

1 Introduction

基于神经网络的风格迁移 (也即神经风格迁移) 是深度学习与神经网络在艺术领域做出的尝试. 将具有艺术信息的图像与具有真实场景或人像的内容图像同时输入到网络中, 最终得到具有艺术图像艺术特征以及内容图像内容特征的风格化图像, 是风格迁移最主要的任务. 与大多数计算机视觉任务一样, 风格迁移随着神经网络的兴起而兴起. 神经风格迁移真正走入大众视野可以追溯到由 Gatys 等人 [5] 于 2016 年使用 VGG 中特定层作为损失函数的核心, 以图像像素作为优化对象进行风格迁移工作.

尽管当前工作已经从效率 [7], 质量 [], 速度 [16], 多模态 [3] 等不同方向有了广泛的发展, 但是依旧无法处理具有不透明度信息的图像. 对于一般用户而言, 不透明度通道可能可有可无, 得到一张完整的风格化图像可以满足他们的需求, 但是对于艺术工作者来说, 一副完整作品需要使用多种背景透明的基本元素作为点缀. 这种基本元素往往具有不规则的外形, 而图像往往是以矩形的方式保存的, 所以透明背景在基本元素中也具有重要的作用—防止遮挡其他图像. 至今大部分风格迁移工作 [4, 6, 8–11, 13–15] 在读入具有不透明度信息的图像时, 大多舍弃不透明度通道, 而是以三通道张量 (R,G,B) 的方式处理他们. 在这中情况下, 原本透明部分由于不透明度通道的丢失, 导致其变成白色或黑色色块. 这些纯色色块会被风格迁移网络误认为是图像风格的一部分, 从而导致风格化图像中存在大量不属于风格信息的色块, 最终生成难以令人满意的迁移结果.

在这份工作中, 我们提出了 **PartAlpha**, 一种新颖的端到端风格迁移模型, 在能正常处理无 alpha 通道图像的同时, 专门考虑了具有 alpha 通道图像作为输入的情况. 为了克服传统神经风格迁移方法无法处理 alpha 通道的情况, 本文以部分卷积 [12] 为基础, 因为部分卷积可以很好的忽略图像中不被关心的像素. 但是传统的部分卷积依旧没有处理不透明度信息的方法, 本文为了克服这个困难, 我们使用了一种将 alpha 通道转化为数值在 0 1 之间的浮点掩膜的方法, 成功的将部分卷积扩展到 alpha 通道上. 基于这个方法, 我们还提出了一种浮点掩膜的自动生成手段, 以弥补部分图像不具有的 alpha 通道信息. 同时, 由于当前风格迁移工作数据集往往以 WikiArt [2] 和 MSCOCO [1] 为基础. 这两个数据集中的图像格式基本均为 jpeg, 不包含 alpha 通道信息. 为此, 本工作以 WikiArt 与 MSCOCO 为基础, 从完整图像中拆分外形不规则且具有不透明信息的场景与风格元素, 并以 png 的格式保存, 形成了名为 AlphaImg 的数据集.

参考文献

- [1] COCO - Common Objects in Context — cocodataset.org. <https://cocodataset.org/#home>. [Accessed 25-03-2025].
- [2] WikiArt.org - Visual Art Encyclopedia — wikiart.org. <https://www.wikiart.org/en/App/Home>. [Accessed 25-03-2025].
- [3] Namhyuk Ahn, Junsoo Lee, Chunggi Lee, Kunhee Kim, Daesik Kim, Seung-Hun Nam, and Kibeom Hong. Dreamstyler: Paint by style inversion with text-to-image diffusion models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 674–681, 2024.
- [4] Bin Fu, Junjun He, Jianjun Wang, and Yu Qiao. Neural transformation fields for arbitrary-styled font generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22438–22447, 2023.
- [5] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2414–2423, 2016.
- [6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Communications of the ACM*, 63(11):139–144, 2020.
- [7] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE international conference on computer vision*, pages 1501–1510, 2017.
- [8] Yongcheng Jing, Yezhou Yang, Zunlei Feng, Jingwen Ye, Yizhou Yu, and Mingli Song. Neural style transfer: A review. *IEEE transactions on visualization and computer graphics*, 26(11):3365–3385, 2019.
- [9] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands*,

- October 11-14, 2016, Proceedings, Part II 14*, pages 694–711. Springer, 2016.
- [10] Zhanghan Ke, Yuhao Liu, Lei Zhu, Nanxuan Zhao, and Rynson WH Lau. Neural preset for color style transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14173–14182, 2023.
- [11] Dongyang Li, Hao Luo, Pichao Wang, Zhibin Wang, Shang Liu, and Fan Wang. Frequency domain disentanglement for arbitrary neural style transfer. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 1287–1295, 2023.
- [12] Guilin Liu, Aysegul Dundar, Kevin J Shih, Ting-Chun Wang, Fitsum A Reda, Karan Sapra, Zhiding Yu, Xiaodong Yang, Andrew Tao, and Bryan Catanzaro. Partial convolution for padding, inpainting, and image synthesis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(5):6096–6110, 2022.
- [13] Eric Risser, Pierre Wilmot, and Connelly Barnes. Stable and controllable neural texture synthesis and style transfer using histogram losses. *arXiv preprint arXiv:1701.08893*, 2017.
- [14] Artsiom Sanakoyeu, Dmytro Kotovenko, Sabine Lang, and Bjorn Ommer. A style-aware content loss for real-time hd style transfer. In *proceedings of the European conference on computer vision (ECCV)*, pages 698–714, 2018.
- [15] Licheng Tang, Yiyang Cai, Jiaming Liu, Zhibin Hong, Mingming Gong, Minhu Fan, Junyu Han, Jingtuo Liu, Errui Ding, and Jingdong Wang. Few-shot font generation by learning fine-grained local styles. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7895–7904, 2022.
- [16] Zhanjie Zhang, Quanwei Zhang, Wei Xing, Guangyuan Li, Lei Zhao, Jiakai Sun, Zehua Lan, Junsheng Luan, Yiling Huang, and Huaizhong Lin. Artbank: Artistic style transfer with pre-trained diffusion model

and implicit style prompt bank. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pages 7396–7404, 2024.