

---

# The Supplementary Material for "Artistic Style Transfer with Internal-external Learning and Contrastive Learning"

---

Haibo Chen   Lei Zhao\*   Zhizhong Wang   Huiming Zhang  
Zhiwen Zuo   Ailin Li   Wei Xing\*   Dongming Lu

College of Computer Science and Technology, Zhejiang University  
{cshbchen, cszh1, endywon, qinglanwuji, zzwcs, liailin, wxing, ldm}@zju.edu.cn

## 1 Societal Impacts

**Positive impacts.** From our point of view, there may be three kinds of people that can benefit from this research: (i) Researchers who are interested in style transfer. Our research may inspire them to develop some more effective and remarkable methods in the future. (ii) Artists. They may draw some inspiration from the artistic images generated by our approach to augment their own work. (iii) Ordinary people who are interested in art. They can create stunning artistic images without any knowledge backgrounds about art.

**Negative impacts.** One potential negative impact is that the proposed style transfer method may replace some human jobs.

## 2 Involved Assets

Existing assets that we used in this work mainly include: 1) the codes of Gatys *et al.* [1], AdaIN [2], WCT [5], Avatar-Net [8], LST [4], and SANet [7], and 2) the MS-COCO dataset [6] and WikiArt dataset [3]. We report their URLs and licenses in the following,

- Gatys *et al.*: <https://github.com/anishathalye/neural-style>, GNU GENERAL PUBLIC LICENSE v3.0.
- AdaIN: <https://github.com/naoto0804/pytorch-AdaIN>, MIT License.
- WCT: <https://github.com/eridgd/WCT-TF>, MIT License.
- Avatar-Net: <https://github.com/LucasSheng/avatar-net>, we were unable to find its license.
- LST: <https://github.com/sunshineatnoon/LinearStyleTransfer>, BSD 2-Clause License.
- SANet: <https://github.com/GlebBrykin/SANET>, MIT License.
- MS-COCO: <https://cocodataset.org/#download>, we were unable to find its license.
- WikiArt: <https://www.kaggle.com/c/painter-by-numbers>, we were unable to find its license.

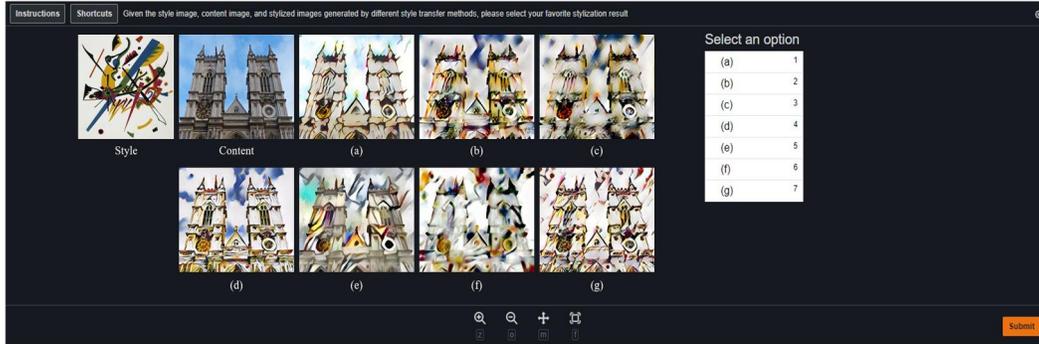
Note that MS-COCO and WikiArt have been widely used in a lot of existing works, and as far as we know, they do not contain personally identifiable information or offensive content.

---

\*Corresponding author

### 3 Research with Human Subjects

(i) In this work, we conducted user studies to evaluate the performance of our method. Section 4.3 in our main paper has presented the detailed text of instructions. Here, we provide the corresponding screenshots in Figure 1.



(a)



(b)

Figure 1: The screenshots of user study in terms of (a) preference score and (b) deception score.

(ii) As far as we know, our user studies do not involve any potential participant risks.

(iii) As for the wage, we paid each participant \$0.5 for completing an assignment within 40 minutes. We determined this wage mainly based on our research about other similar tasks that are also published on the Amazon Mechanical Turk (AMT) platform.

### 4 More Experimental Results

In this section, we present more experimental results: (i) We present more ablation study results in Figure 2 to demonstrate the effectiveness of the proposed internal-external learning and contrastive learning scheme. (ii) More qualitative comparisons between our method and the baselines are presented in Figure 3 and Figure 4.

### 5 Resources

During our experiments, 2 GeForce GTX 1080 Ti GPUs and 2 GeForce RTX 3090 GPUs are used.

### References

[1] 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.



Figure 2: Ablation studies of external learning (*abbr.* EL) and contrastive learning (*abbr.* CL) on image style transfer. It can be observed that our full model has the best performance. Please zoom in for a better view and details.

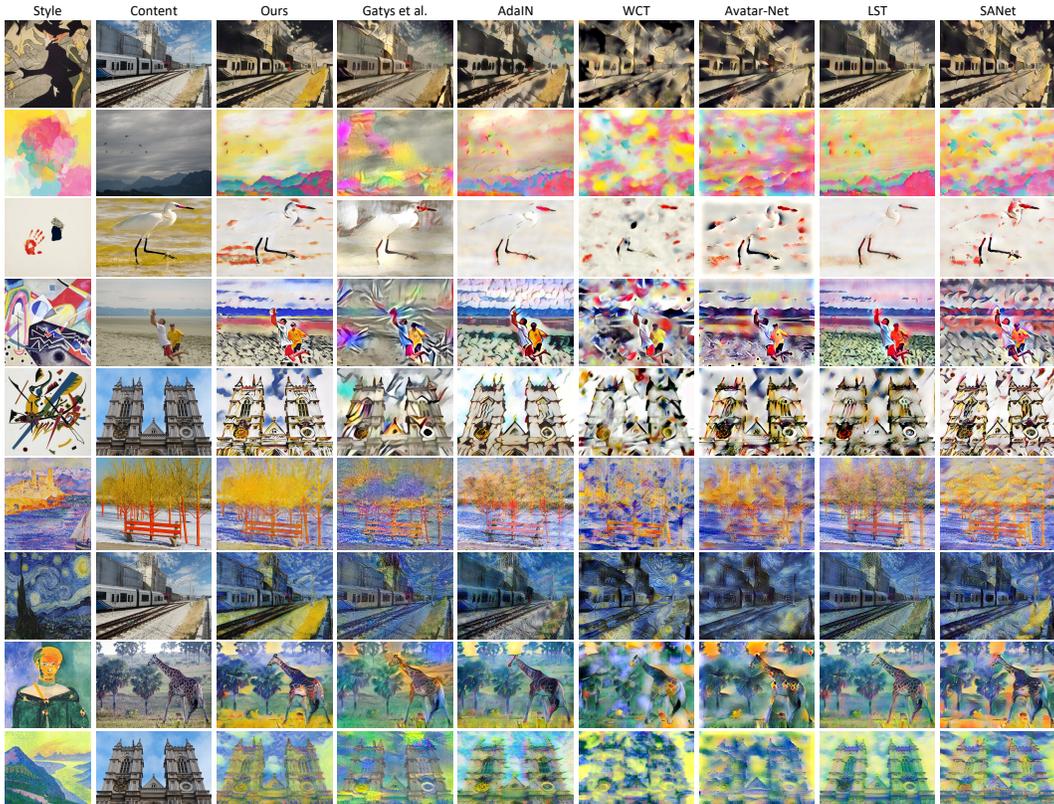


Figure 3: Stylization examples. The first and second columns show the style and content images, respectively. The rest of the rows show the stylization results generated with different methods.



Figure 4: Qualitative comparisons on video style transfer. The first row shows several video frames and the style image. The rest of the rows show the stylization results generated with different methods. The last column shows the heat maps of differences between different frames.

- [2] 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.
- [3] Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller. Recognizing image style. *arXiv preprint arXiv:1311.3715*, 2013.
- [4] Xueting Li, Sifei Liu, Jan Kautz, and Ming-Hsuan Yang. Learning linear transformations for fast image and video style transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3809–3817, 2019.
- [5] Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, Xin Lu, and Ming-Hsuan Yang. Universal style transfer via feature transforms. In *Advances in neural information processing systems*, pages 386–396, 2017.
- [6] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.
- [7] Dae Young Park and Kwang Hee Lee. Arbitrary style transfer with style-attentional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5880–5888, 2019.
- [8] Lu Sheng, Ziyi Lin, Jing Shao, and Xiaogang Wang. Avatar-net: Multi-scale zero-shot style transfer by feature decoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8242–8250, 2018.

## Checklist

1. For all authors...
  - (a) Do the main claims made in the abstract and introduction accurately reflect the paper’s contributions and scope? [\[Yes\]](#)
  - (b) Did you describe the limitations of your work? [\[Yes\]](#)
  - (c) Did you discuss any potential negative societal impacts of your work? [\[Yes\]](#) Please refer to our supplementary material.
  - (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [\[Yes\]](#)
2. If you are including theoretical results...

- (a) Did you state the full set of assumptions of all theoretical results? [Yes]
  - (b) Did you include complete proofs of all theoretical results? [Yes]
3. If you ran experiments...
- (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes]
  - (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes]
  - (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No] Because it would be too computationally expensive and its significance is very limited for the task of style transfer.
  - (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] Please refer to our supplementary material.
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
- (a) If your work uses existing assets, did you cite the creators? [Yes] Moreover, we report their URLs in our supplementary material.
  - (b) Did you mention the license of the assets? [Yes] Please refer to our supplementary material.
  - (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
  - (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
  - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes] Please refer to our supplementary material.
5. If you used crowdsourcing or conducted research with human subjects...
- (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [Yes]
  - (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [Yes] Please refer to our supplementary material.
  - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [Yes] Please refer to our supplementary material.