Intrinsic Image Transfer for Illumination Manipulation: Supplementary Materials

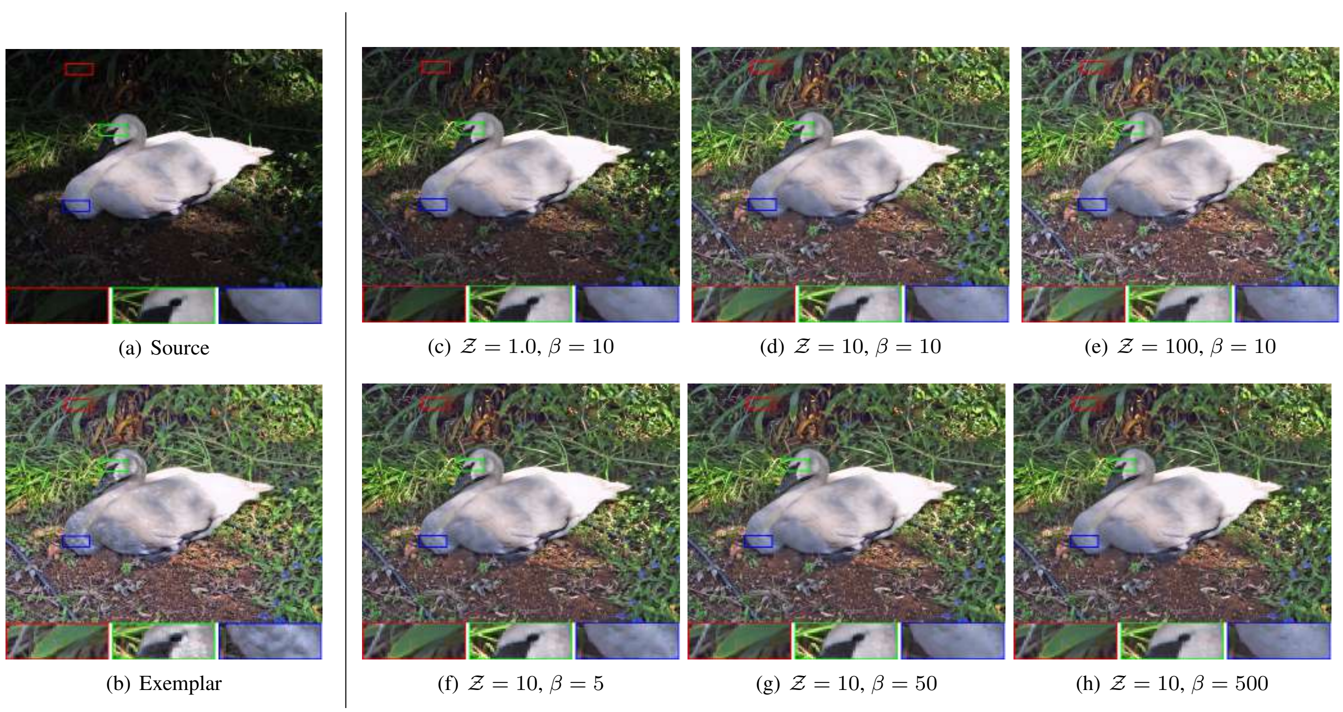
Junqing Huang, Michael Ruzhansky, Qianying Zhang, Haihui Wang[1]
zcfeng@buaa.edu.cn, ruzh@ic.ac.uk, zhang_qy@sz.jnu.edu.cn, whhmath@buaa.edu.cn

1 Derivation of model parameters

We additionally take an ablation study to explore the choice of parameters in Eq. 2 of the main paper. We empirically found that the parameters, including $(\mathcal{N}_i, \delta_s, \delta_r)$ in filter $\mathcal{K}$ and (Ω_i, ϵ) in LLE weights $\boldsymbol{W}$, can be configured or fixed in views of their local property. The global parameters α, β, γ need to be configured to weigh the three photo-realistic losses. As explained in the main paper, the parameters α and γ control the image illumination. If $\alpha \gg \gamma$, the output image illumination is close to the exemplar, otherwise it is close to the source. For simplicity, we introduce $\mathcal{Z} = \frac{\alpha}{\gamma}$ to weigh the importance of the exemplar. We always keep $\alpha + \gamma = 1$ to limit the output results between source and exemplar. As shown in Fig. 1 (Top), a large $\mathcal{Z}$ gives arise a brighter image having a closer illumination to that of the exemplar. On the other hand, β is related to the textural distortions of the exemplar, which can be fixed in a known level of textural distortion. As shown in Figure 1 (Bottom), a large β forces our IIT method to generate images with more consistent local textures. Note that a very large β may cause over-smoothing textures. We set $\beta \leq \beta_{max} = 1000$ in all our experiments.

2 Additional Experimental Results

For completeness, we provide additional experimental results. We compare our IIT method with the state-of-art in each dataset. All parameter-settings are the main paper without specification. In Fig. 2, we show the results with the exemplar given by the NASA Retinex method [2]. It is notable that the NASA Retinex method is capable of producing high-quality results with vivid



(a) Source (c) $\mathcal{Z} = 1.0, \beta = 10$ (d) $\mathcal{Z} = 10, \beta = 10$ (e) $\mathcal{Z} = 100, \beta = 10$

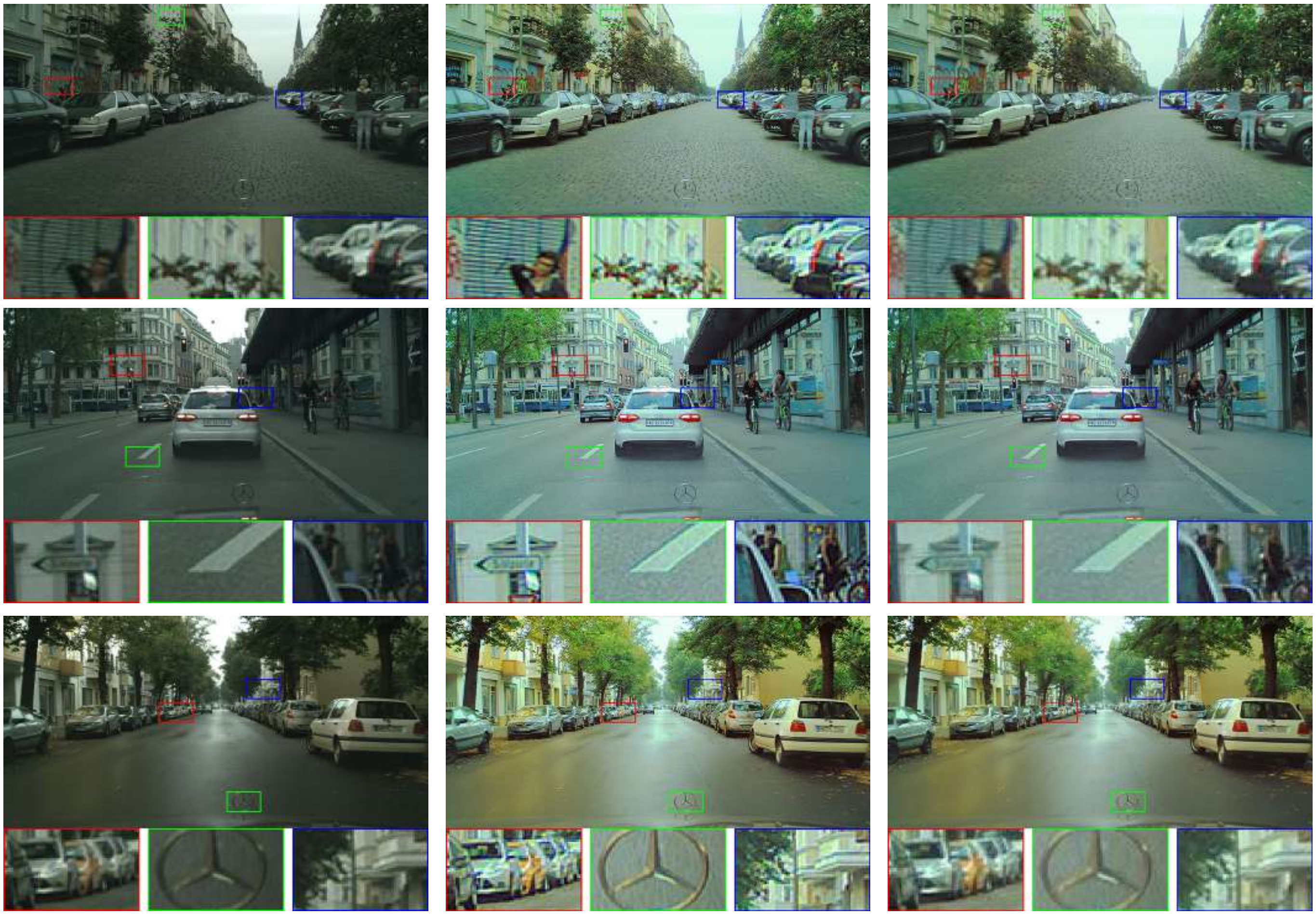
(b) Exemplar (f) $\mathcal{Z} = 10, \beta = 5$ (g) $\mathcal{Z} = 10, \beta = 50$ (h) $\mathcal{Z} = 10, \beta = 500$

Fig. 1. Visual results of our IIT algorithm by using exemplars with different levels of textural distortions. (a) Source image, (b) CLAHE exemplar [1], and (c) $\sim$ (e) the corresponding results.

color and contrast, but the achievement of fine-balanced image illumination is at expense of exaggerating noise. In this case, we set $\alpha = 0.99$, $\beta = 100$ and $\gamma = 0.01$ for the proposed IIT method in consideration of the fine-balanced illumination of the exemplar. As we can see, the output results reveal identical brightness and vivid color as that of the exemplar, and the strong noise in which are also remarkably suppressed, giving high-quality consistent textures as that of the input image. In Fig. 3, we show the results on Cityscape dataset [4], where the low-resolution images exhibit obvious degradation in color, saturation and contrast due to the unappropriated illumination conditions. As mentioned in the main paper, the deep learning-based WESPE photo enhancer [3] can produce visual-pleasant results outperforming a wide range of traditional photo methods and commercial software. However, some textural distortions may occur around the main structures and salient edges due to the amplification of high-frequency details during the WESPE training process. In this case, we show that the proposed IIT method not only preserves the local textures consistent with the source but removes these non-consistent structures and high-frequency artifacts. In Fig.4, we show the results on DPED dataset [5], where the selected images reveal much better quality with higher resolution and the degradation of image illumination is not so serious as that of Cityscape dataset [4]. As the result, we here adopt using the CLAHE exemplar and compare the results with the WESPE method [3]. As we can see, the WESPE method [3] may lose the tiny details in the bright regions due to the over-espouse problem. While, our IIT method faithfully preserves these details and gives natural-like results.

(a) Input (b) NASA Retinex [2] (c) Ours (IIT + BF)

Fig. 2. Visual comparison on NASA dataset. (a) Source; (b) NASA Retinex [2]; (c) Our IIT results. The local noises are significantly reduced in comparison of the NASA Retinex (b).



(a) Input (b) WESPE [3] (c) Ours (IIT+BF)

Fig. 3. Visual comparison on Cityscapes dataset [4]. (a) Source; (b)WESPE [3]; (c) Our results. Notice that the state-of-art WESPE reveals strong color artifacts around the salient edges, which can be suppressed using our IIT scheme.

Finally, we show our IIT method extended to image stylization with the same configuration as explained in the main paper. The stylized exemplars are provided by the deep learning method [6], which is capable of producing favorable style results under the reference styles. As shown in Fig. 5 and 6, we take two typical image-to-image translation cases into account. In the first case, the natural images are transferred for the photo-realistic results for photo-realistic results; while the image-to-image translation in the second case attempts to keep the strokes and textures in artistic images invariant. In both cases, the proposed IIT method helps to generate the stylized results with high-quality consistent textures as the input image despite the strong non-consistent textures in exemplars. These results, again, demonstrate the performance and robustness of the proposed IIT method to the exemplar and imply that it is not necessary to pay much attention to the local details when applying the exemplar in our IIT framework.

We additionally show the performance of our IIT algorithm on the endoscope image enhancement, in which illumination manipulation is always a challenging problem due to the occlusion or unsuitable exposure compensation. We take the CLAHE output as exemplars in this case. As shown in Fig. 7, these endoscopic images can be recovered with high-fidelity structures which are very important for nerves, blood vessels, and tissues in the medical image processing field.

References

[1] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics gems*, pp. 474–485, 1994.

[2] "Nasa retinex," https://dragon.larc.nasa.gov/retinex/.

[3] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, "Wespe: weakly supervised photo enhancer for digital cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 691–700.

[4] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

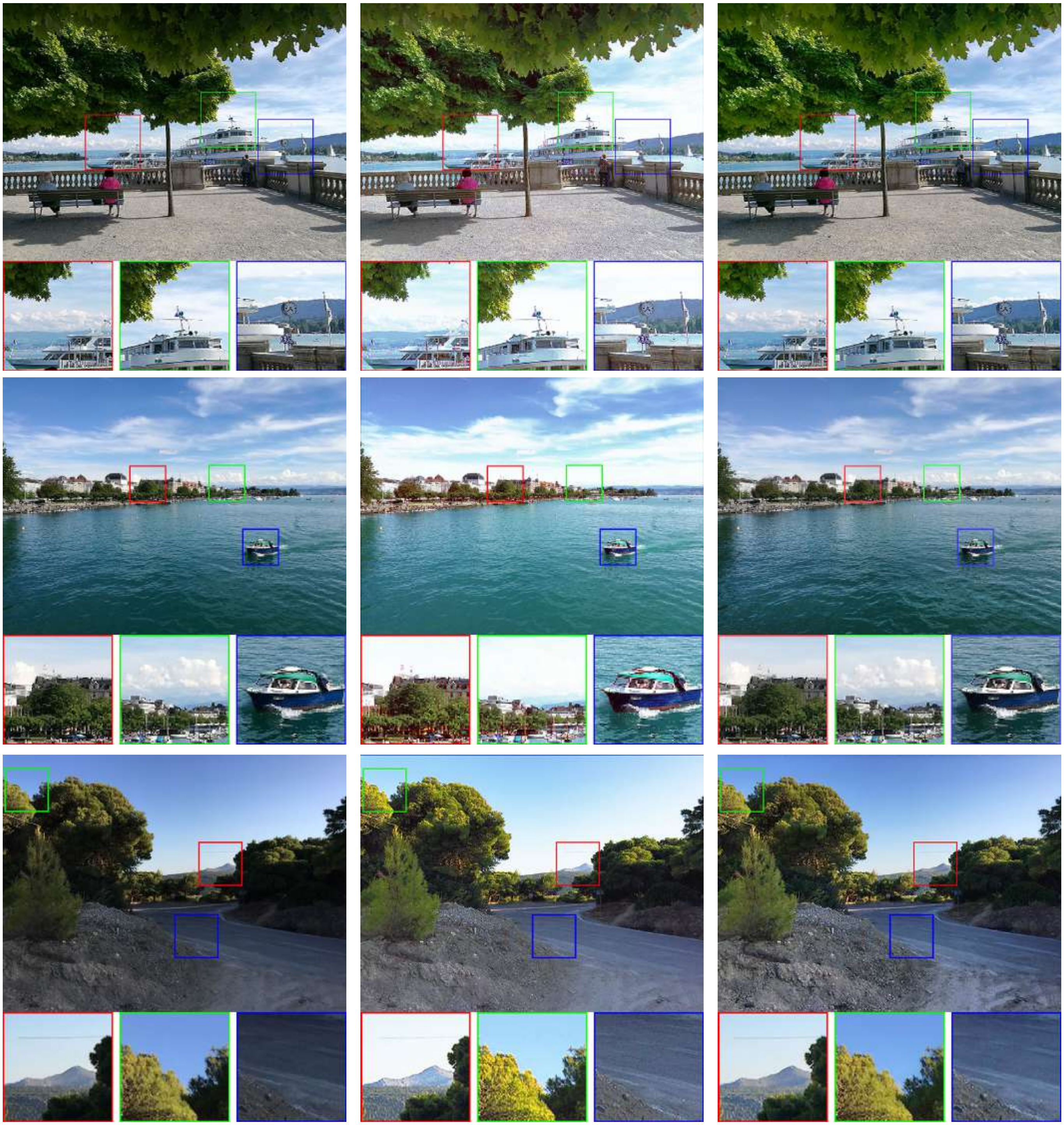
[5] A. Ignatov, N. Kobyshev, K. Vanhoey, R. Timofte, and L. Van Gool, "Dslr-quality photos on mobile devices with deep convolutional networks," in *Proceedings of the International Conference on Computer Vision (ICCV)*, 2017.

[6] X. Huang and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 1501–1510.

[7] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2414–2423.

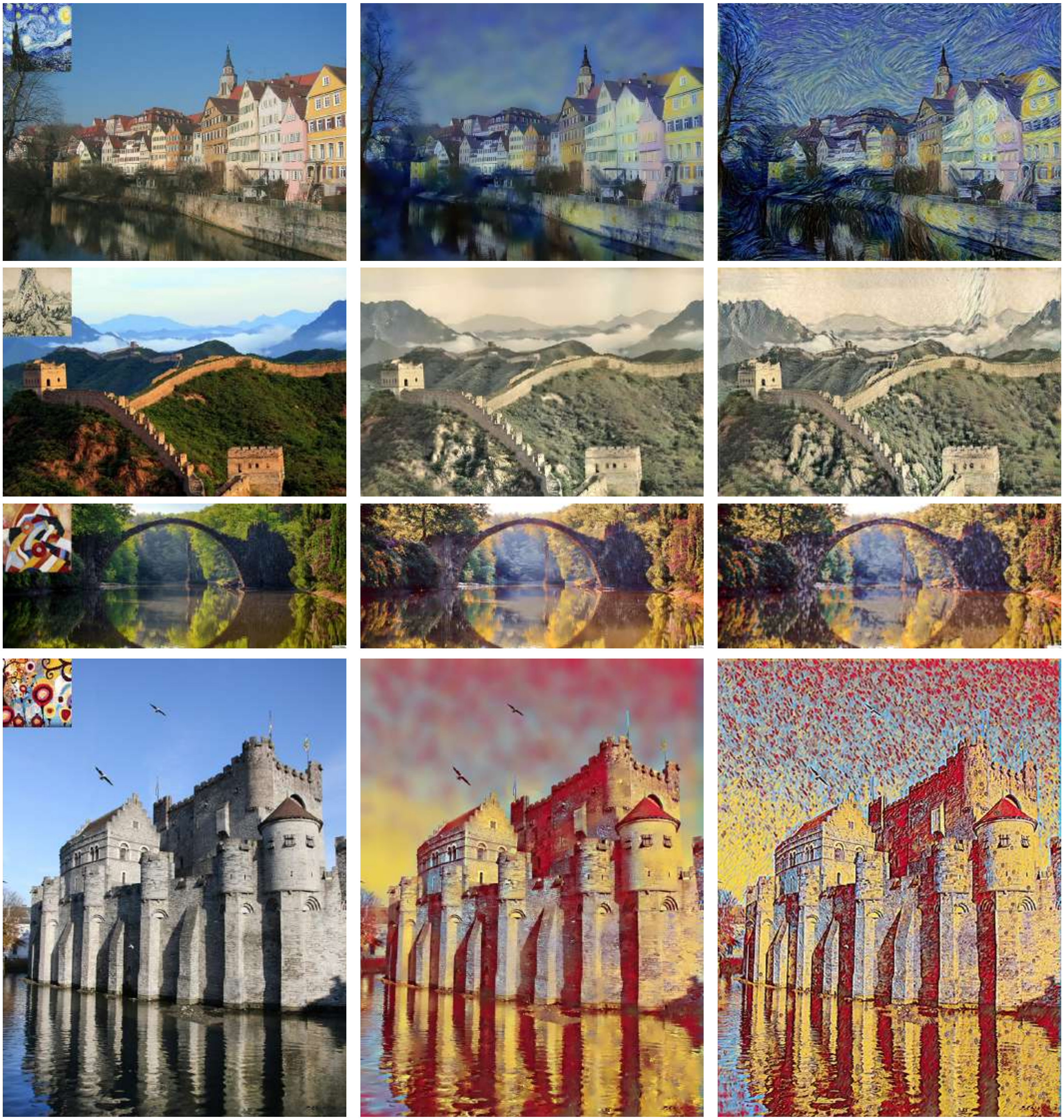
[8] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, and M. Song, "Neural style transfer: A review," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 11, pp. 3365–3385, 2019.

[9] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M.-H. Yang, "Universal style transfer via feature transforms," *Advances in neural information processing systems*, vol. 30, 2017.



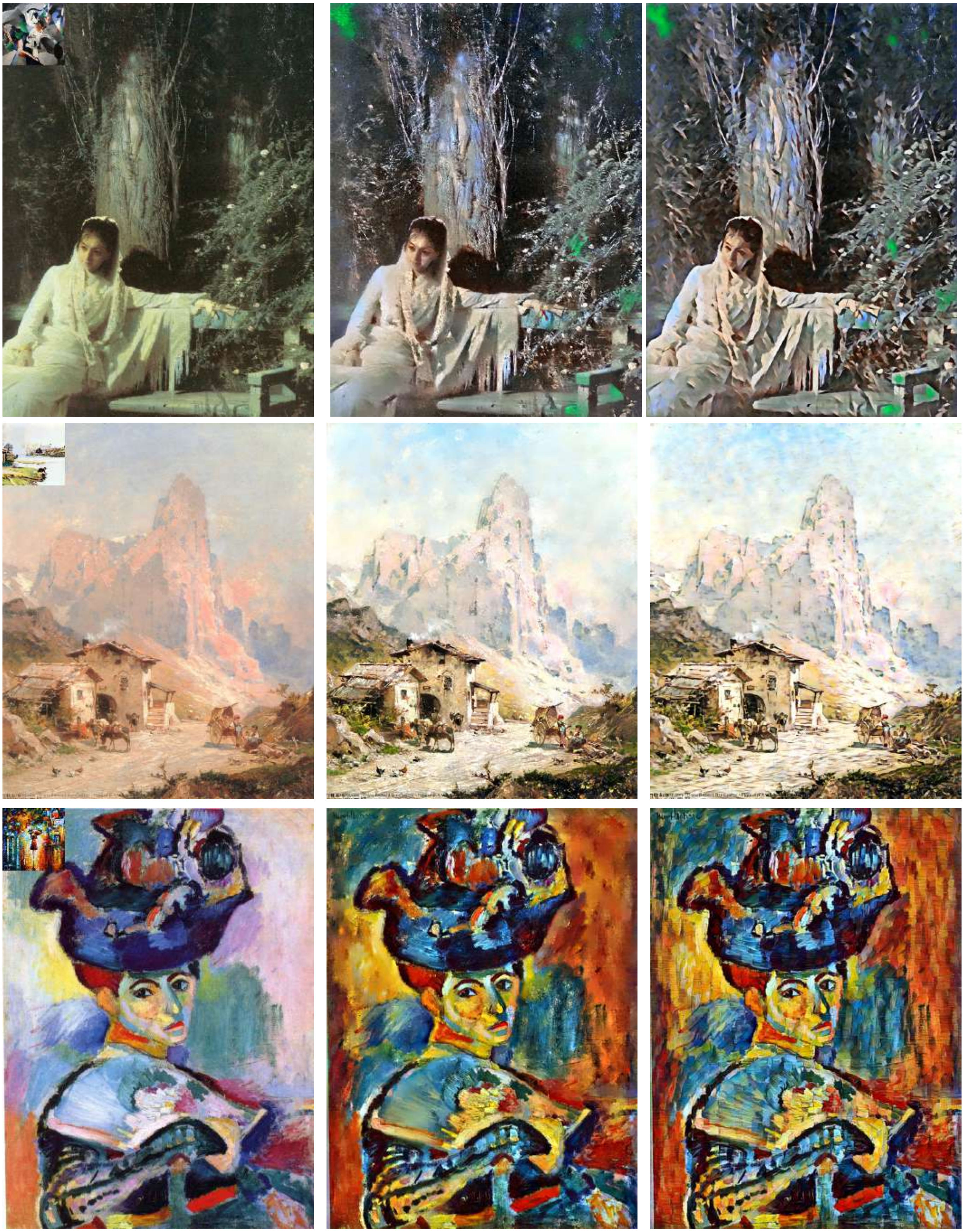
(a) Input (b) WESPE [3] (c) Ours (IIT+BF)

Fig. 4. Visual comparison on DPED dataset [5] . (a) Source; (b)WESPE [3]; (c) Our results. Notice that the state-of-art WESPE method gives an over-exposure in local, while our IIT method can preserve these details based on the CLAHE exemplars.



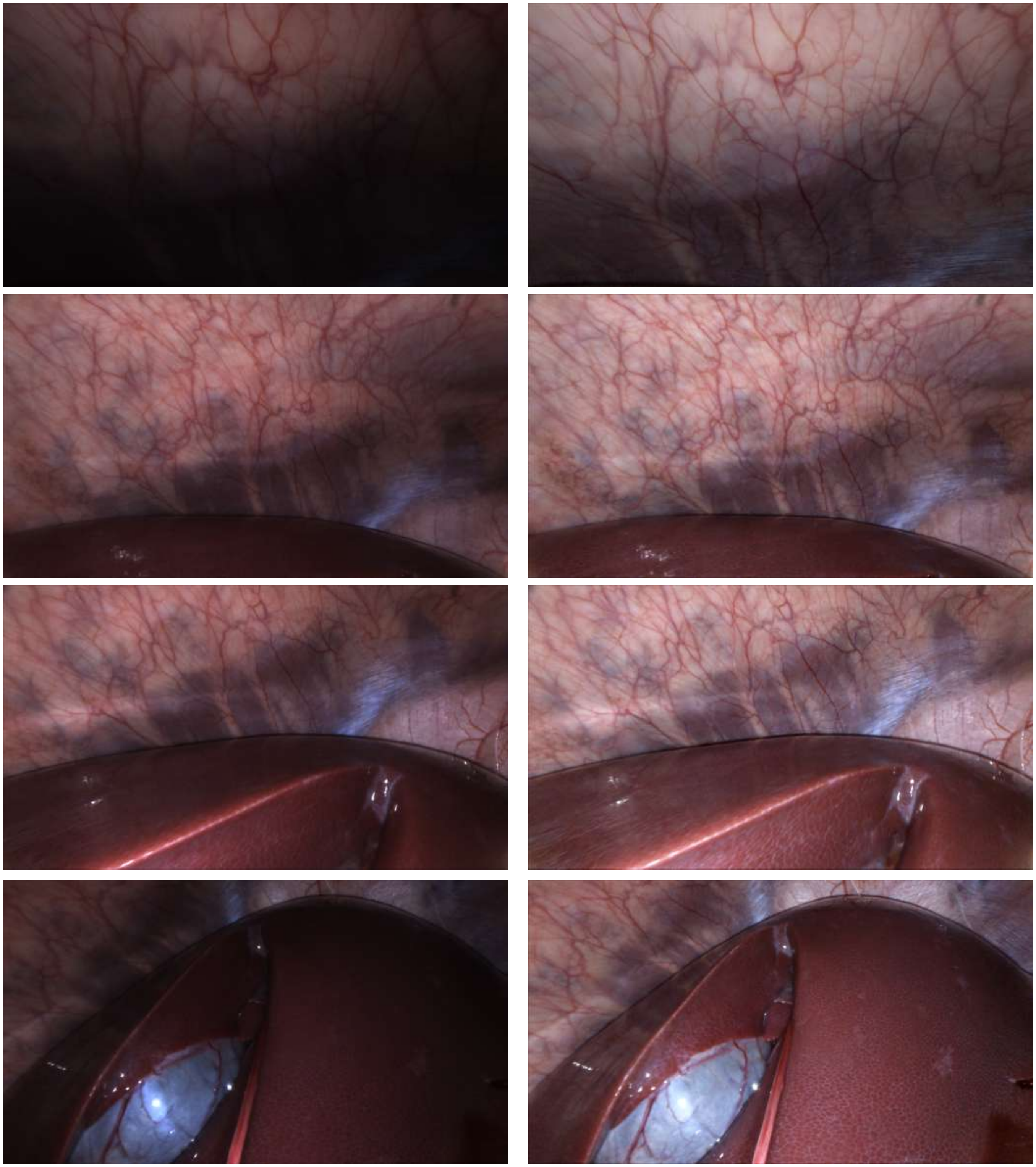
(a) Input (b) Ours (IIT + GF) (c) Exemplar

Fig. 5. Photorealistic style transfer. Given an input image (a) and a reference style (up-left), deep-learning methods [6], [7], [8], [9] produce the stylized results (c), which are then used as the exemplars for our IIT model to generate the photorealistic result (b).



(a) Input (b) Ours (IIT + GF) (c) Exemplar

Fig. 6. Photorealistic style transfer. Given an input image (a) and a reference style (up-left), deep-learning method [6] produces the stylized results (c), which are then used as the exemplars for our IIT model to generate the photorealistic result (b).



(a) Input (b) Results (IIT+BF)

Fig. 7. Visual results on endoscope image enhancement. (a) Source and (b) Our results. We use the CLAHE exemplars in this case.