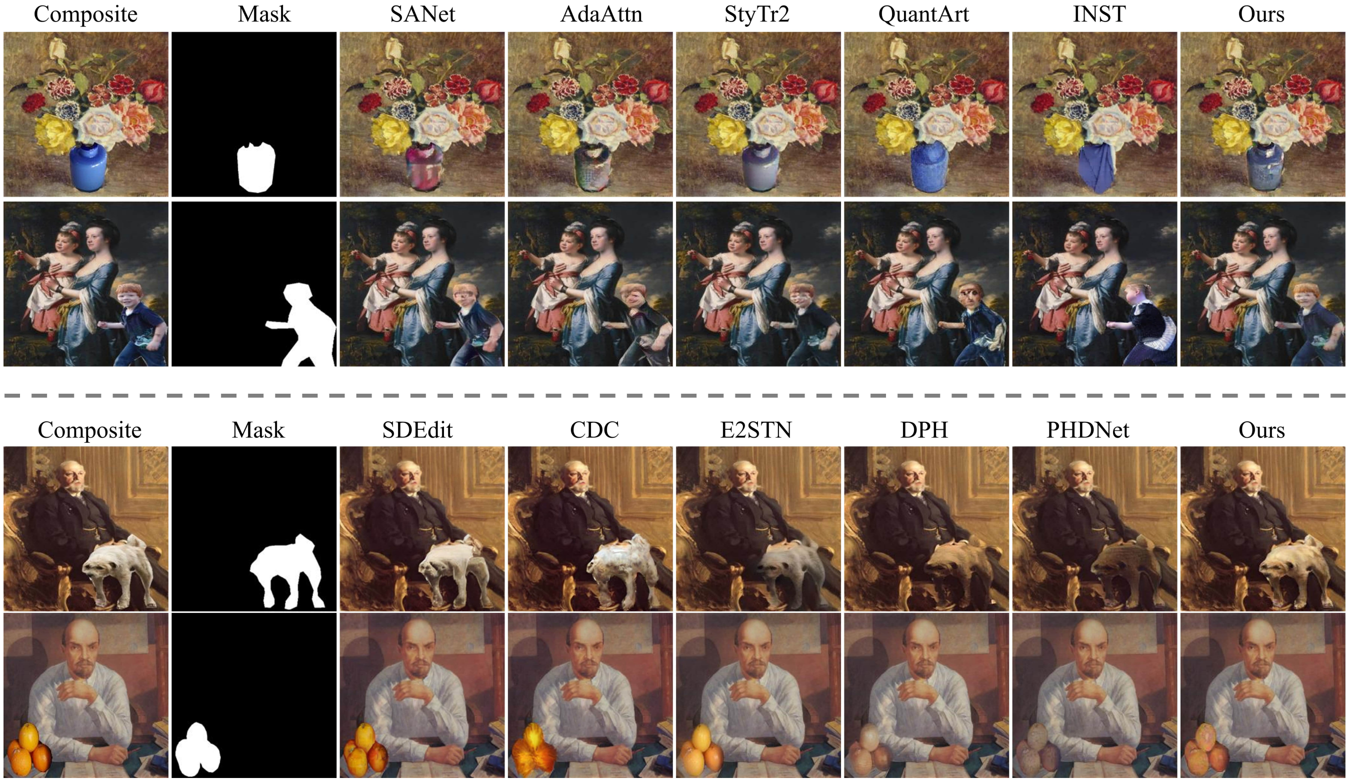

Figure 4: In the upper part, we compare with style transfer baselines SANet (**?**), AdaAttN (**?**), StyTr2 (**?**), QuantArt (**?**), INST (**?**). In the lower part, we compare with painterly image harmonization baselines SDEdit (**?**), CDC (**?**), E2STN (**?**), DPH (**?**), PHDNet (**?**).

port the inference time and FLOPs. We evaluate with the image size 256×256 and the inference time is averaged over 100 test images on a single RTX 3090 GPU. The optimization-based method DPH is time-consuming, due to iteratively updating the input composite image. Diffusion-based methods are much slower than the other feed-forward methods, which limits their real-world applicability. Our method is relatively efficient compared with the competitive baselines (**??**).

4.3 Hallucinated Object Style

To intuitively demonstrate the effectiveness of our hallucinated object styles, we sample some photographic objects in the test set and find their reference painterly objects in the test set, similar to the process of constructing training image pairs (see Section 3.1). Specifically, given a photographic object and a painterly object in a painterly image, we can obtain a composite image as illustrated in Figure 3, followed by calculating the L_2 distance between the composite object feature $\hat{\boldsymbol{f}}^{c,o}$ and painterly object feature $\hat{\boldsymbol{f}}^{p,o}$. For each photographic object, we retrieve its nearest painterly object as reference object and the corresponding painterly image as reference image. In Figure 5, we show several examples of photographic objects and the retrieved reference objects. It can be seen that the photographic objects have similar color and semantics with their reference objects, which verifies the effectiveness of our learnt domain-invariant object features.

Based on the composite image composed by the photographic object and its reference image, we further adjust the shape of photographic object to fully cover the reference object in the reference image, to exclude the influence of reference object. Then, based on our trained model, we apply three types of style vectors to the composite object to get the harmonized images. The first one is the background style vector $\boldsymbol{s}_l^{c,b}$. The second one is the style vector $\boldsymbol{s}_l^{p,o}$ of reference object. The third one is our hallucinated style vector $\tilde{\boldsymbol{s}}_l^{c,o}$. We apply the above three style vectors to the foreground feature map using AdaIN operation. The obtained harmonized images are denoted as “BG”, “RO”, and “Ours” respectively, as shown in Figure 5. The harmonized results “BG” and “RO” differ a lot, which demonstrates the huge gap between background style and object style. Directly applying background style may destroy the original color of composite object or bring in noticeable artifacts.

Since the composite object and reference object have similar object information and background style, the target style $\tilde{\boldsymbol{s}}_l^{c,o}$ of composite object is expected to approach the reference object style $\boldsymbol{s}_l^{p,o}$. Based on Figure 5, the harmonized results “Ours” are close to “RO”, which proves that our model can hallucinate ideal target styles.

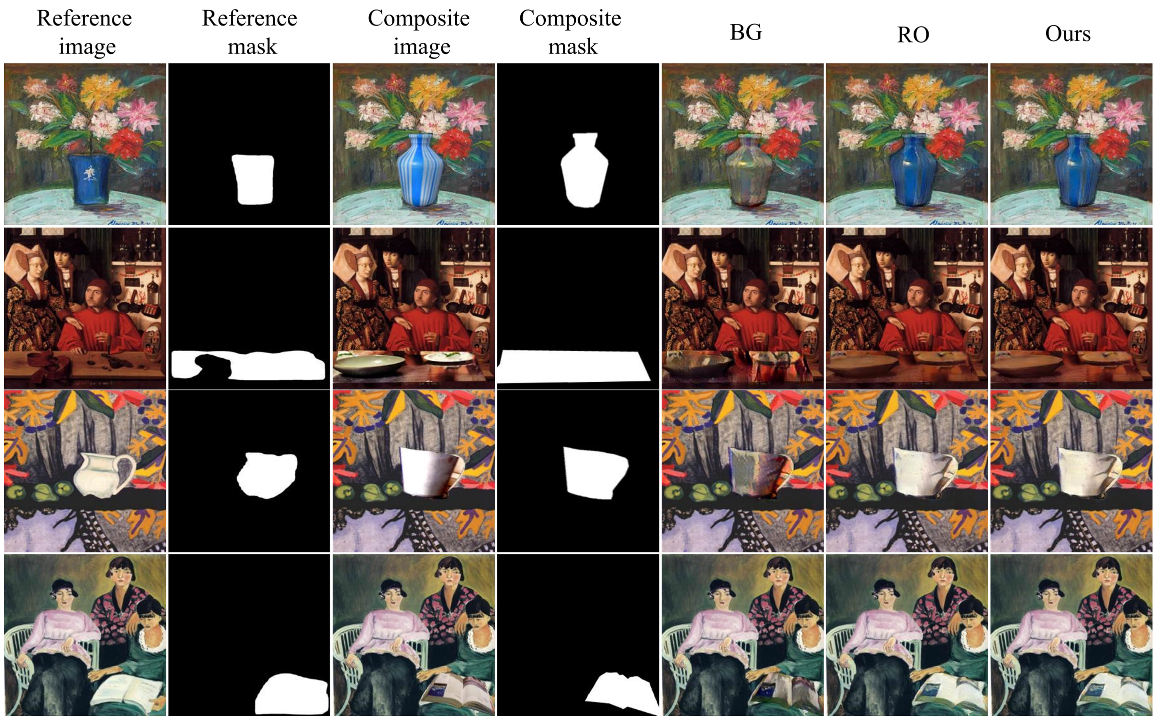

Figure 5: From left to right, we show the reference image, the mask of reference object, the composite image, the mask of composite object, and the harmonized results obtained using different style vectors. “BG” uses background style vector, “RO” uses reference object style vector, and “Ours” uses our hallucinated style vector.

4.4 More Results in Supplementary

In the supplementary, we will provide ablation study results, more visual comparison with baselines, and discussion on failure cases. We will also show the results beyond COCO (**?**) dataset to demonstrate the generalization ability of our method across different datasets and different object categories.

5 Conclusion

In this work, we have explored learning from painterly objects for painterly image harmonization. Based on the annotated pairs of composite images and reference painterly images, we have succeeded in hallucinating the target style of composite object, leading to visually pleasing harmonization results. Extensive experiments on the benchmark dataset have proved the advantage of our proposed ArtoPIH.

Acknowledgments

The work was supported by the National Natural Science Foundation of China (Grant No. 62076162), the Shanghai Municipal Science and Technology Major/Key Project, China (Grant No. 2021SHZDZX0102, Grant No. 20511100300). Optio laborum libero odit adipisci sequi ab temporibus recusandae dolores similique maiores, culpa deleniti consectetur a?Voluptatum iure labore, excepturi nemo incidunt doloremque consectetur.Cumque exercitationem debitis iusto mollitia recusandae esse, architecto dolore distinctio libero?Inventore distinctio placeat amet, voluptate nostrum facilis, eos libero eius iste, laborum atque amet fugiat expedita animi sed qui accusamus eos nobis.Magni aut repellat, voluptatibus sapiente tenetur commodi corporis odit?Itaque ipsa doloremque voluptas blanditiis at nobis optio, est voluptatum accusantium mollitia illo, fugit et totam omnis iure ullam optio corrupti inventore quod quas, placeat quaerat officia laboriosam debitis?Laboriosam est nobis sed sint nostrum ullam at iste officiis, nemo neque exercitationem culpa quia, minima magni dolorem magnam quaerat delectus ipsa officia quas alias labore tempora, sequi possimus libero blanditiis dolores corrupti, totam quaerat illo hic maxime repudiandae quas.Eveniet ipsum ea nobis numquam, quo fugit sequi harum.Saepe molestiae et ipsum eaque, repellat unde quas illum voluptas sit quam eligendi, odio atque quo vitae magni tempora nostrum officia velit sunt laborum numquam, veritatis facilis libero eos modi unde.Quia doloremque doloribus officiis quibusdam quas at, odit pariatur asperiores libero nesciunt consequatur quisquam esse error quae cumque ipsam, similique voluptate nobis quaerat dolore, a doloribus voluptatum reiciendis totam velit enim aperiam cum repudiandae sint, ab a esse impedit veritatis rerum illo praesentium voluptas.Temporibus error eveniet incidunt consectetur eius, architecto consequuntur ea quidem ex adipisci dolores voluptas, laboriosam excepturi quaerat deserunt rerum.