Inversion-Based Style Transfer with Diffusion Models – Supplementary Materials

Yuxin Zhang[1,2] Nisha Huang[1,2] Fan Tang[3] Haibin Huang[4]
Chongyang Ma[4] Weiming Dong[1,2*] Changsheng Xu[1,2]
[1]MAIS, Institute of Automation, Chinese Academy of Sciences [2]School of AI, UCAS
[3]Institute of Computing Technology, Chinese Academy of Sciences [4]Kuaishou Technology

Appendix

In this appendix, we provide additional details and analysis of our approach. We give our user study results and more explanation in Section 1. Social impacts are provided in Section 2. Finally, we present additional results in 3.

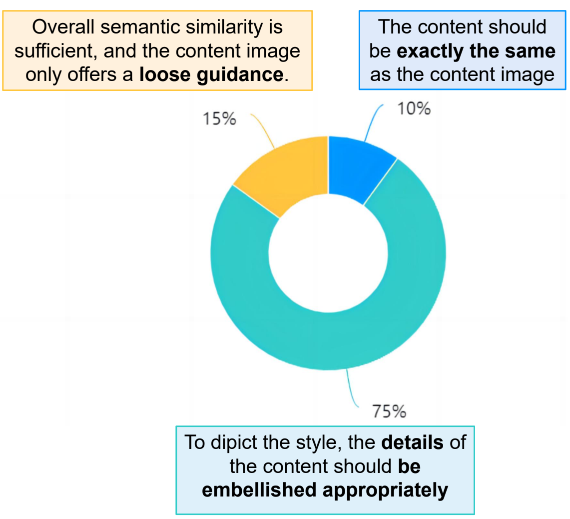

Figure 1. Preference of the content guidance strength.

1. User Study

We conducted a survey of 60 participants on the preferences of the the content image guidance strength and artistic visual effects. In the case of a content image existing, users tend to consider that "To depict the artistic style, the details of the content should be embellished appropriately". We then invite the participants to rank the factors of their expected visual effect. The average comprehensive score of the options in the sorting question is automatically calculated based on the ranking of the options by all the participants. The higher the score, the higher the comprehensive ranking. The scoring rule is formulated as:

$$score = \frac{(\sum frequency \times weight)}{participantes}, \tag{1}$$

where $score$ denotes the average comprehensive score of the options, $participantes$ denotes the number of people who complete this question, $frequency$ denotes the frequency that the option is selected by users, $weight$ denotes the weight which is determined by the option's ranking. The ranking results (rank by score from highest to lowest):

- Similar artistic effect on semantic corresponding subjects (s=5.4);
- With the same paint material ($score$=3.65);
- Having similar brushstrokes ($score$=3.2);
- Having typical shapes ($score$=2.65);
- With the same decorative elements ($score$=2.1);
- Sharing the same color ($score$=1.4).

More details are shown in Figures 1 and 2.

2. Social Impacts

Our approach tackle the problem that the artistic visual effect is difficult to describe in the artistic image generation. These can lower the barrier for the public to create artwork, and help professionals increase efficiency and inspire inspiration. We hope our method can help more people get in touch with art appreciation and creation. On the other side, we recommend people always use this technology in the right way, not to create artworks with offensive contents.

3. More Image-to-Image Generation Results

We show image-to-image generation results in Figures 3-5. Several classic paintings of famous painters are tested to demonstrate that our method can be competent for a variety of art forms from classical to modern.

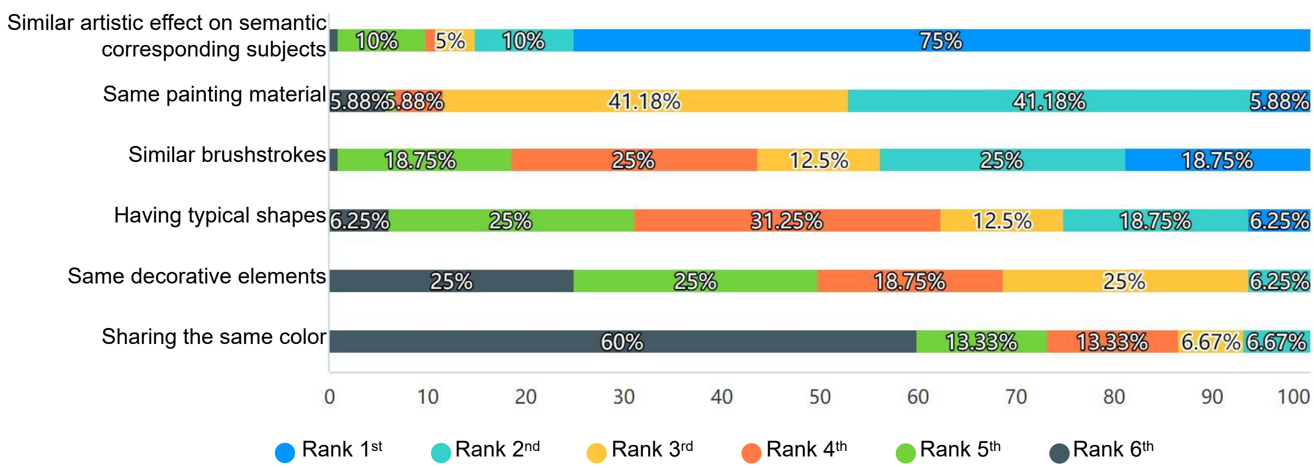

Figure 2. Ranking results of visual factor importance.

Figure 3. Additional image-to-image style transfer results of our method. The content of each image depends on the input image and the guided text "[C]" of the corresponding painting image shown in the left column.

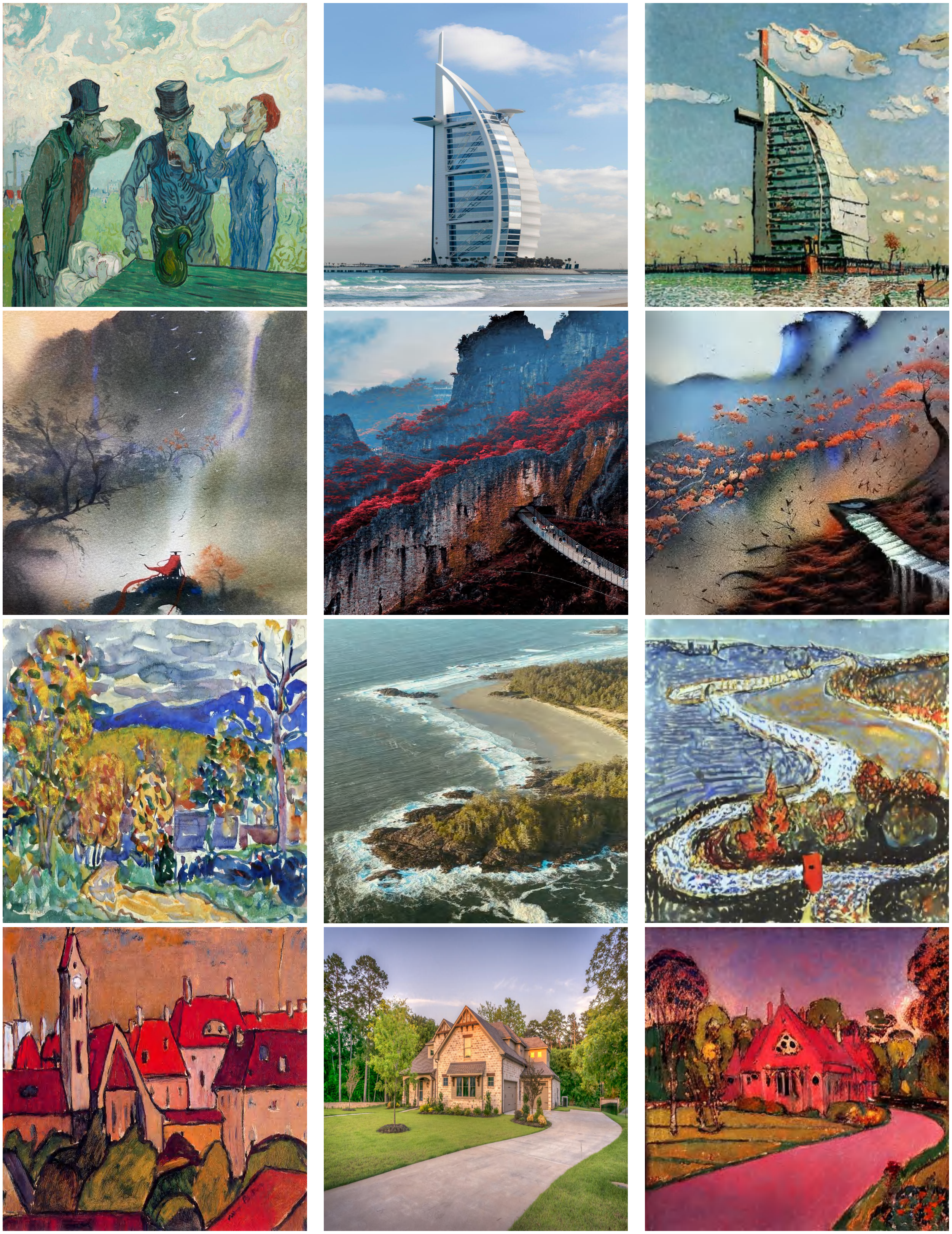

Figure 4. Additional image-to-image style transfer results of our method. The content of each image depends on the input image and the guided text "[C]" of the corresponding painting image shown in the left column.

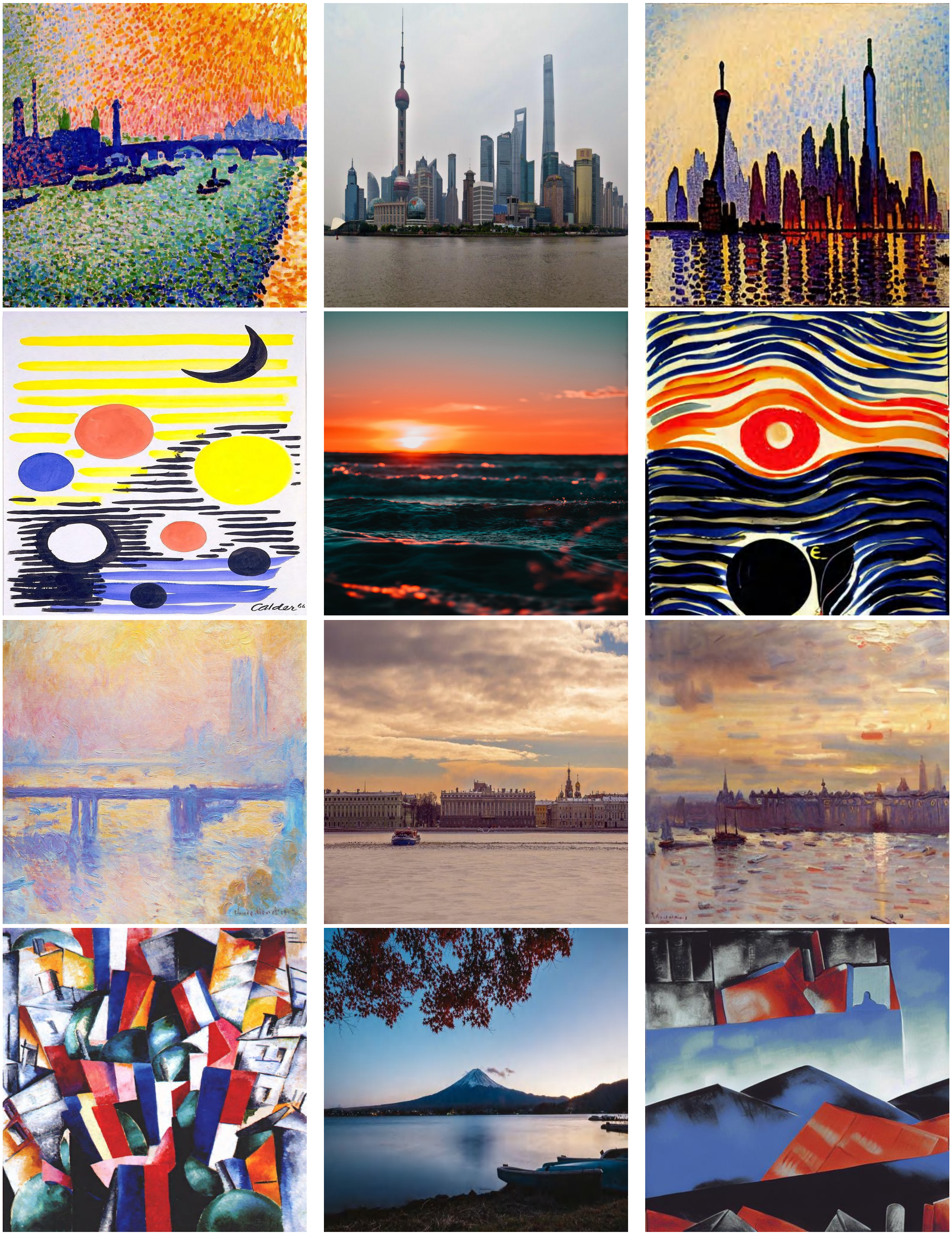

Figure 5. Additional image-to-image style transfer results of our method. The content of each image depends on the input image and the guided text "[C]" of the corresponding painting image shown in the left column.