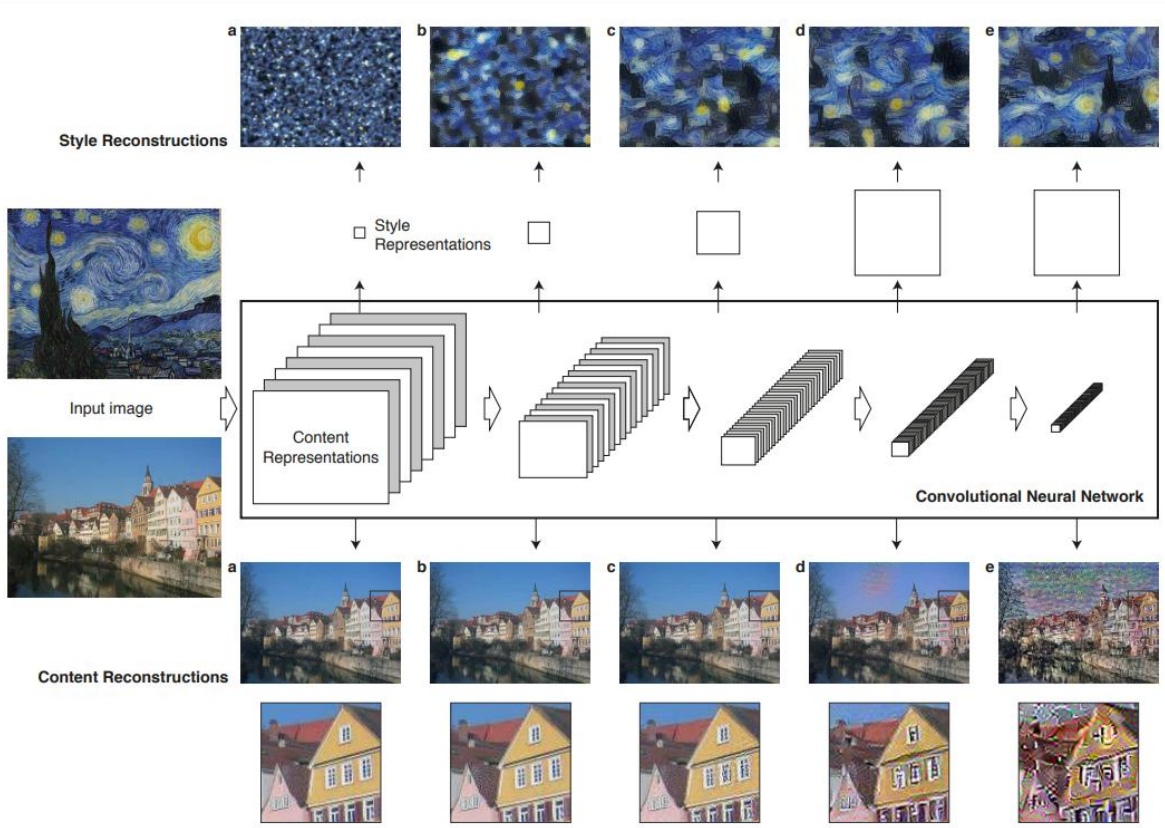


Leon A. Gatys, Alexander S. Ecker, Matthias Bethge: Image Style Transfer Using Convolutional Neural Networks. **CVPR** 2016: 2414-2423

神经风格迁移，Nerual Style Transfer

诀窍：基于深度特征的内容风格度量

$$\mathcal{L}_{total} = \alpha\mathcal{L}_{content} + \beta\mathcal{L}_{style}$$



$$loss = \underbrace{|\text{参考图片的风格} - \text{生成图片的风格}|}_{style\ loss} + \underbrace{|\text{原始图片的内容} - \text{生成图片的内容}|}_{content\ loss}$$

内容度量：embedding similarity

$$\mathcal{L}_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2}\sum_{i,j}\left(F_{ij}^{l} - P_{ij}^{l}\right)^2$$

风格度量：Gram similarity

$$G_{ij}^{l} = \sum_{k} F_{ik}^{l} F_{jk}^{l}$$

$$E_l = \frac{1}{4N_l^2 M_l^2}\sum_{i,j}\left(G_{ij}^{l} - A_{ij}^{l}\right)^2$$

$$\mathcal{L}_{\text{style}}(\vec{a}, \vec{x}) = \sum_{l=0}^{L} w_l E_l$$

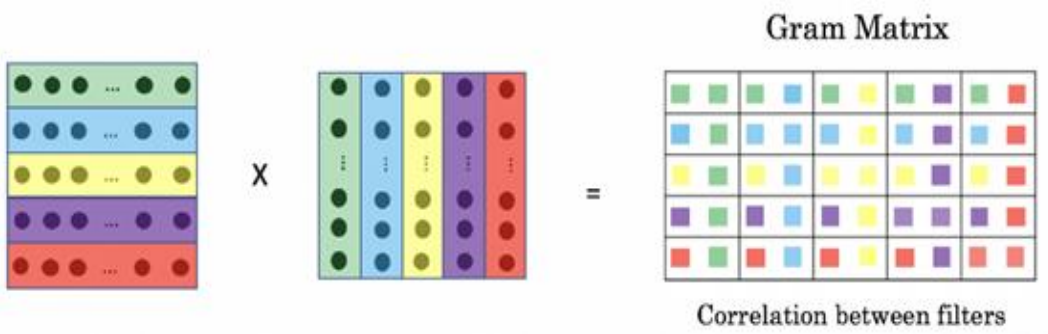
Leon A. Gatys, Alexander S. Ecker, Matthias Bethge: Image Style Transfer Using Convolutional Neural Networks. **CVPR** 2016: 2414-2423

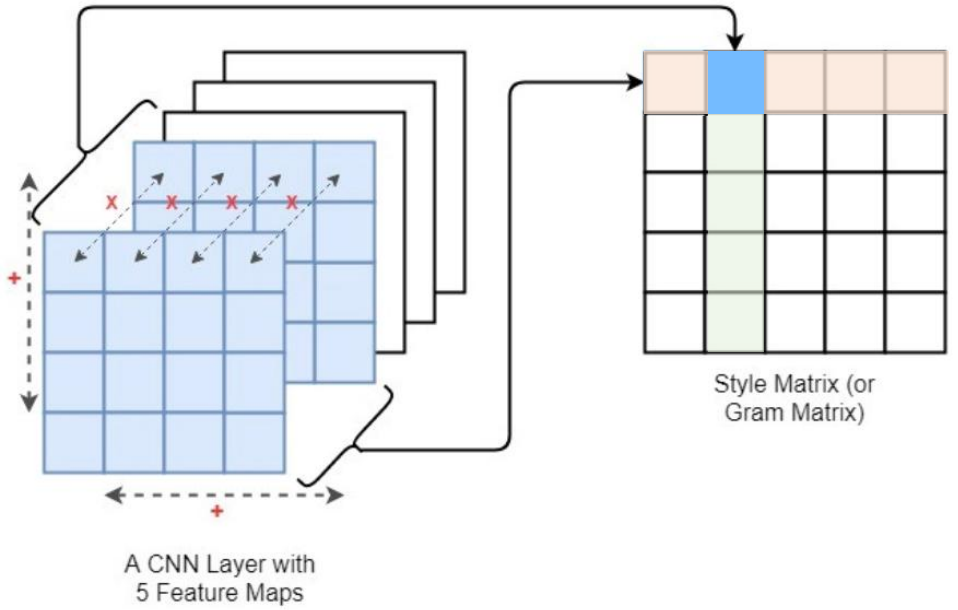
神经风格迁移，Nerual Style Transfer

核心：基于格莱姆矩阵的风格表示

k个向量之间两两的内积所组成的矩阵，称为这k个向量的格拉姆矩阵(Gram matrix)。

$$\Delta(\alpha_1,\alpha_2,\ldots,\alpha_k)=\begin{pmatrix}(\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \ldots & (\alpha_1,\alpha_k)\\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \ldots & (\alpha_2,\alpha_k)\\ \ldots & \ldots & \ldots & \ldots\\ (\alpha_k,\alpha_1) & (\alpha_k,\alpha_2) & \ldots & (\alpha_k,\alpha_k)\end{pmatrix}$$





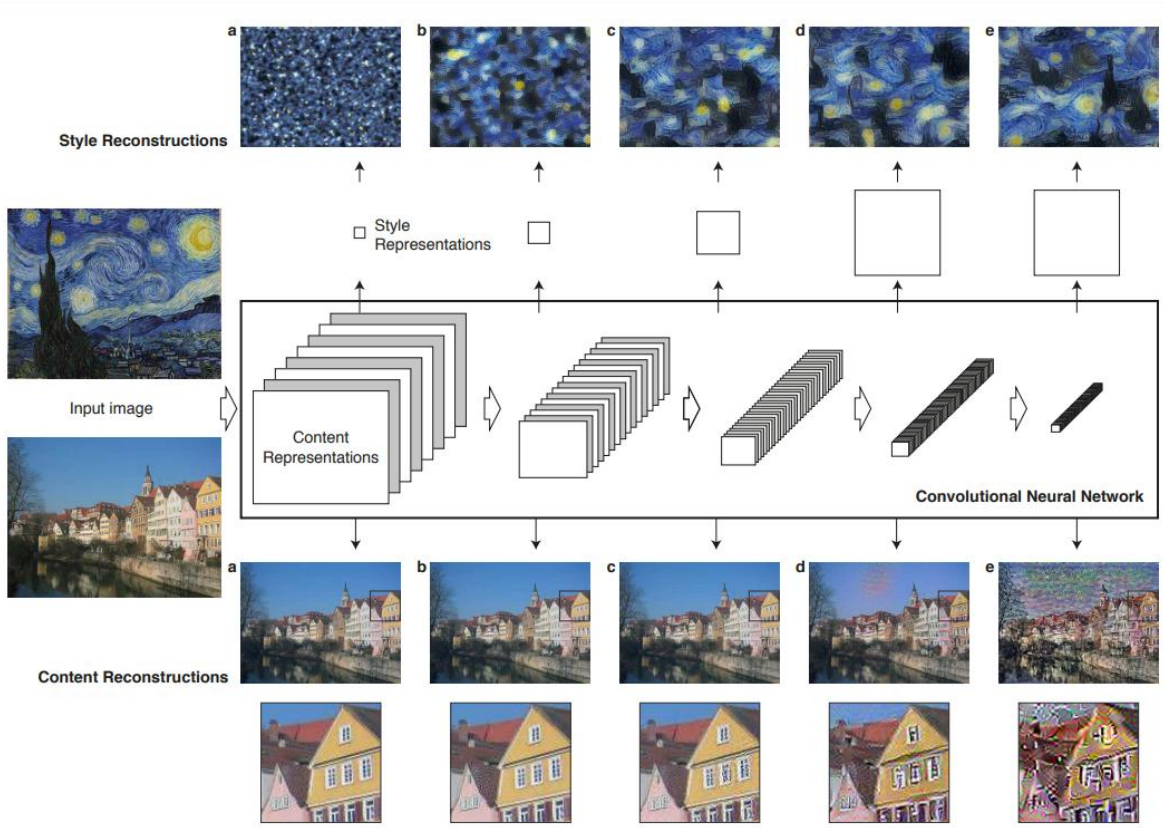
$$G^l_{ij}=\sum_k F^l_{ik}F^l_{jk}$$

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge: Image Style Transfer Using Convolutional Neural Networks. CVPR 2016: 2414-2423

神经风格迁移，Nerual Style Transfer

诀窍：基于深度特征的内容风格度量

$$\mathcal{L}_{total} = \alpha\mathcal{L}_{content} + \beta\mathcal{L}_{style}$$



$$loss = \underbrace{|\text{参考图片的风格} - \text{生成图片的风格}|}_{style\ loss} + \underbrace{|\text{原始图片的内容} - \text{生成图片的内容}|}_{content\ loss}$$

内容度量：embedding similarity

$$\mathcal{L}_{\text{content}}(\vec{p}, \vec{x}, l) = \frac{1}{2}\sum_{i,j}\left(F_{ij}^{l} - P_{ij}^{l}\right)^2$$

风格度量：Gram similarity

$$G_{ij}^{l} = \sum_{k} F_{ik}^{l} F_{jk}^{l}$$

$$E_l = \frac{1}{4N_l^2 M_l^2}\sum_{i,j}\left(G_{ij}^{l} - A_{ij}^{l}\right)^2$$

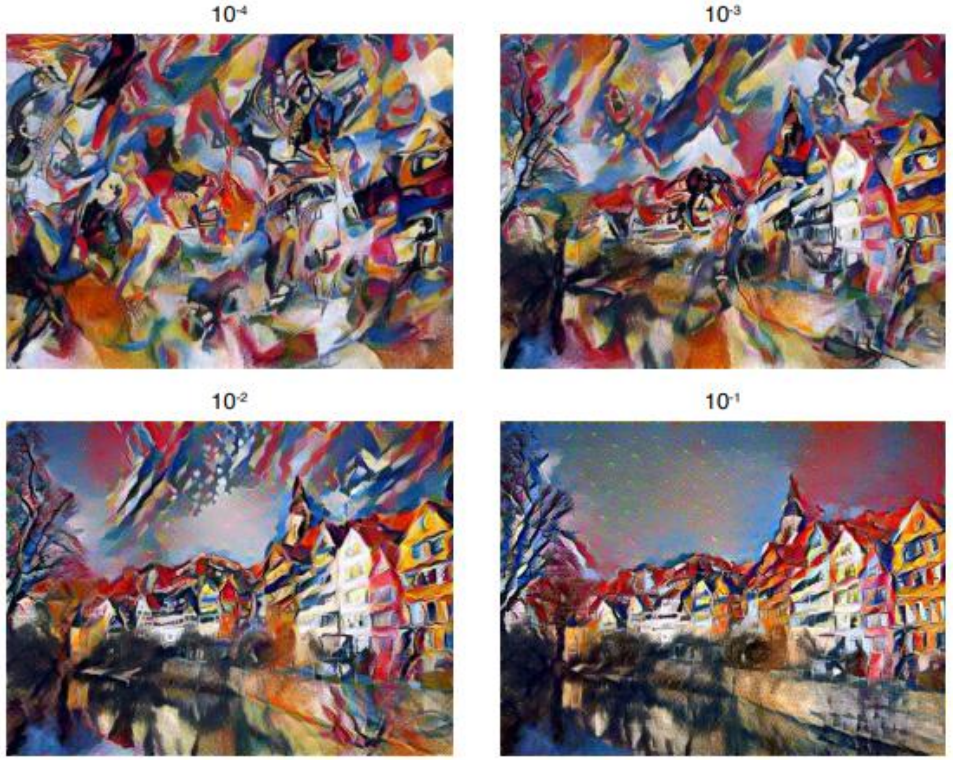
$$\mathcal{L}_{\text{style}}(\vec{a}, \vec{x}) = \sum_{l=0}^{L} w_l E_l$$

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge: Image Style Transfer Using Convolutional Neural Networks. **CVPR** 2016: 2414-2423

神经风格迁移，Nerual Style Transfer

诀窍：基于深度特征的内容风格度量

$$\mathcal{L}_{total} = \alpha\mathcal{L}_{content} + \beta\mathcal{L}_{style}$$

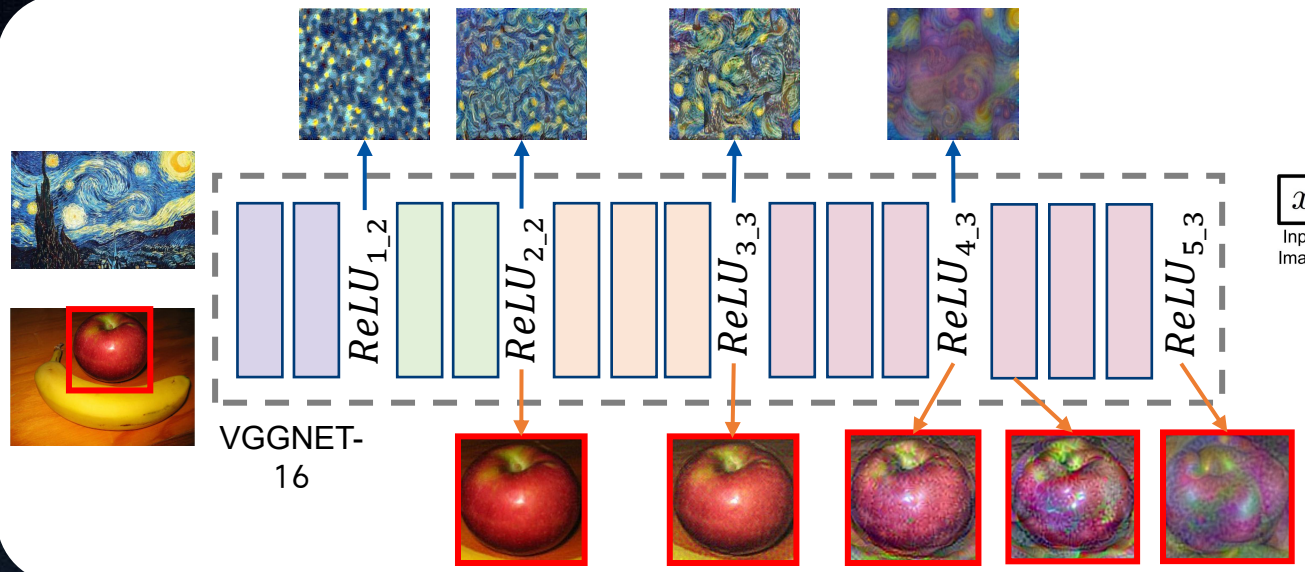


α/β.

Initialisation.

Leon A. Gatys, Alexander S. Ecker, Matthias Bethge: Image Style Transfer Using Convolutional Neural Networks. **CVPR** 2016: 2414-2423

训练网络实现实时风格迁移



Perceptual Loss

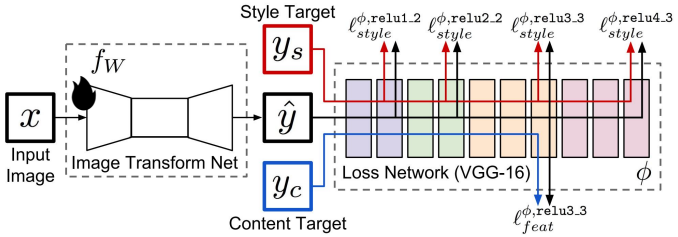
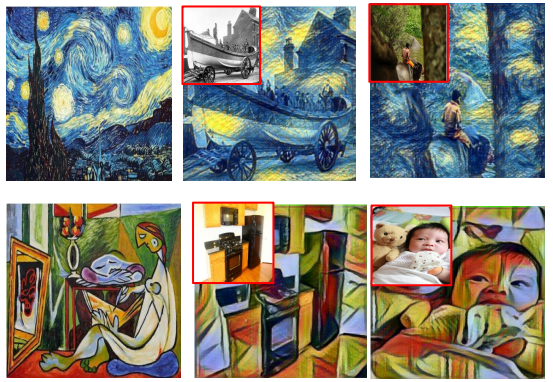


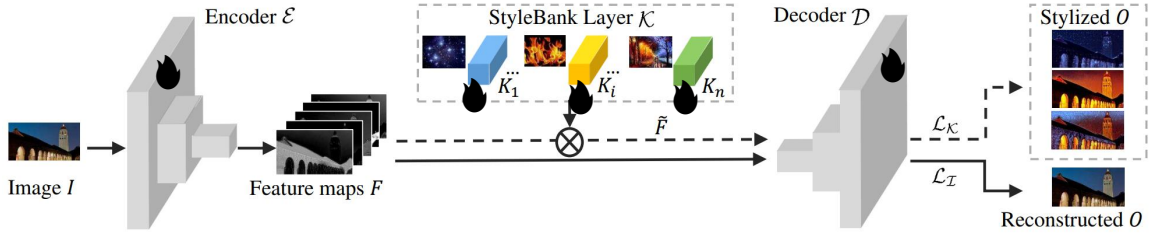
Image Size	Gatys *et al* [10] 100	300	500	Ours
256 × 256	3.17	9.52s	15.86s	**0.015s**
512 × 512	10.97	32.91s	54.85s	**0.05s**
1024 × 1024	42.89	128.66s	214.44s	**0.21s**



单网络单风格

J. Johnson, A. Alahi, and L. Fei-Fei, "C," in Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. Springer, 2016, pp. 694–711.

StyleBank: 为每一种风格学习一个StyleBank Layer



单网络多风格

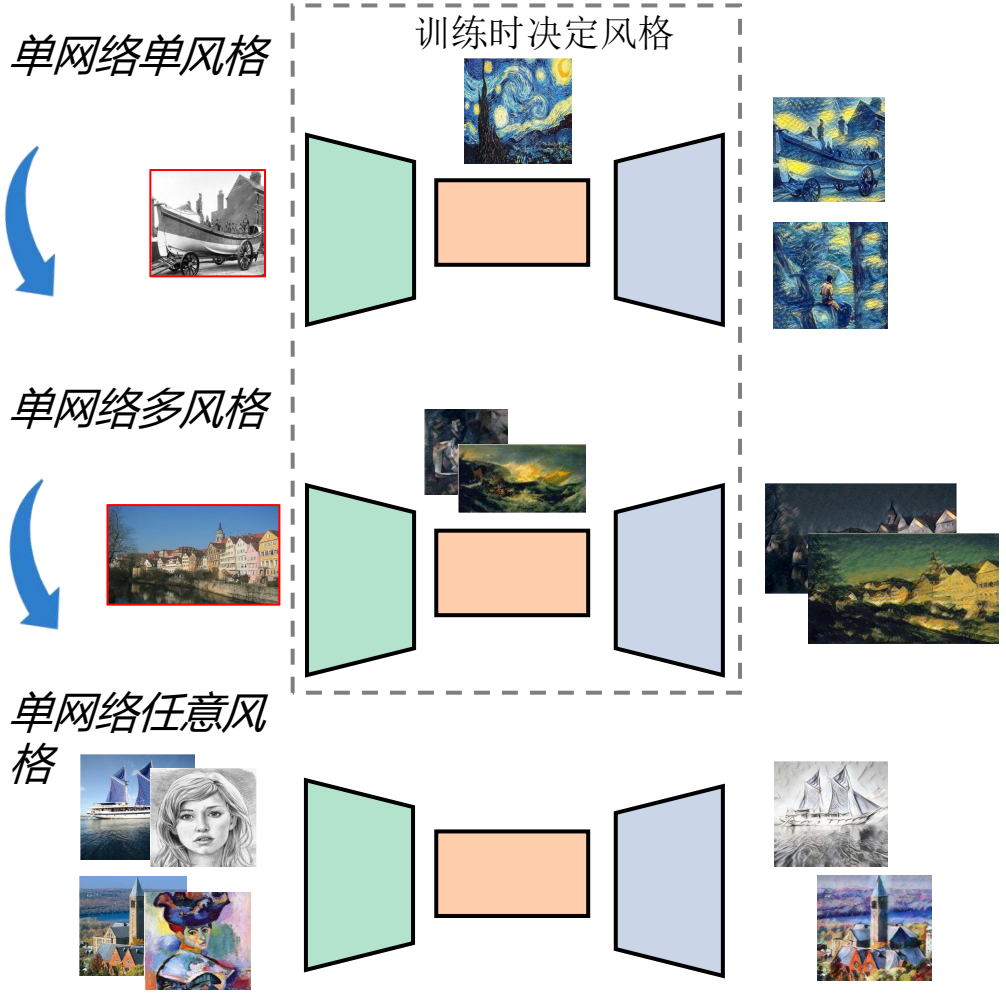
StyleBank

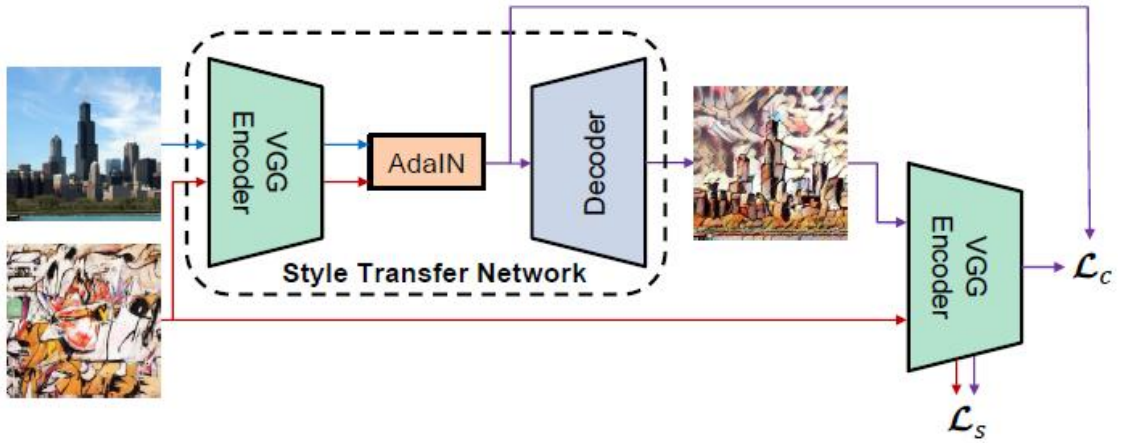
Perceptual loss

Dongdong Chen, Lu Yuan, Jing Liao, Nenghai Yu, and Gang Hua. Stylebank: An explicit representation for neural image style transfer. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 1897–1906, 2017. 2

任意风格图像风格化

诀窍：自适应实例归一化实现跨域融合





$$\text{AdaIN}(x, y) = \sigma(y)\left(\frac{x - \mu(x)}{\sigma(x)}\right) + \mu(y)$$

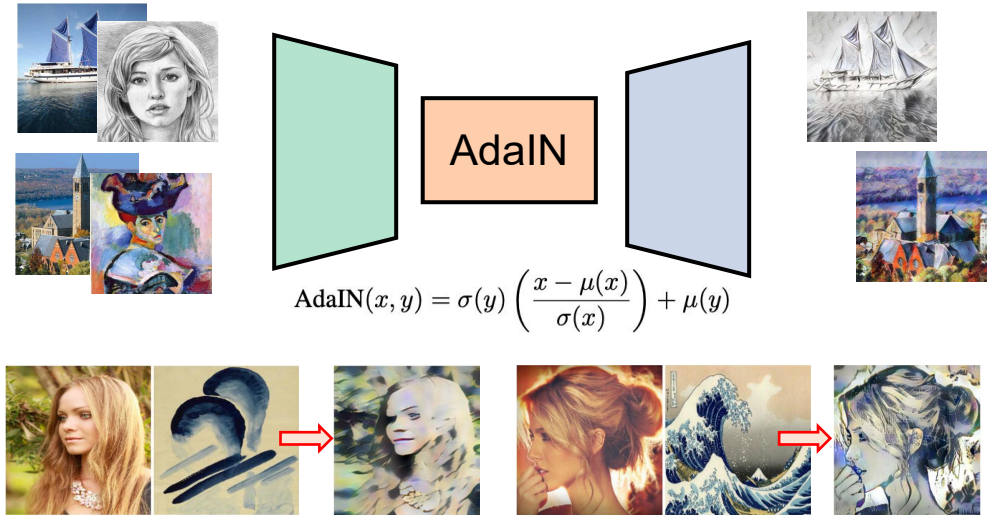
内容实例特征 (x) 的均值和方差

对齐

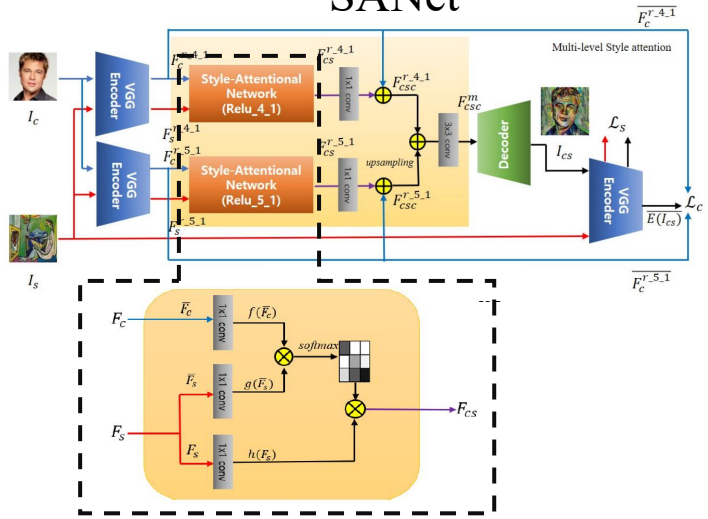
风格实例特征(y)的均值和方差

Xun Huang and Belongie Serge. Arbitrary style transfer in real-time with adaptive instance normalization. In IEEE International Conference on Computer Vision (ICCV), pages 1501–1510, 2017.

Encoder-Transformer-Decoder



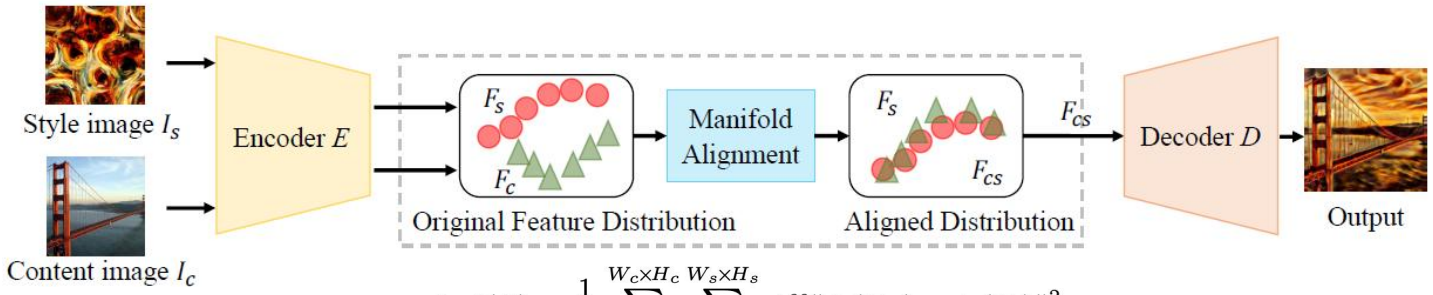
SANet



D. Y. Park and K. H. Lee, "Arbitrary style transfer with style attentional networks," in proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 5880–5888.

MAST

假设多个语义区域的图像遵循多流形分布，风格化任务可以视为基于流形对齐的任务。



$$\min_P J(P) = \frac{1}{N} \sum_{i=1}^{W_c \times H_c} \sum_{j=1}^{W_s \times H_s} A_{ij}^{cs} \|\phi_i(F_{cs}) - \phi_j(F_s)\|_2^2,$$

J. Huo, S. Jin, W. Li, J. Wu, Y.-K. Lai, Y. Shi, and Y. Gao, "Manifold alignment for semantically aligned style transfer," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 14 861–14 869.

Rethinking the content represention for AST


Content


Style


Round 1


Round 20


Content


Style


Round 1


Round 20

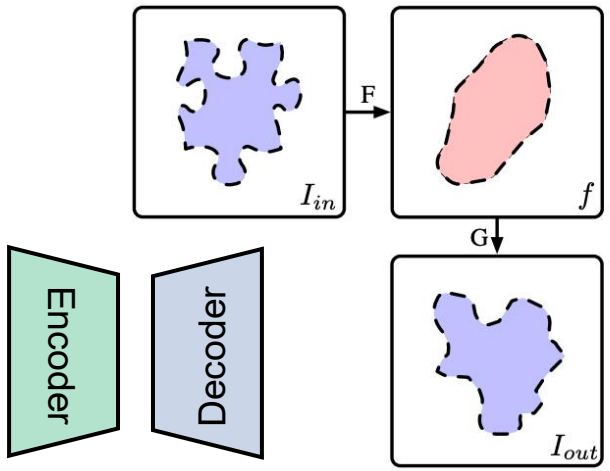

Content

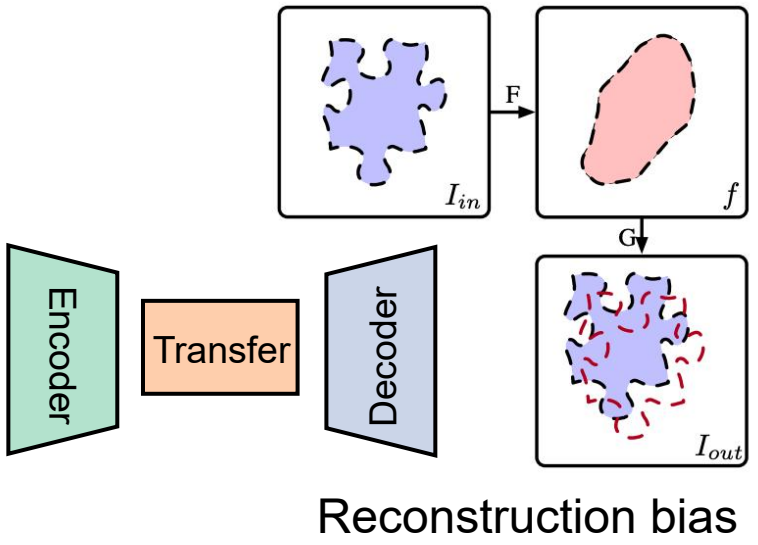

Style


Round 1

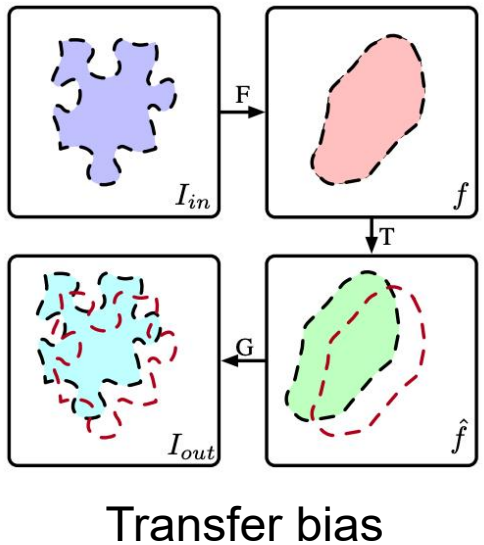

Round 20


Round 40





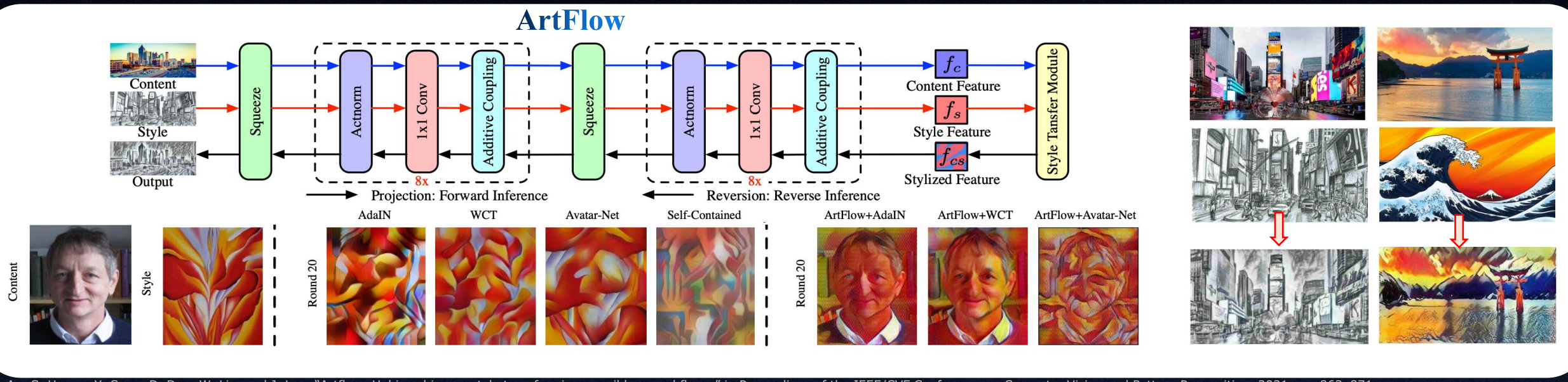
Reconstruction bias



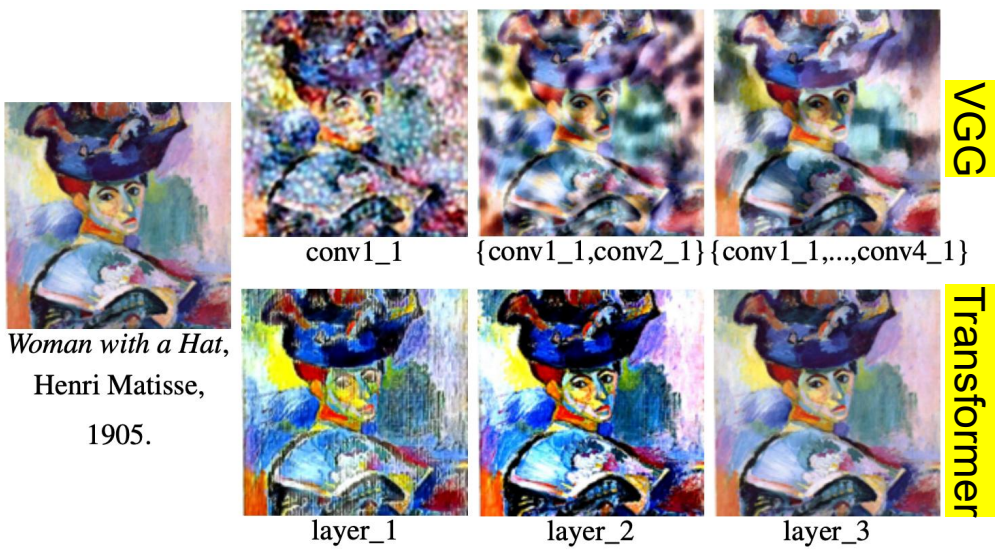
Transfer bias

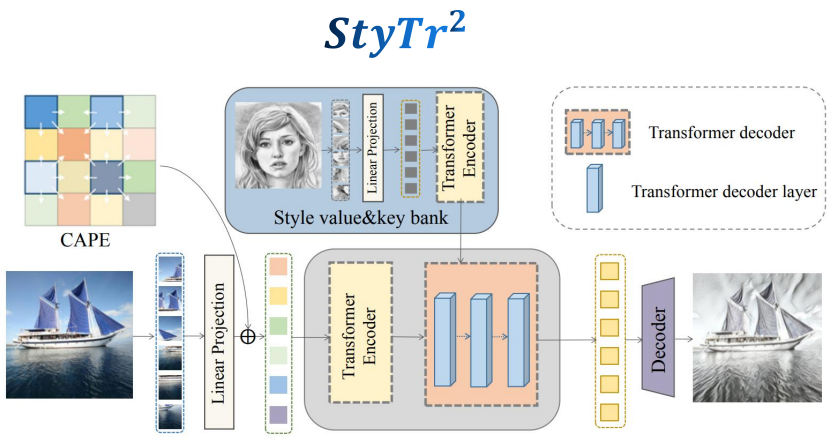
D. Y. Park and K. H. Lee, "Arbitrary style transfer with style attentional networks," in proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 5880–5888.

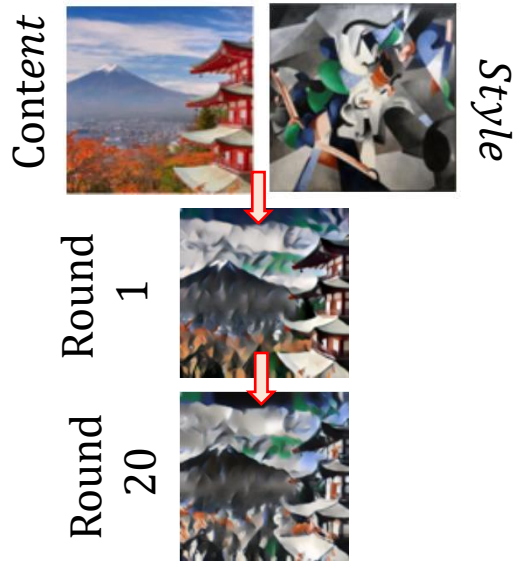
Rethinking the content represention for AST



J. An, S. Huang, Y. Song, D. Dou, W. Liu, and J. Luo, "Artflow: Unbiased image style transfer via reversible neural flows," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 862–871.







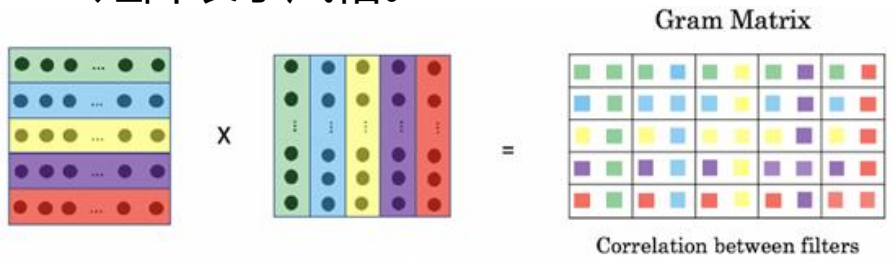
Y. Deng, F. Tang, W. Dong, C. Ma, X. Pan, L. Wang, and C. Xu, "Stytr2: Image style transfer with transformers," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 11 326–11 336.

How about “style”？

传统风格迁移任务中的风格表示方式

- **用低层次手工标注特征**：基于纹理合成和图像滤波，例如基于笔画的渲染和图像过滤通常使用低级的手工制作的特征。

- **用Gram矩阵作为风格表征**：随着深度卷积神经网络的发展，Gatys 等首次提出使用表示卷积网络提取到的特征图通道间相关性的Gram矩阵表示风格。

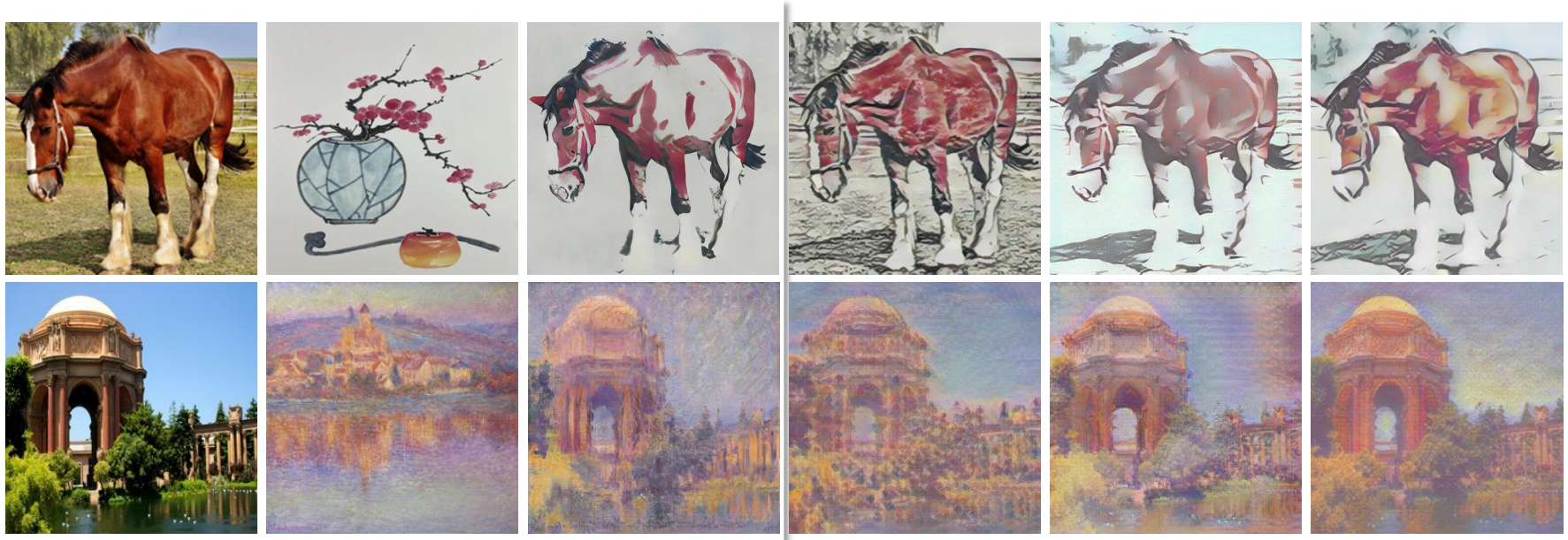


- **用特征二阶统计量作为风格表征**：Li等通过理论推导发现，最小化图像间的Gram统计量差异，等价于令图像特征表达二阶统计分布尽可能接近。并且，作者提出用逐通道间特征的均值和方差作为另一种风格表征形式。

基于自适应对比学习的任意图像风格迁移

创新点

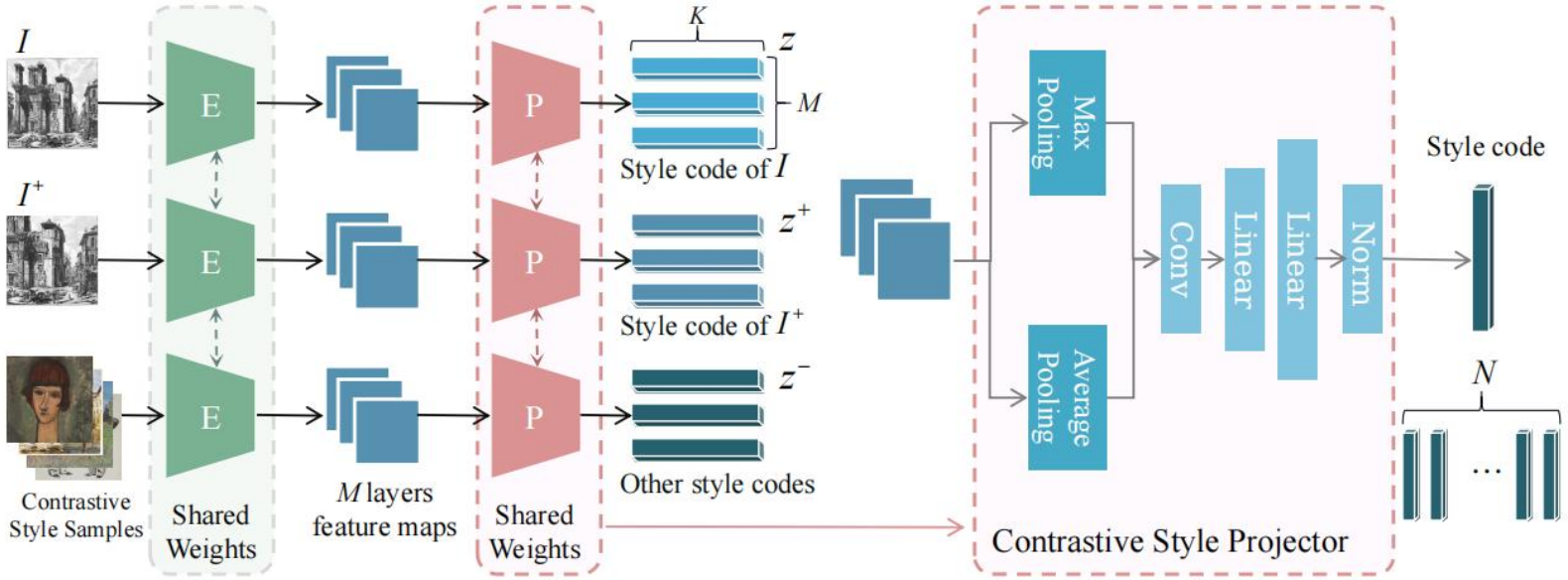
提出一种新颖的自适应对比学习任意图像风格迁移框架，基于对比学习挖掘图像间相似性关系，获得图像艺术风格统一表示模型；提出一种双线并行的风格表示和风格一致性度量方法，在一次训练是同时完成风格表示和风格迁移；支持基于编码器-解码器、流模型和Transformer的等多种风格迁移模块嵌入。



内容图　风格图　我们的结果　AdaAttN ICCV2021　ArtFlow CVPR2021　StyTr² CVPR2022

基于自适应对比学习的任意图像风格迁移

核心：对比学习+风格编码



梵高绘画在风格空间的分布可视化结果

艺术图像的增强版本

最大互信息

“靠近”

目标艺术图像

最小互信息

“远离”

其他艺术图像

基于自适应对比学习的任意图像风格迁移

核心：对比学习+风格度量

艺术图像的增强版本

最大互信息

“靠近”

目标艺术图像

最小互信息

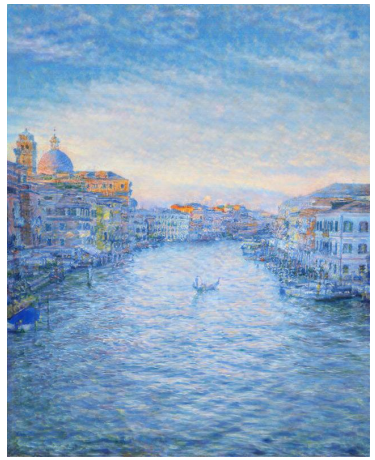
“远离”

其他艺术图像

目标艺术图像

最大互信息

“靠近”



生成风格化图像

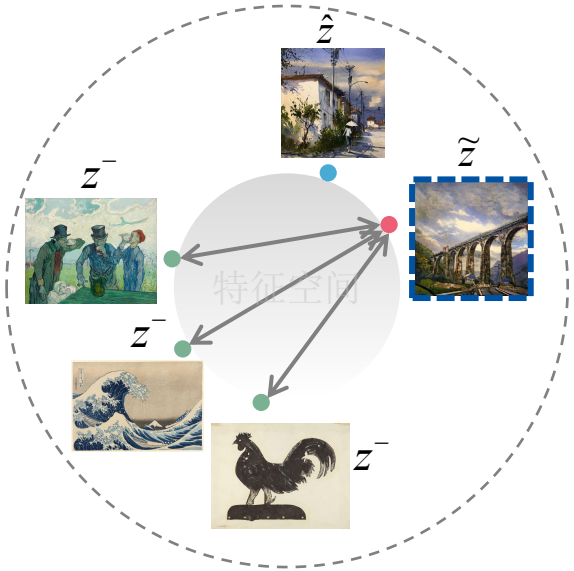
最小互信息

“远离”

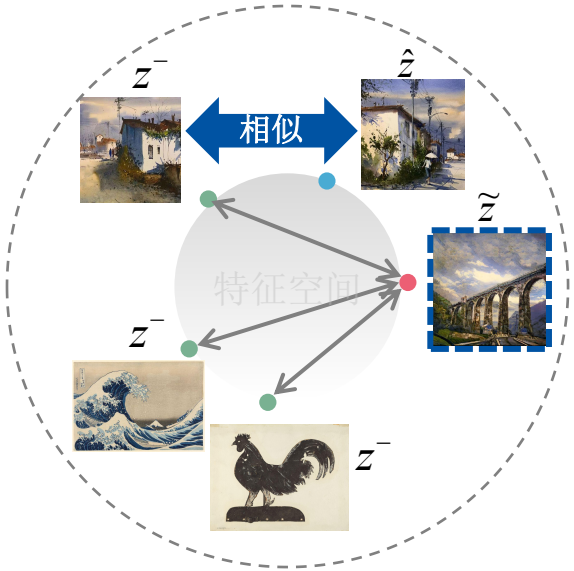
其他艺术图像

基于自适应对比学习的任意图像风格迁移

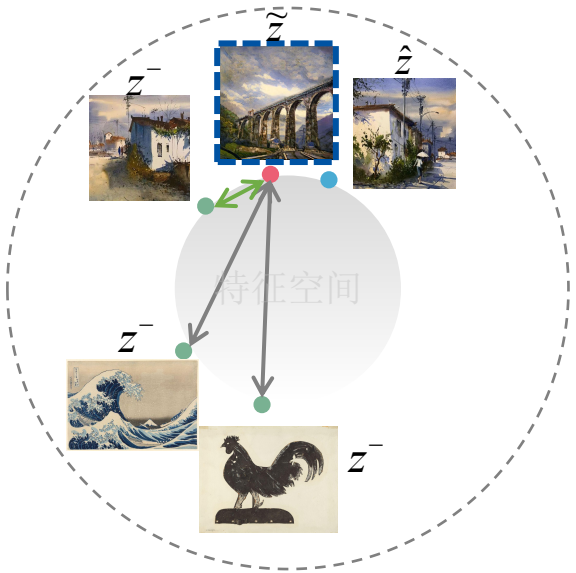
诀窍：自适应温度系数



(a) 固定温度系数
场景1



(b) 固定温度系数
场景2



(c) 自适应温度系数
场景2

输入相关的自适应温度系数对比损失

$$\mathcal{L}_{contra}^{G} = -\sum_{i=1}^{M} \log \frac{\exp(s_i^+/\tau^+)}{\exp(s_i^+/\tau^+) + \sum_{j=1}^{N} \exp(s_{ij}^-/\tau^-)}$$

$$\frac{\partial \mathcal{L}_{contra}^{G}}{\partial s_i^+} = -\sum_{i=1}^{M} \frac{1}{\tau^+} \cdot \frac{\sum_{j=1}^{N} \exp(s_{ij}^-/\tau^-)}{\exp(s_i^+/\tau^+) + \sum_{j=1}^{N} \exp(s_{ij}^-/\tau^-)}$$

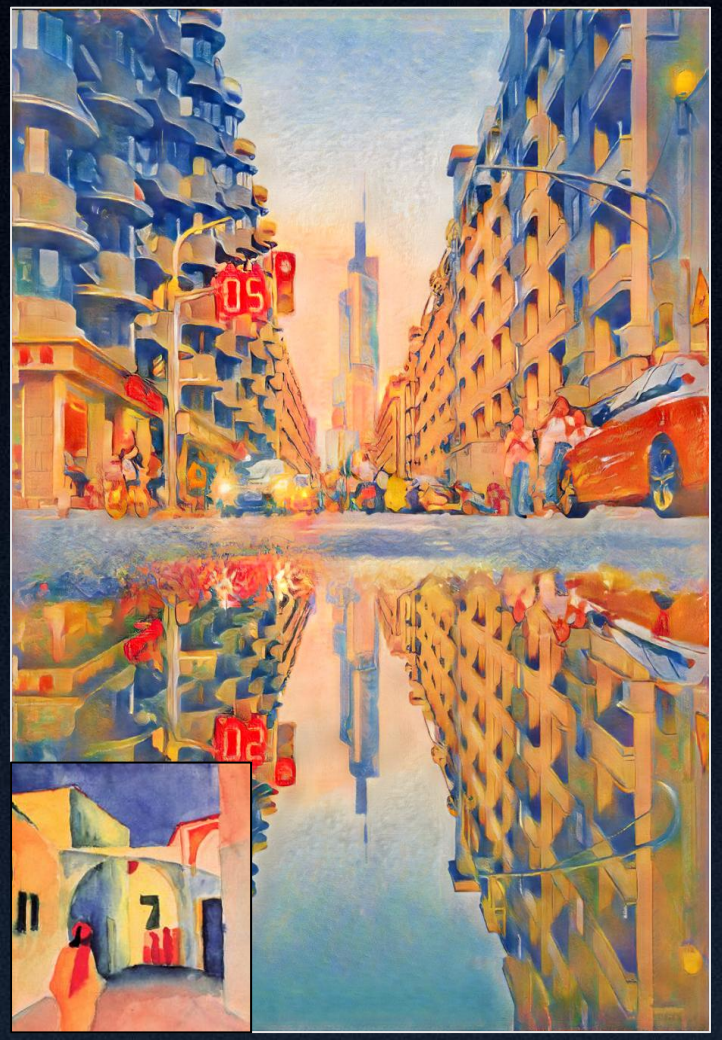
$$\frac{\partial \mathcal{L}_{contra}^{G}}{\partial s_{ij}^-} = -\sum_{i=1}^{M} \frac{1}{\tau^-} \cdot \frac{\exp(s_{ij}^-/\tau^-)}{\exp(s_i^+/\tau^+) + \sum_{j=1}^{N} \exp(s_{ij}^-/\tau^-)}$$

内容图

野兽派

Yuxin Zhang, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Changsheng Xu: Domain Enhanced Arbitrary Image Style Transfer via Contrastive Learning. **SIGGRAPH** 2022: 12:1-12:8

内容图



奥费主义

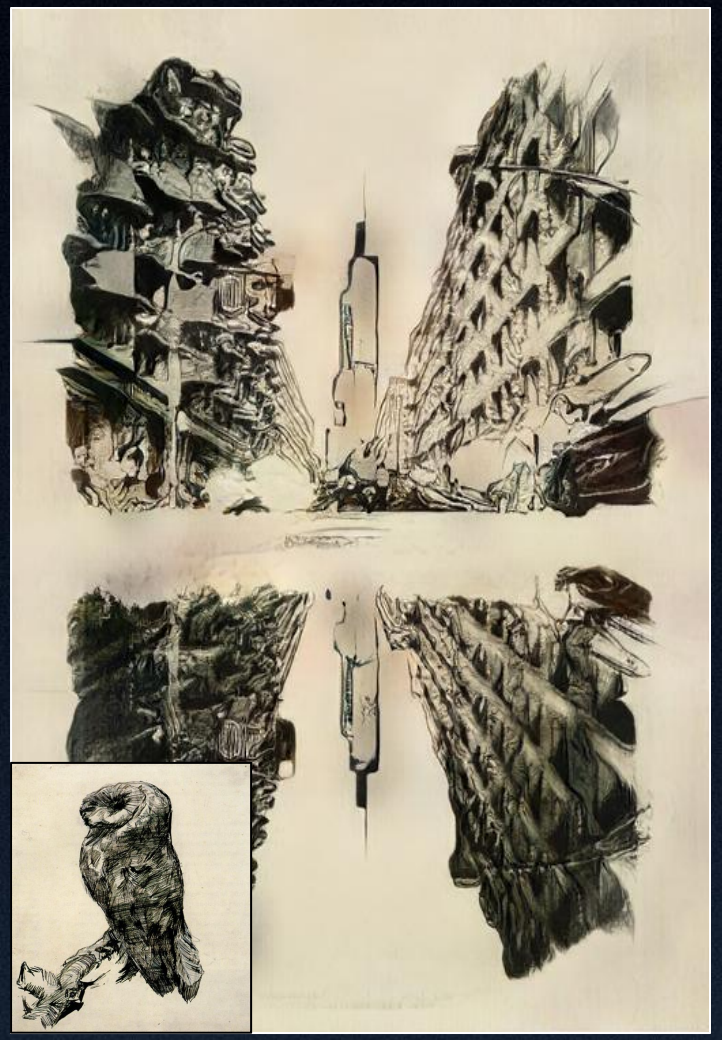
Yuxin Zhang, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Changsheng Xu: Domain Enhanced Arbitrary Image Style Transfer via Contrastive Learning. **SIGGRAPH** 2022: 12:1-12:8

内容图

浮世绘

Yuxin Zhang, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Changsheng Xu: Domain Enhanced Arbitrary Image Style Transfer via Contrastive Learning. **SIGGRAPH** 2022: 12:1-12:8

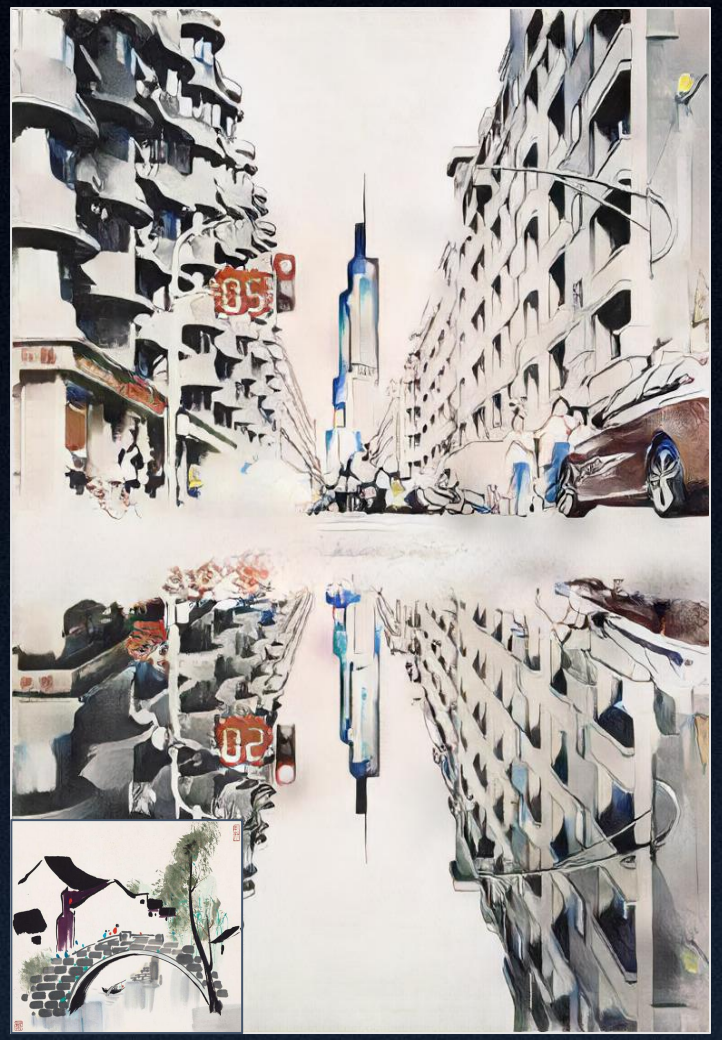
内容图



现实主义

Yuxin Zhang, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Changsheng Xu: Domain Enhanced Arbitrary Image Style Transfer via Contrastive Learning. **SIGGRAPH** 2022: 12:1-12:8

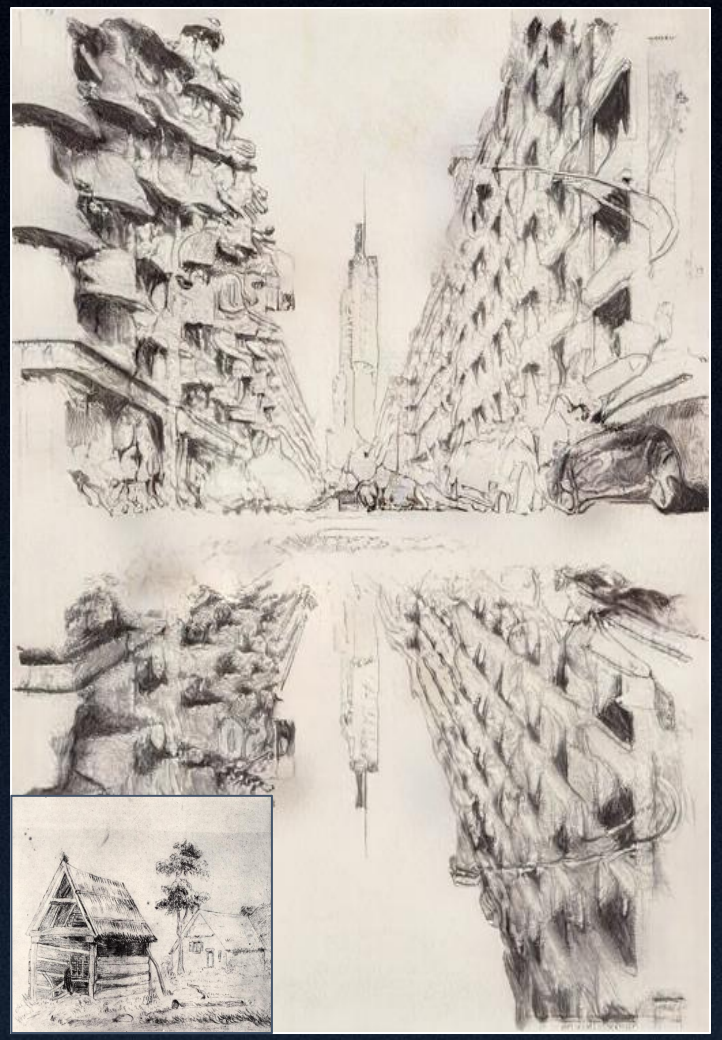
内容图



水墨

Yuxin Zhang, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Changsheng Xu: Domain Enhanced Arbitrary Image Style Transfer via Contrastive Learning. **SIGGRAPH** 2022: 12:1-12:8

内容图

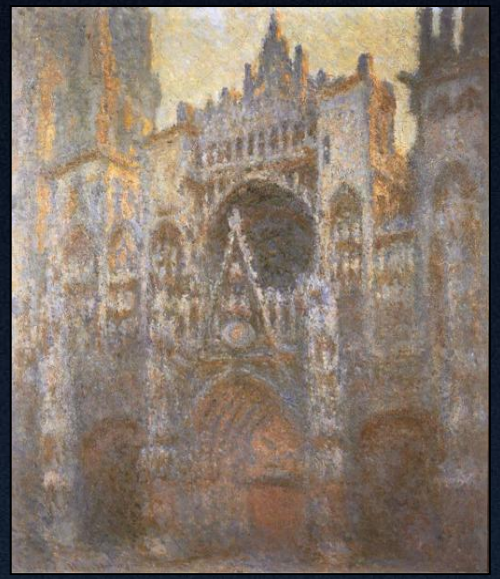


素描

Yuxin Zhang, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Changsheng Xu: Domain Enhanced Arbitrary Image Style Transfer via Contrastive Learning. **SIGGRAPH** 2022: 12:1-12:8

内容图

风格图



内容图

风格图


(a) Content
(b) Style
(c) AdaIN
(d) UCAST+AdaIN
(e) StyTr^2
(f) UCAST+StyTr^2
(g) ArtFlow
(h) UCAST+ArtFlow
Inputs
CNN-based method
ViT-based method
Flow-based method

电影《至爱梵高》125位画家7年时间手绘65000帧油画

1 *v.s.* 24

Repeat Image NST 24 times?

第 i 帧

第 i+1 帧

残差图

AdaIN

SANet

Xiaoyu Kong, Yingying Deng, Fan Tang et al: Exploring the Temporal Consistency of Arbitrary Style Transfer: A Channelwise Perspective. **IEEE TNNLS. 2023.**

视频风格化

Rethinking: 帧间不一致性从哪里产生?

Encoder-Transformer-Decoder

Encoder-Decoder

原始视频

AE 重建

Transformer

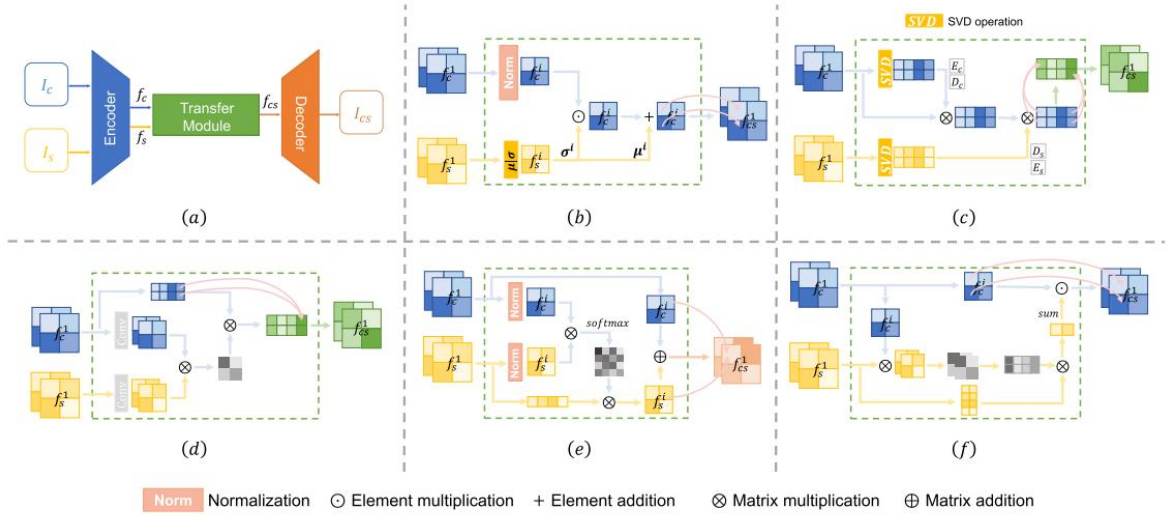


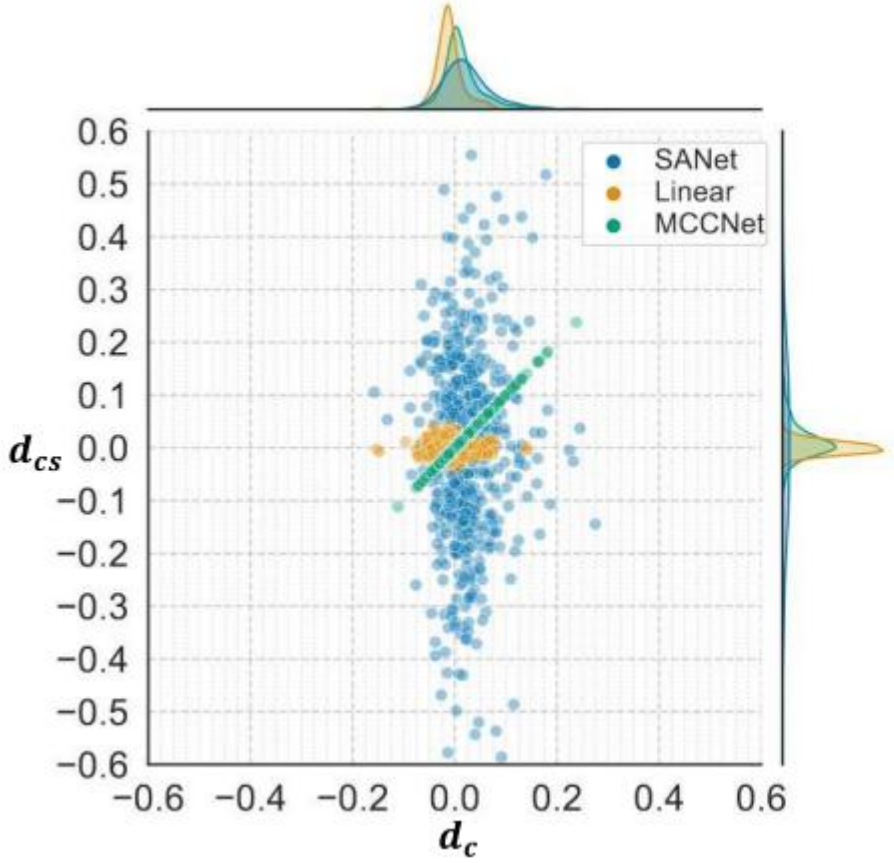
Fig. 2. Overall structure of different style transfer modules. The pink arrows indicate the alignment of features. (a) Encoder–transform–decoder. (b) AdaIN. (c) WCT. (d) Linear. (e) SANet. (f) MSSNet.

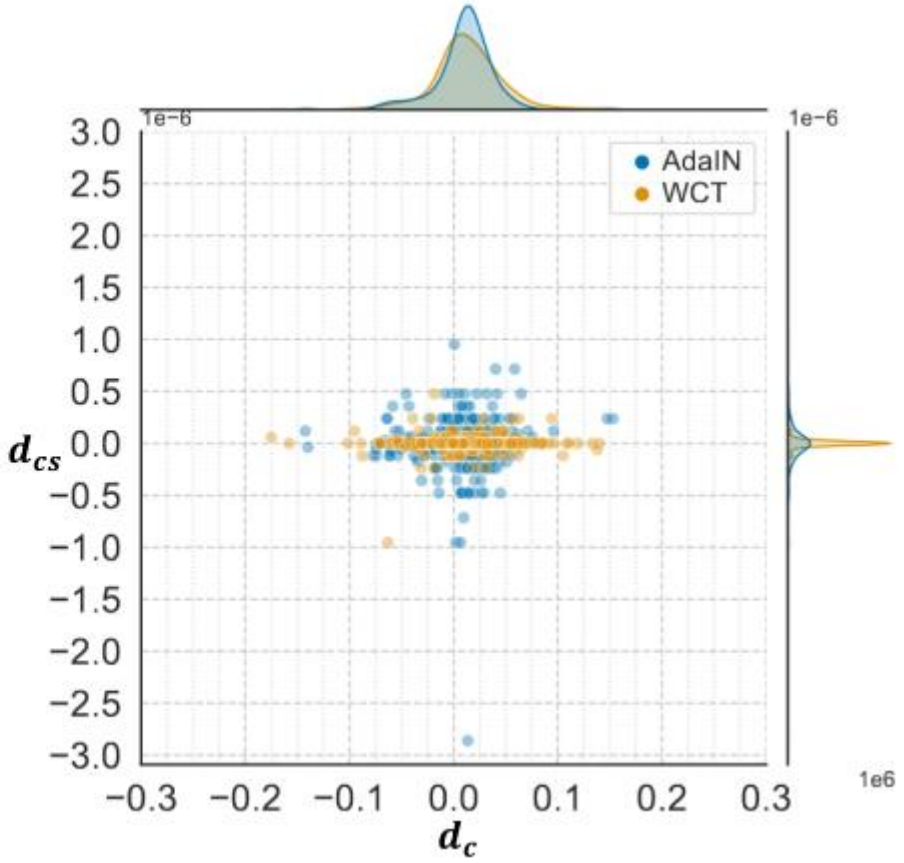
Xiaoyu Kong, Yingying Deng, Fan Tang et al: Exploring the Temporal Consistency of Arbitrary Style Transfer: A Channelwise Perspective. **IEEE TNNLS. 2023.**

视频风格化

Rethinking: 帧间不一致性从哪里产生?

Encoder-Transformer-Decoder



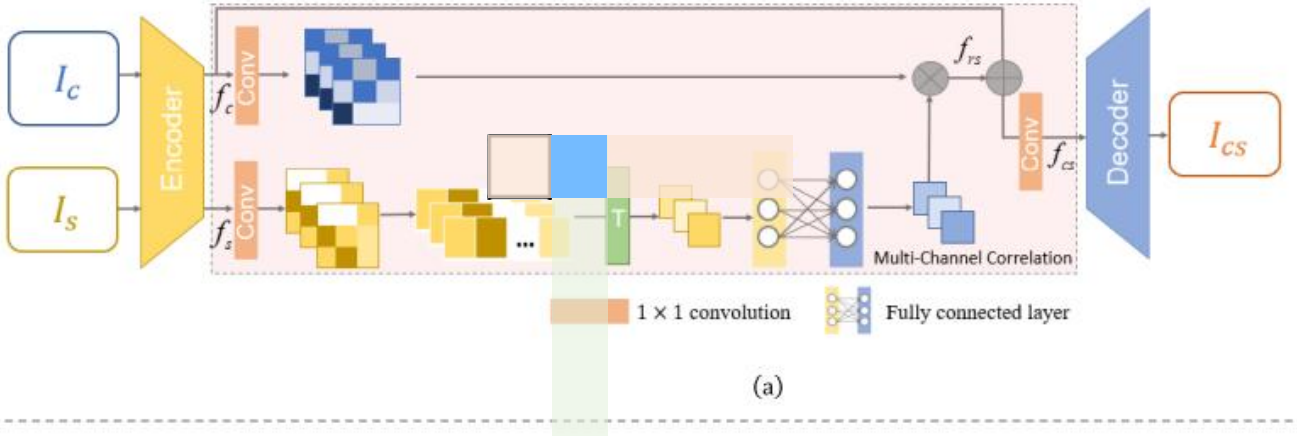


Xiaoyu Kong, Yingying Deng, Fan Tang et al: Exploring the Temporal Consistency of Arbitrary Style Transfer: A Channelwise Perspective. **IEEE TNNLS. 2023.**

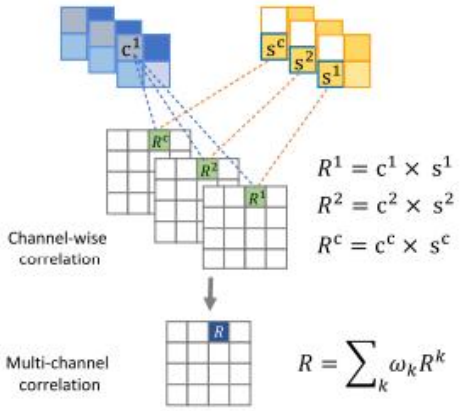
视频风格化

基于多通道相关性的视频风格化方法

$$f_{cs} = F(f_c, f_s) \propto exp(f(f_c), g(f_s))h(f_s)$$



(a)



(b)

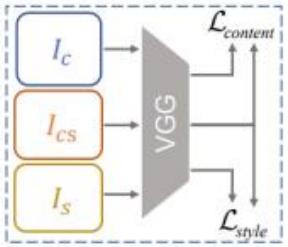
$$CO^i = f_c^{iT} \otimes f_s^i.$$

$$f_{rs}^i = f_s^i \otimes CO^{iT} = \|f_s^i\|_2 f_c^i$$

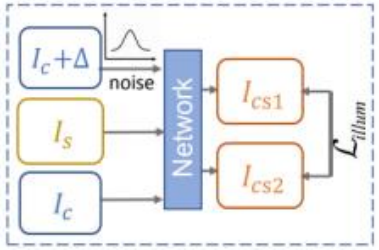
$$f_{cs}^i = f_c^i + f_{rs}^i = (1 + \|f_s^i\|_2) f_c^i$$

$$f_{cs}^i = f_c^i + f_{rs}^i = (1 + \sum_{k=1}^{C} w_k \|f_s^k\|_2) f_c^i$$

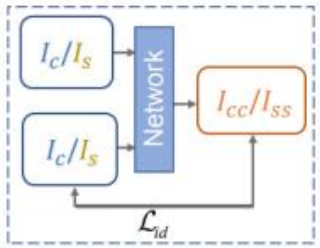
$$f_{cs} = K f_c$$



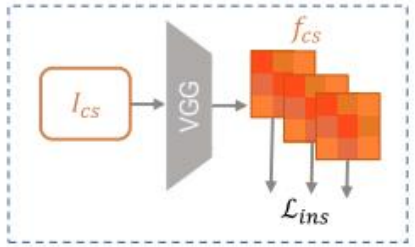
Perceptual loss



Illumination loss



Identity loss

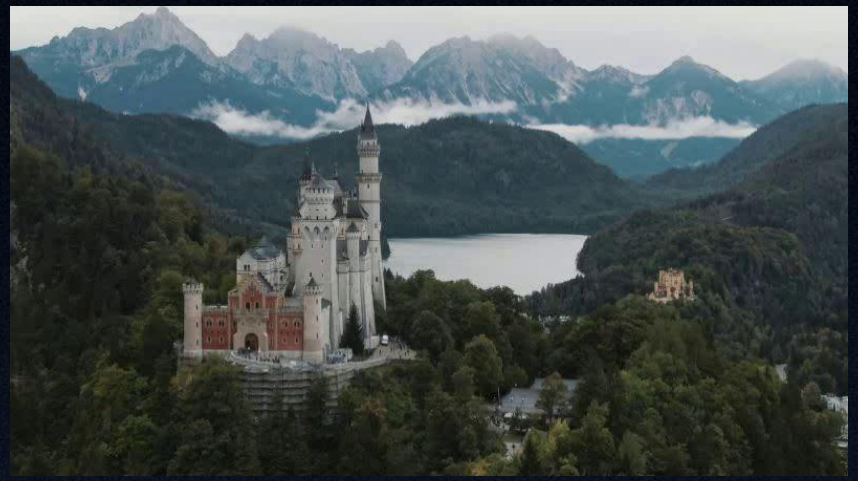


Inner channel similarity loss

Xiaoyu Kong, Yingying Deng, Fan Tang et al: Exploring the Temporal Consistency of Arbitrary Style Transfer: A Channelwise Perspective. **IEEE TNNLS. 2023.**

风格图

风格图



风格图

风格化，就这？

有点意思，不多。

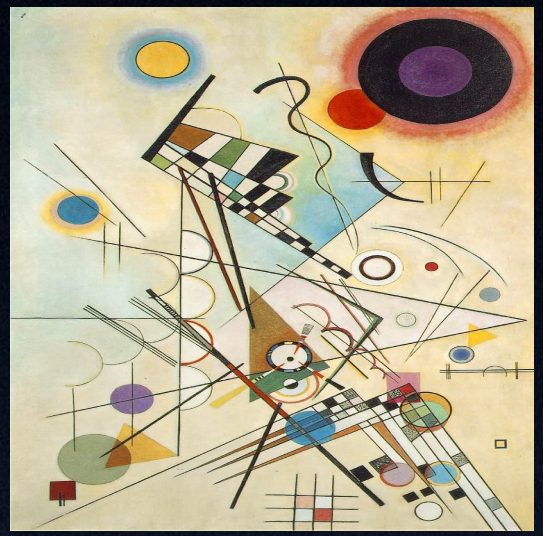
“罗曼-朱安多的油画
《云中飞翔的蒸汽朋克屋》”

Nisha Huang, Fan Tang, Weiming Dong, Changsheng Xu: Draw Your Art Dream: Diverse Digital Art Synthesis with Multimodal Guided Diffusion. ACM Multimedia 2022

“ 一幅描绘太阳系的绘画 ”

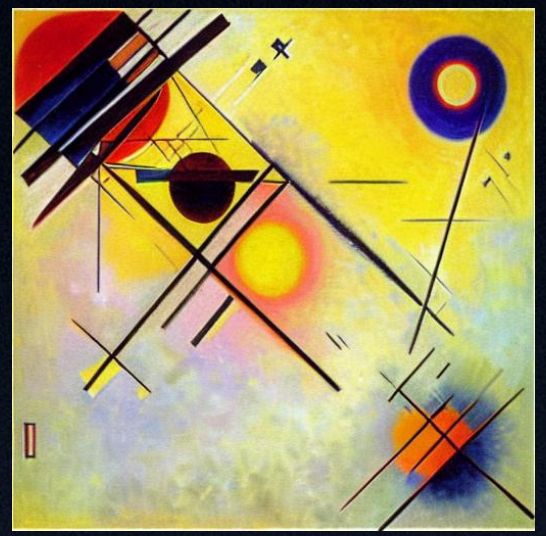
Painting name:
Small Worlds IV

“ 一幅描绘太阳系的绘画，
参考图片的风格 ”



Reference

The space ship



The Sun

Reference

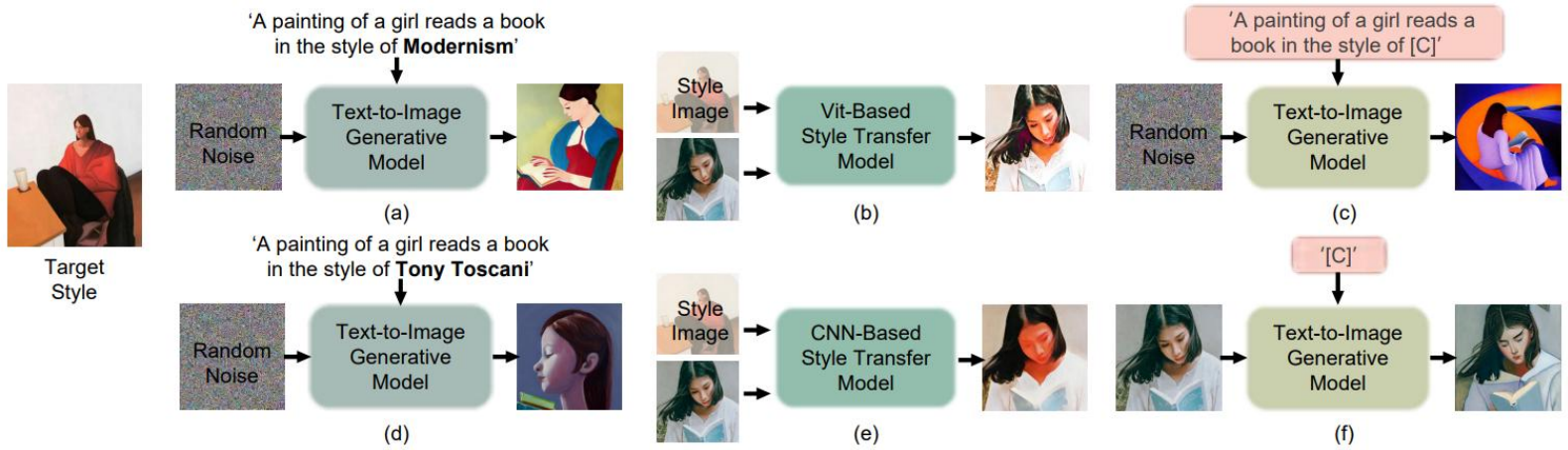
The Solar System

Flowers

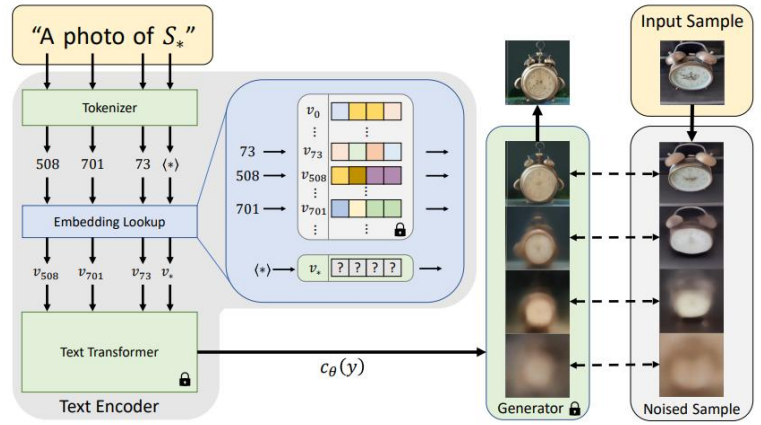
Zhang Y, Huang N, Tang F, et al. Inversion-Based Creativity Transfer with Diffusion Models[J]. arXiv preprint arXiv:2211.13203, 2022.

基于扩散模型的图像风格化

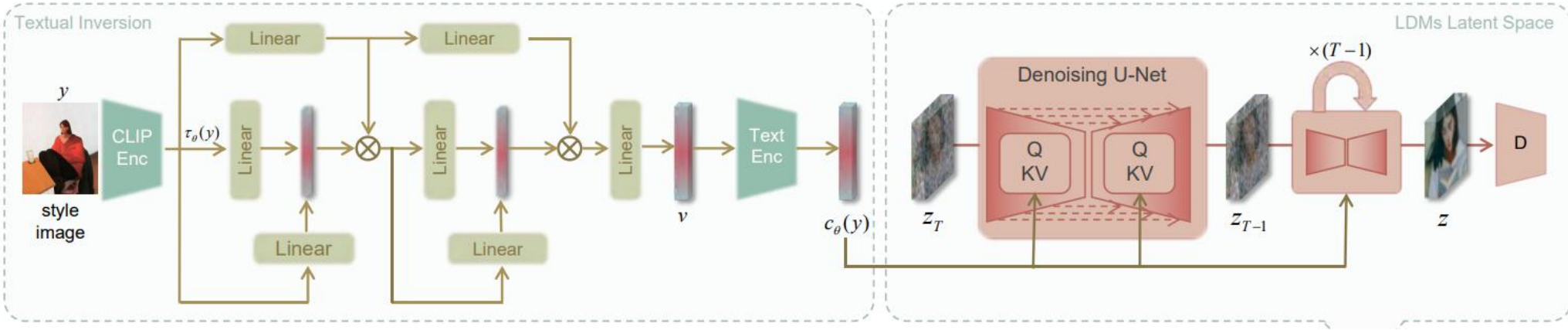
核心: 基于文本反演的风格化图像生成



各种风格化框架

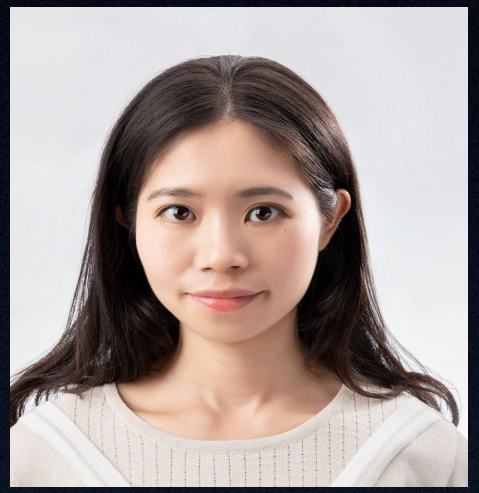


基于TI的概念定制化



Yuxin Zhang, NishaHuang, Fan Tang, Chongyang Ma, Haibin Huang, Weiming Dong, Changsheng Xu. Inversion-Based Creativity Transfer with Diffusion Models. CVPR 2023

Yuxin Zhang, NishaHuang, Fan Tang, Chongyang Ma, Haibin Huang, Weiming Dong, Changsheng Xu. Inversion-Based Creativity Transfer with Diffusion Models. CVPR 2023



Yuxin Zhang, NishaHuang, Fan Tang, Chongyang Ma, Haibin Huang, Weiming Dong, Changsheng Xu. Inversion-Based Creativity Transfer with Diffusion Models. CVPR 2023

Yuxin Zhang, NishaHuang, Fan Tang, Chongyang Ma, Haibin Huang, Weiming Dong, Changsheng Xu. Inversion-Based Creativity Transfer with Diffusion Models. CVPR 2023

基于扩散模型的图像风格化

进阶：可解释的逐时间步反演

属性增加

Layout/Content/Color　　Material


ball


+ yarn
in step 0-200


ball of yarn


+ yarn
in step 200-400


+ yarn
in step 400-600


+ yarn
in step 600-800


+ yarn
in step 800-1000

标准文字条件空间

$$\mathcal{P} = \{p\}$$

对于一次图像生活过程，原始条件空间仅包含一个文字条件样本

ProSpect文字条件空间

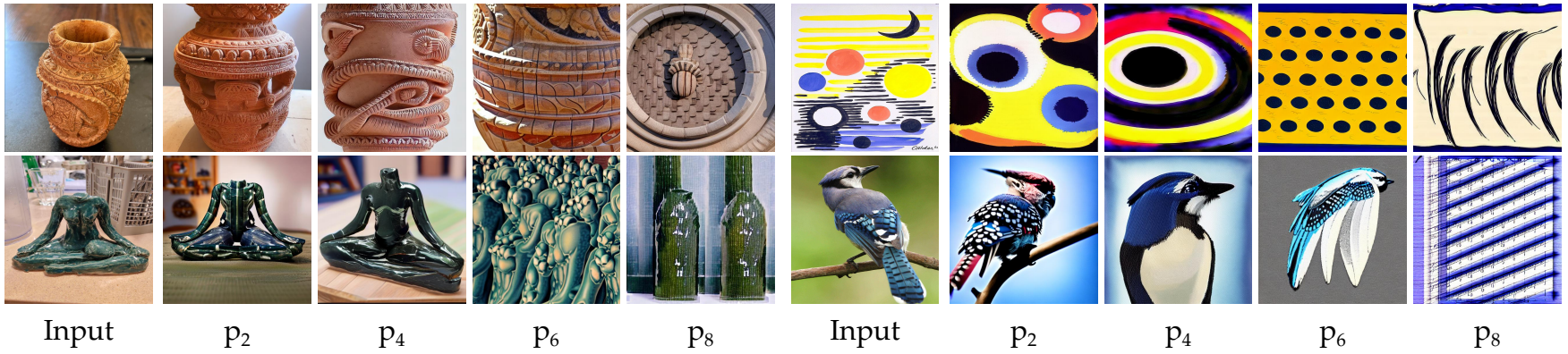
$$\mathcal{P}^* = \{p_1, p_2, ..., p_n\}$$

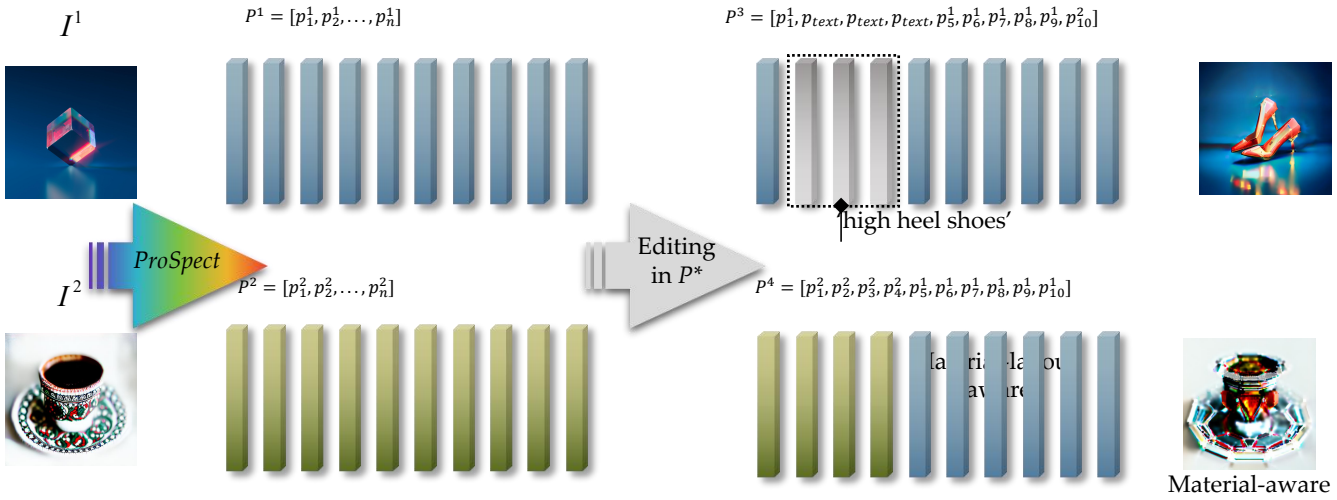
从时间步扩展的文字条件空间，对于一次图像生活过程，原始条件空间包含一组文字条件样本

Yuxin Zhang, Weiming Dong, Fan Tang, Nisha Huang, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Oliver Deussen, Changsheng Xu. Prompt spectrum for attribute-aware personalization of diffusion models. ACM Transactions on Graphics (SIGGRAPH ASIA 2023)

基于扩散模型的图像风格化

进阶：可解释的逐时间步反演

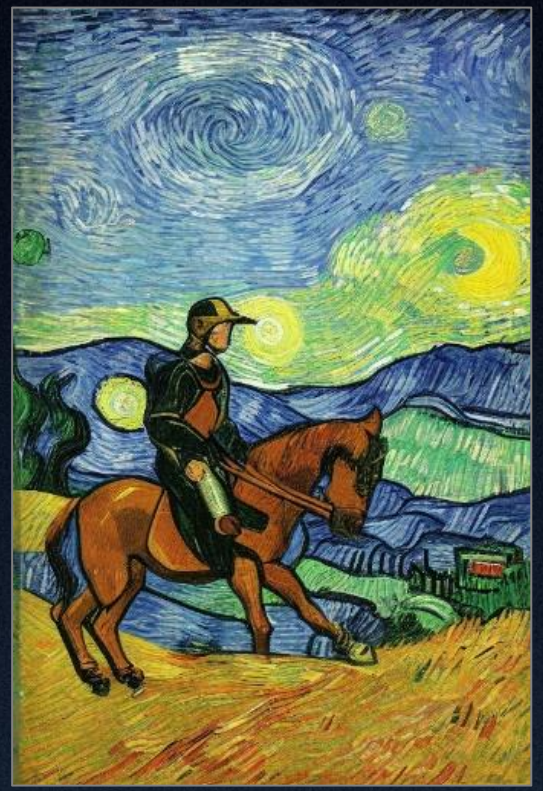




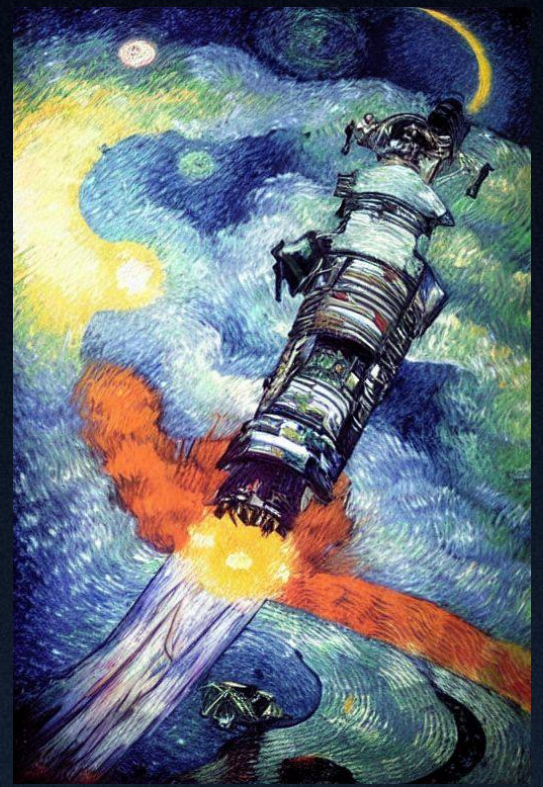
Yuxin Zhang, Weiming Dong, Fan Tang, Nisha Huang, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Oliver Deussen, Changsheng Xu. Prompt spectrum for attribute-aware personalization of diffusion models. ACM Transactions on Graphics (SIGGRAPH ASIA 2023)

基于扩散模型的图像风格化

梵高： *The starry night*



‘骑马的宇航员’



‘宇宙飞船’

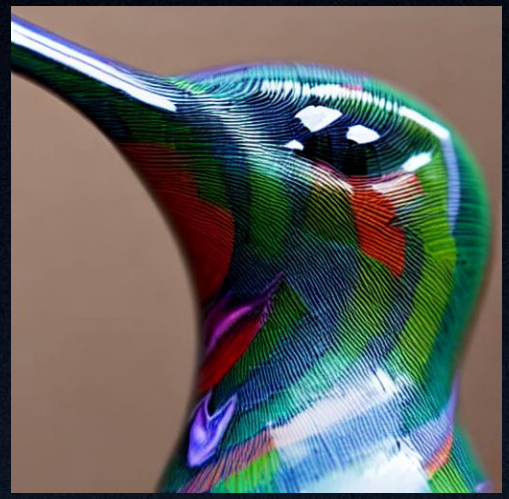
Yuxin Zhang, Weiming Dong, Fan Tang, Nisha Huang, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Oliver Deussen, Changsheng Xu. Prompt spectrum for attribute-aware personalization of diffusion models. ACM Transactions on Graphics (SIGGRAPH ASIA 2023)

基于扩散模型的图像风格化

输入图

我们的结果

‘毛线’



‘陶瓷’

‘孔雀羽毛’

InstructPix2Pix

‘变成毛线’

‘变成陶瓷’

‘变成孔雀羽毛’

Yuxin Zhang, Weiming Dong, Fan Tang, Nisha Huang, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Oliver Deussen, Changsheng Xu.
Prompt spectrum for attribute-aware personalization of diffusion models. ACM Transactions on Graphics (SIGGRAPH ASIA 2023)

基于扩散模型的图像风格化

文字到图片生成

参考图

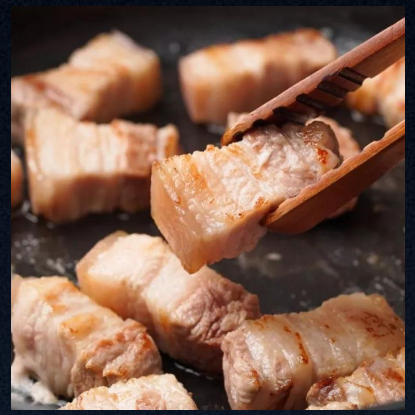
‘草莓杯子蛋糕 ‘

参考图

‘草莓杯子蛋糕 ‘

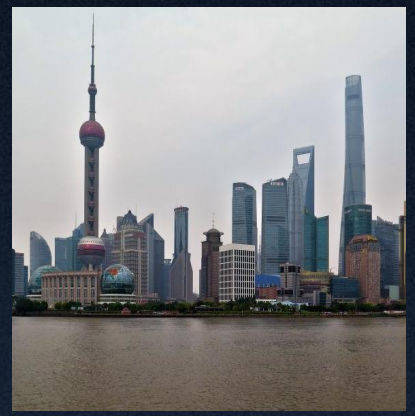
图片到图片生成

输入图片



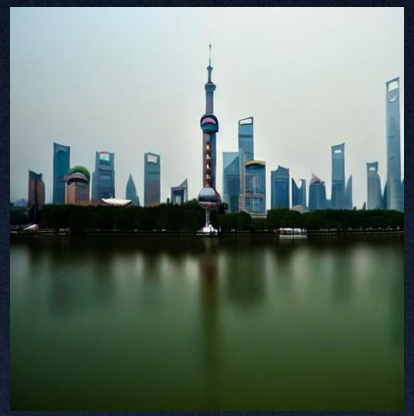
布局引导图

生成结果



输入图片

布局引导图



生成结果
(轴对称、反射的视觉效果)

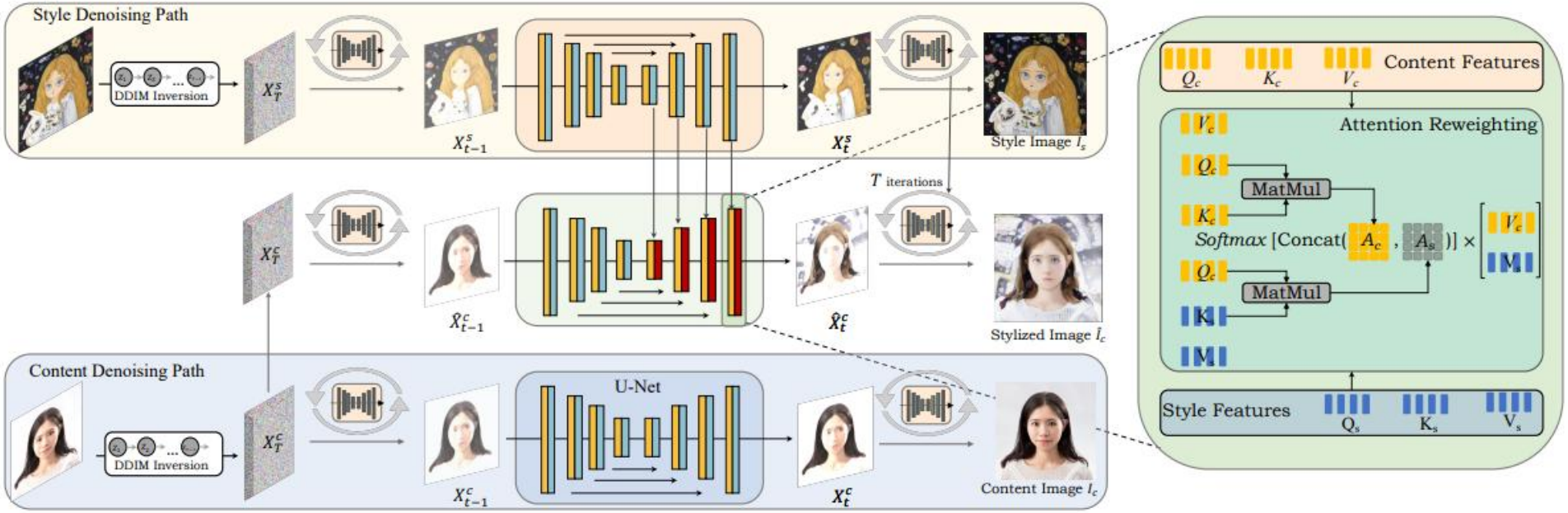
Yuxin Zhang, Weiming Dong, Fan Tang, Nisha Huang, Haibin Huang, Chongyang Ma, Tong-Yee Lee, Oliver Deussen, Changsheng Xu.
Prompt spectrum for attribute-aware personalization of diffusion models. ACM Transactions on Graphics (SIGGRAPH ASIA 2023)

基于扩散模型的图像风格化

核心：大模型知识激活，training free

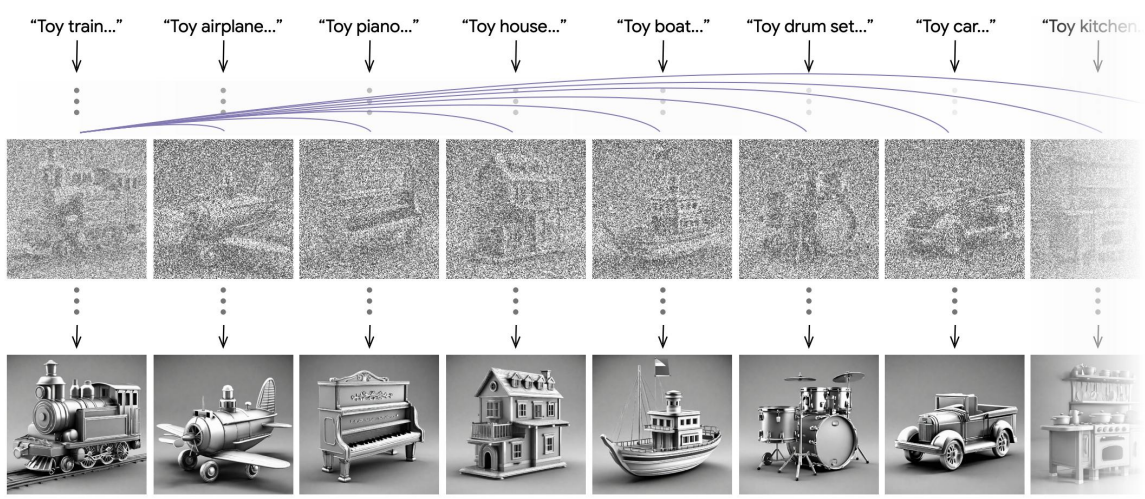
创新点

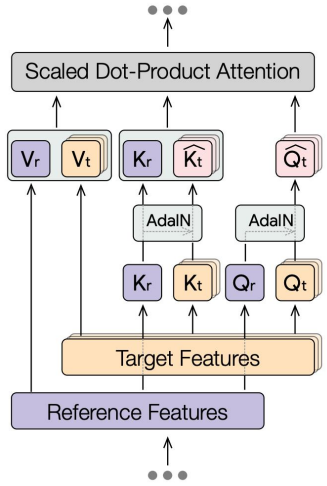
提出一种新颖的无微调风格化图像生成框架，基于风格内容双分支重建，获得图像风格、内容表示；提出一种基于注意力重加权的跨域特征交互方法，通过在潜空间设计特征交互策略，实现风格图像和内容图像的融合。

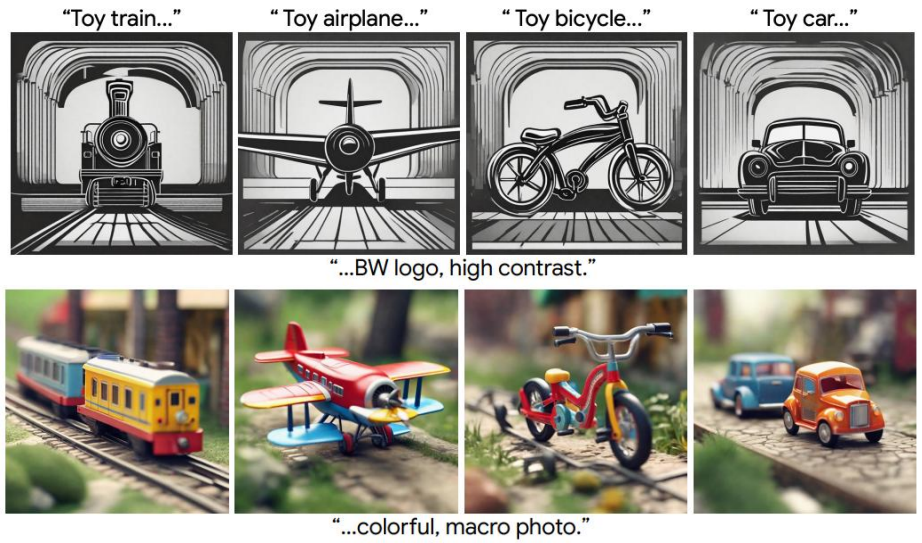


Yingying Deng, Xiangyu He, Fan Tang, Weiming Dong. Zero-shot Style Transfer via Attention Reweighting. CVPR 2024.

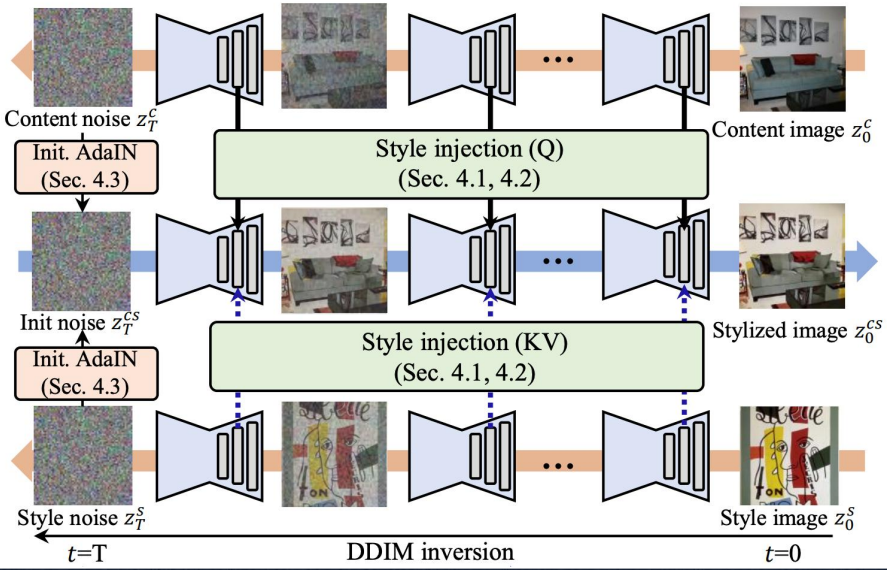
Parallel works

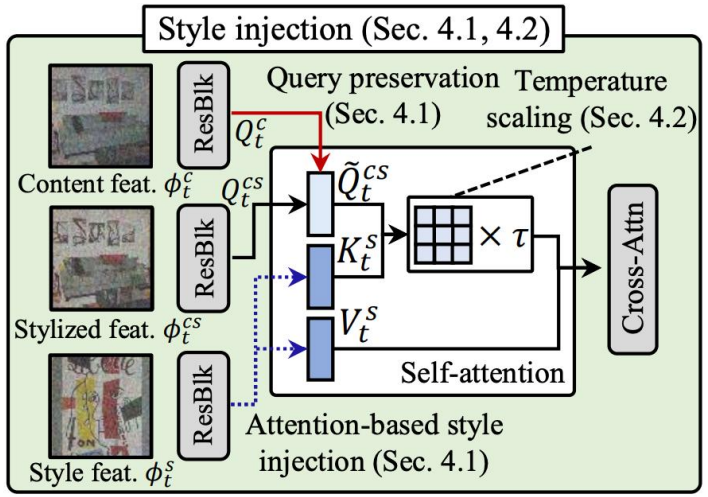






HERTZ A., VOYNOV A., FRUCHTER S., COHEN-OR D.: Style aligned image generation via shared attention, CVPR 2024. URL: https://arxiv.org/abs/2312.02133, arXiv:2312. 02133. 1

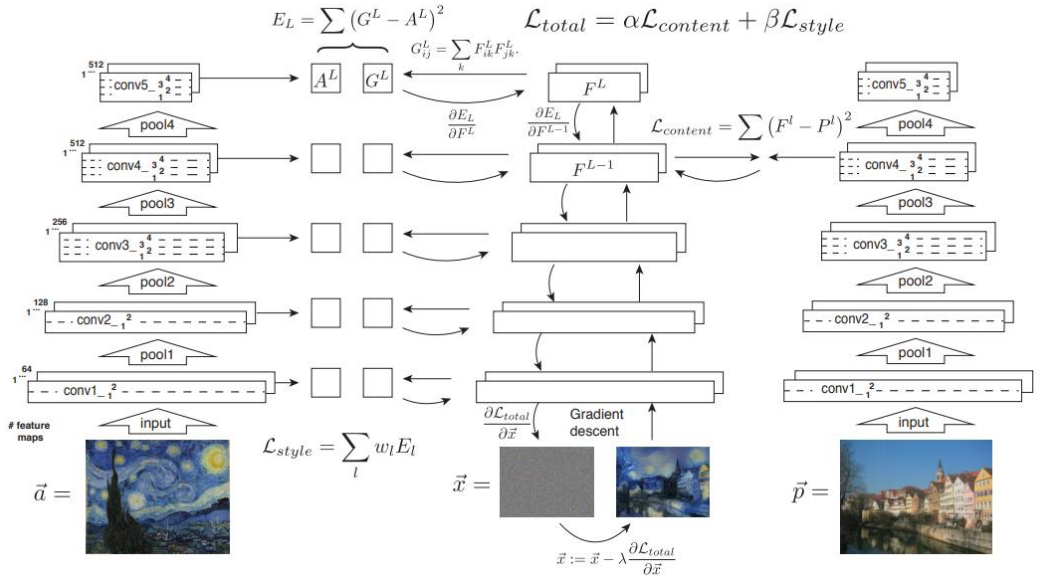




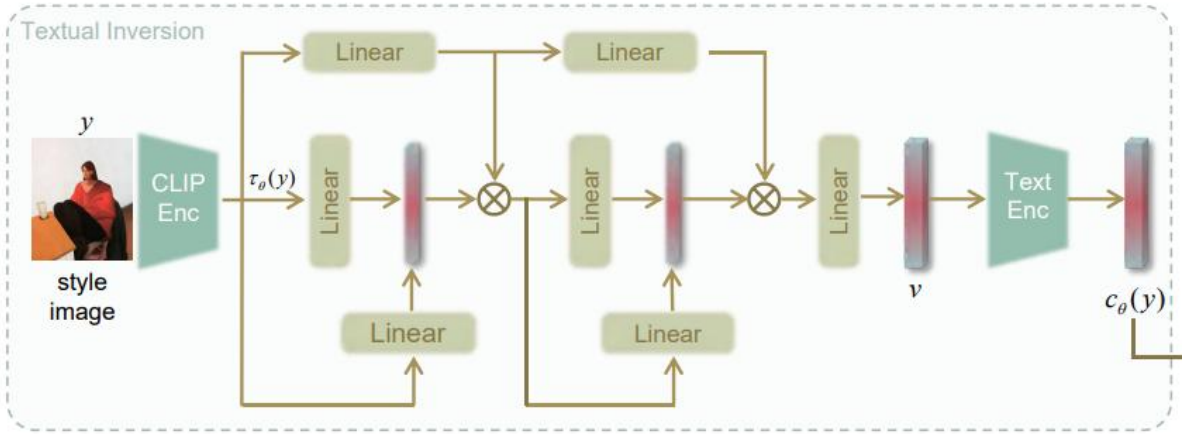
CHUNG J., HYUN S., HEO J.-P.: Style injection in diffusion: A training-free approach for adapting large-scale diffusion models for style transfer, CVPR 2024. URL: https://arxiv.org/abs/2312.09008, arXiv:2312.09008. 1

Style tranfer v.s. GenAI

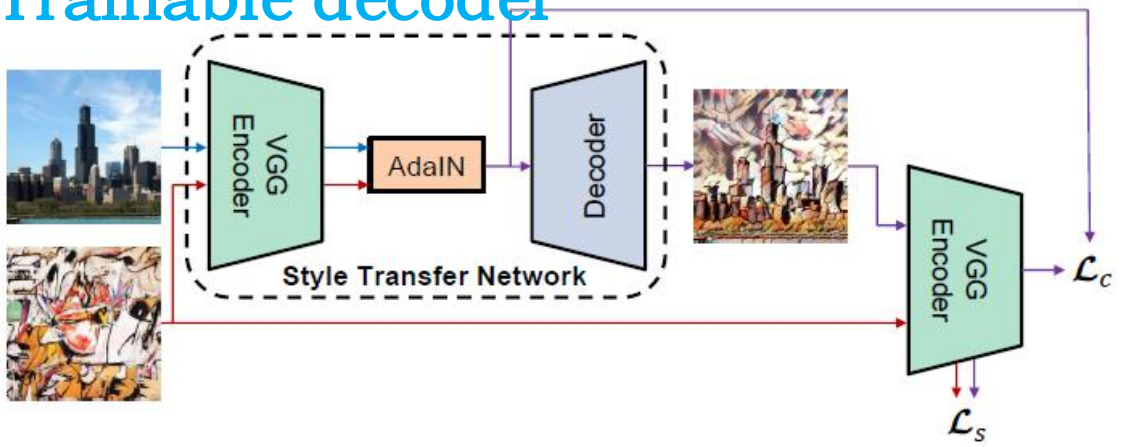
From NST to AdaIN, from InST to Z-Star



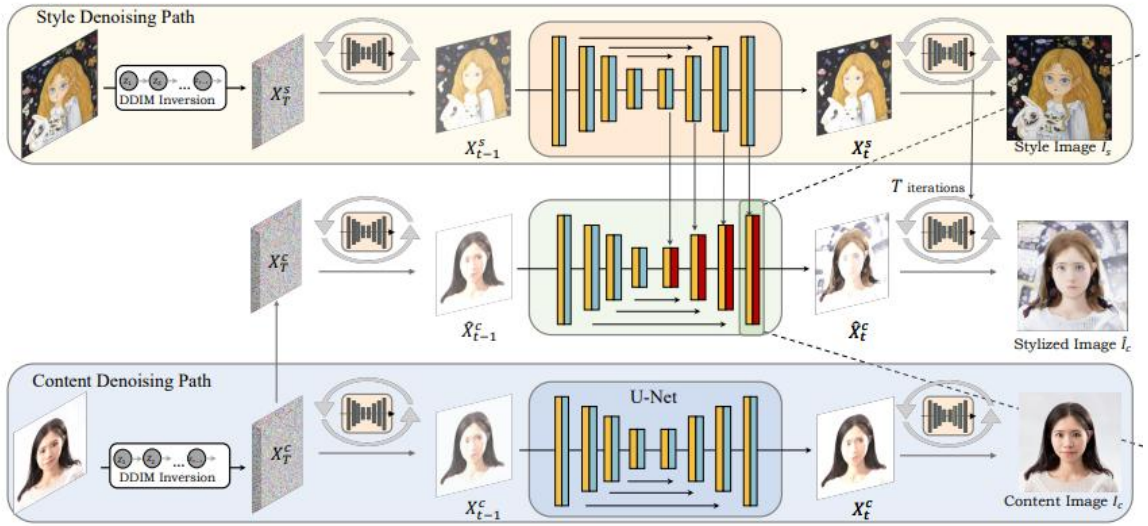
Pretrained+优化



Trainable decoder



先验激发



展 望

Take away for NST

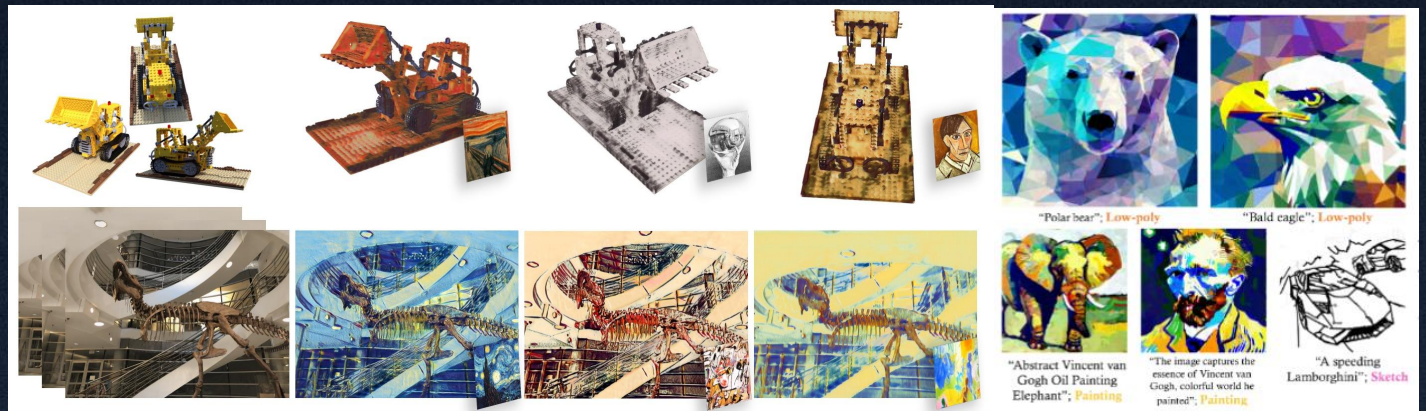
风格迁移

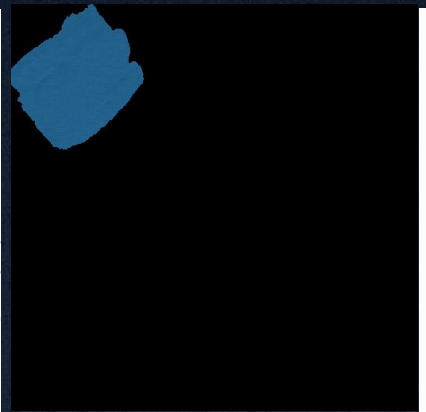
风格理解 → 传统图像风格化方法基于<content-style> pair，后续发展出collection-based, text-based，如果帮助用户表达创作需求?

知识激发 → 基于预训练模型，研究概念注入方法
DB，TI，LoRa，TrainingFree

Novel views与视频 → 结合Nerf、3DGS
结合视频生成先验

SVG stylization → model-based stylization
矢量化风格化





谢谢大家！

欢迎交流与批评指正