TCTL-Net: Template-free Color Transfer Learning for Self-Attention Driven Underwater Image Enhancement (Supplementary Material)

Kunqian Li, *Member, IEEE*, Hongtao Fan, Qi Qi*, Chi Yan, Kun Sun, *Member, IEEE*, and Q. M. Jonathan Wu, *Senior Member, IEEE*

***Abstract*—This supplementary material presents more visual comparisons of the enhancements made on the UIEB, RUIE, EUVP, and SQUID datasets. Additionally, we provide further discussion and analysis based on the performance of each method across all test datasets, focusing more on the overall effectiveness of the enhancement methods on various underwater image datasets. Finally, we also present additional results of the proposed method for other types of image enhancement tasks.**

I. More Visual Comparison of Underwater Image Enhancement

In this supplementary material, we present additional visual comparisons of the enhancements achieved through different approaches. The results on the **UIEB** (including the **UIEB Challenging set**) [2], **RUIE** [8], **EUVP** [10], and the **SQUID** [11] datasets are presented in Figures 1, 2, 3, and 4, respectively. The diverse scenes and degradation types of underwater images in the UIEB and EUVP datasets, as shown in Figures 1 and 3, allow for a comprehensive evaluation of the performance of underwater enhancement methods, including color correction capability, detail recovery, and robustness. The RUIE and SQUID datasets contain more cases with severe visual degradation, which can be used to verify the applicability of each enhancement method to extreme examples.

By evaluating their performance on various underwater image datasets collectively, we present the following additional comments and discussion regarding the tested methods' performance:

(1) Among the traditional methods we tested, the method based on prior knowledge of the imaging model (i.e., ULAP) struggles to handle complex and diverse underwater visual degradation and experiences more failures compared to other methods.

(2) In contrast, traditional methods based on heuristic priors such as color balance (e.g., CBF [5]), contrast improvement, and the gray-world hypothesis (e.g., TwoStep [3]) outperform the method based on a single imaging model prior and perform well in some severely degraded examples. It also suggests that such generalized assumptions are applicable in partially realistic underwater imaging environments.

(3) However, the constraints imposed by the aforementioned heuristic priors are sometimes too strict, resulting in obvious overcorrection and color distortion from time to time. Furthermore, while using the gray-world hypothesis to eliminate color distortion, the CBF method inevitably sacrifices the diversity and vibrancy of the original colors.

(4) Deep-learning methods have demonstrated superior robustness compared to the tested traditional methods, but they have also encountered failures when dealing with images that exhibit extreme degradation (e.g., the SQUID dataset). As discussed in the main text, deep-learning methods perform more robustly than traditional methods because volunteer voting screens out enhancements that do not match visual preferences while acquiring reference images for learning. It effectively prevents deep-learning methods from introducing obvious over-correction and visual flaws.

(5) The cases that lead to the failure of existing deep-learning methods mostly exhibit extreme color distortions, such as a heavy greenish appearance, which are less frequently observed in the training set. With only a few samples, it is difficult for deep-learning methods to learn complicated distortion-to-clear mappings. While with the same training set, it is much easier to learn to predict color transfer parameters. As a result, color transfer-based enhancement methods such as LCL-Net and TCTL-Net perform better on challenging examples.

(6) Comparing LCL-Net with TCTL-Net, we find that the results of the latter introduce much less new color distortion and over-correction. This is due to the fact that the learning of transfer parameters provides more flexibility for the color transfer-based enhancement strategy, reduces the dependence on the pre-prepared template pooling, and avoids the negative impact of using inappropriate templates when there is no preferred template. In particular, compared to the template-based color transfer of LCL-Net, TCTL-Net introduces the self-attention mechanism

Kunqian Li, Hongtao Fan and Chi Yan are with the College of Engineering, Ocean University of China, Qingdao 266404, China (likunqian@ouc.edu.cn; fht@stu.ouc.edu.cn; yanchi@stu.ouc.edu.cn).

Qi Qi is with the School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266520, China (qiqi@qut.edu.cn).

Kun Sun is with the School of Computer Science, China University of Geosciences, Wuhan 430078, China (sunkun@cug.edu.cn).

Q. M. Jonathan Wu is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada. (jwu@uwindsor.ca).

The research has been supported by the National Natural Science Foundation of China under Grant 62371431, 61906177, 62176242, in part by the Natural Science Foundation of Shandong Province under Grant ZR2019BF034, and in part by the Fundamental Research Funds for the Central Universities under Grants 202262004.

* Corresponding author: Qi Qi (qiqi@qut.edu.cn)

and predicts biased transfer parameter matrices, providing differentiated visual enhancement. This effectively avoids local overexposure and excessive color correction.

(7) For multiple images collected in similar scenes, TCTL-Net can often generate more stable and consistent enhancement results compared to LCL-Net. The results demonstrate that, in comparison to the transfer parameters acquisition method based on template selection, the transfer parameters obtained through parameter prediction learning exhibit greater robustness.

In Fig. 5, we present more visual comparison of the enhancements on images that exhibit obvious uneven degradation. By comparing TCTL-Net with the ablated models -w/o-DPF, -w/o-DPP and -w/o-SA, and LCL-Net, its advantages in improving uneven color distortion caused by target distance and wavelength-based light absorption, and recovering lost details are quite obvious. The reason for this lies in the effective utilization of self-attention guidance and biased transfer parameter matrices, which exhibit a remarkable capability to perceive uneven degradation and local detail deterioration. These local differentiated constraints serve as valuable complements to the global guidance provided by either the predicted basic transfer parameters or the extracted ones from the selected template.

II. More Visual Comparison of Other Enhancement Tasks

As mentioned in the main text, we have applied the enhancement strategy proposed in this paper to other types of enhancement tasks and have verified its good general applicability. In Fig. 6, we present additional results to provide a more comprehensive perspective on its performance and characteristics. Both hazy and sand-dust images we presented here were collected from the Internet. It can be observed that these images are photographed from diverse outdoor scenes and exhibit varying levels of degradation. They experience reduced contrast, blurred details, dull appearance, distorted colors, etc. By adding only 200 pairs of hazy and sand-dust image samples each to the original underwater image enhancement training set, TCTL-Net achieves quite good enhancement results for such degraded images. The detailed information is revealed again, and the severe color distortions of some sand-dust images are effectively corrected. It demonstrates that the proposed method has good compatibility and universality for various types of degraded images, indicating a promising application prospect.

References

[1] S. Zhao, L. Zhang, S. Huang, Y. Shen, and S. Zhao, "Dehazing evaluation: Real-world benchmark datasets, criteria, and baselines," *IEEE Transactions on Image Processing*, vol. 29, pp. 6947–6962, 2020.

[2] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao, "An underwater image enhancement benchmark dataset and beyond," *IEEE Transactions on Image Processing*, vol. 29, pp. 4376–4389, 2020.

[3] X. Fu, Z. Fan, M. Ling, Y. Huang, and X. Ding, "Two-step approach for single underwater image enhancement," in *2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2017, pp. 789–794.

[4] W. Song, Y. Wang, D. Huang, and D. Tjondronegoro, "A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration," in *Pacific Rim Conference on Multimedia*, 2018, pp. 678–688.

[5] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 379–393, 2018.

[6] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren, "Underwater image enhancement via medium transmission-guided multi-color space embedding," *IEEE Transactions on Image Processing*, vol. 30, pp. 4985–5000, 2021.

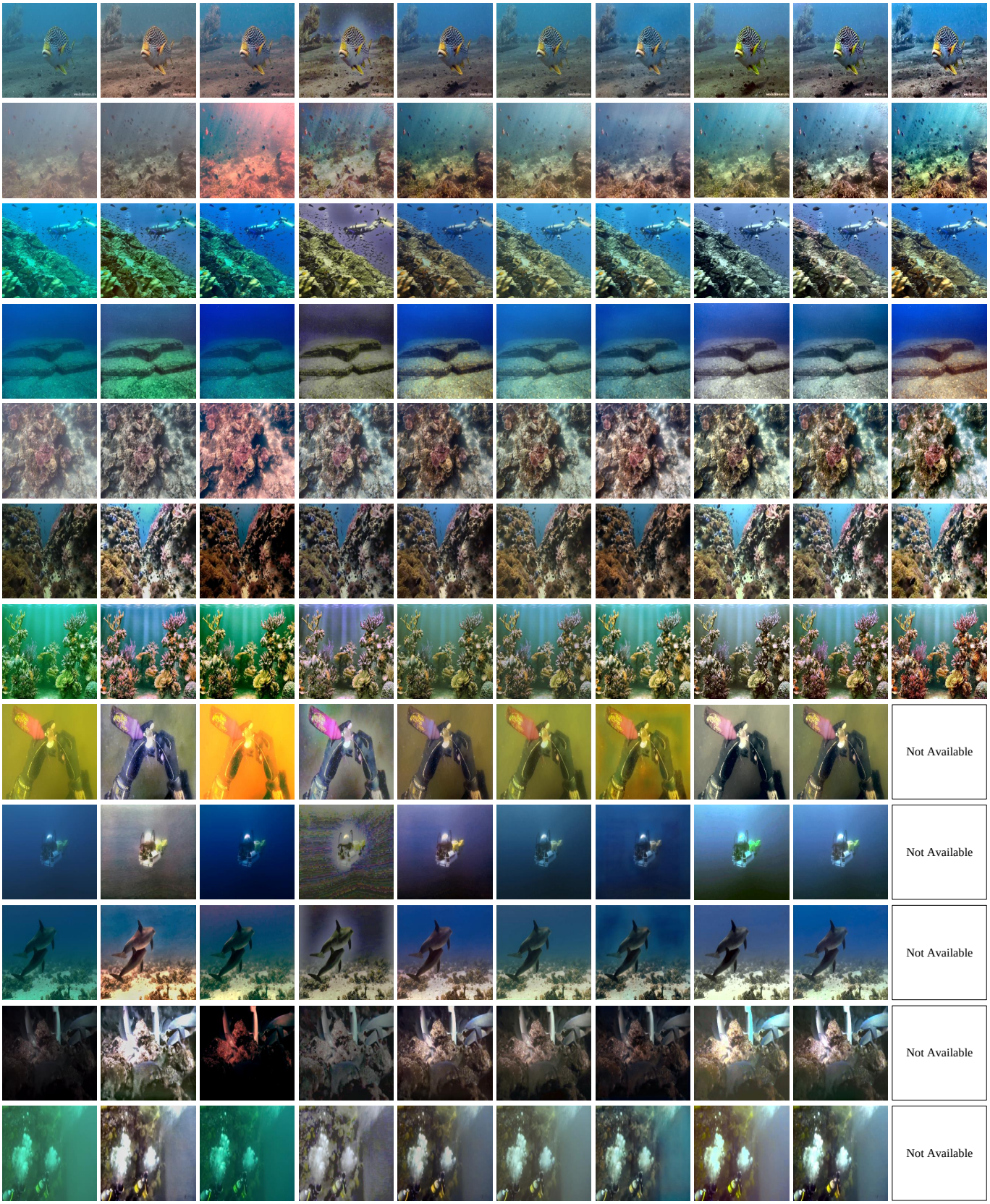
[7] Q. Qi, Y. Zhang, F. Tian, Q. M. Jonathan Wu, K. Li, X. Luan, and D. Song, "Underwater image co-enhancement with correlation feature matching and joint learning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 3, pp. 1133–1147, 2022.

[8] R. Liu, X. Fan, M. Zhu, M. Hou, and Z. Luo, "Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 12, pp. 4861–4875, 2020.

[9] H. Yang, F. Tian, Q. Qi, Q. J. Wu, and K. Li, "Underwater image enhancement with latent consistency learning-based color transfer," *IET Image Processing*, vol. 16, no. 6, pp. 1594–1612, 2022.

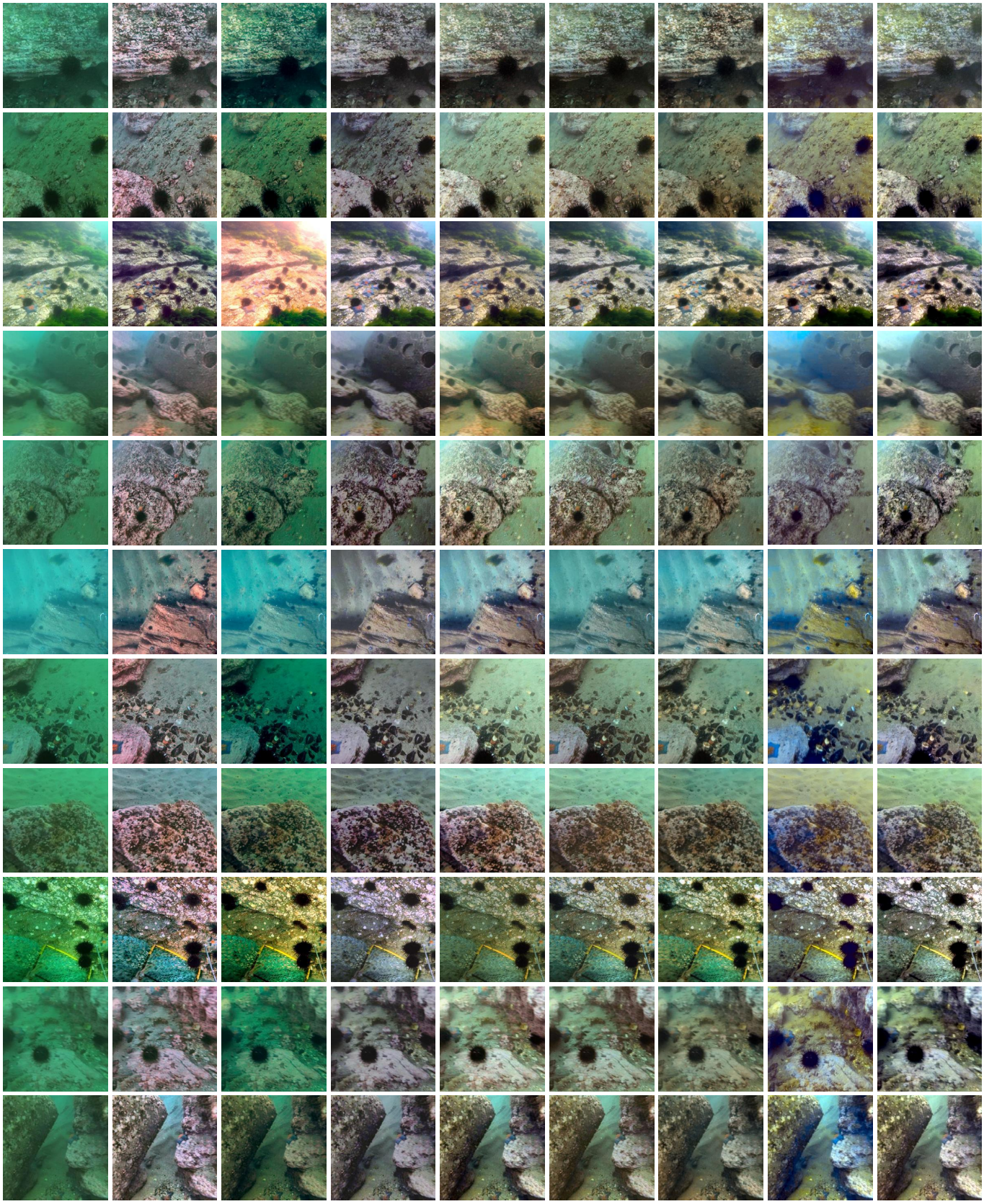
[10] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3227–3234, 2020.

[11] D. Berman, D. Levy, S. Avidan, and T. Treibitz, "Underwater single image color restoration using haze-lines and a new quantitative dataset," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 8, pp. 2822–2837, 2021.



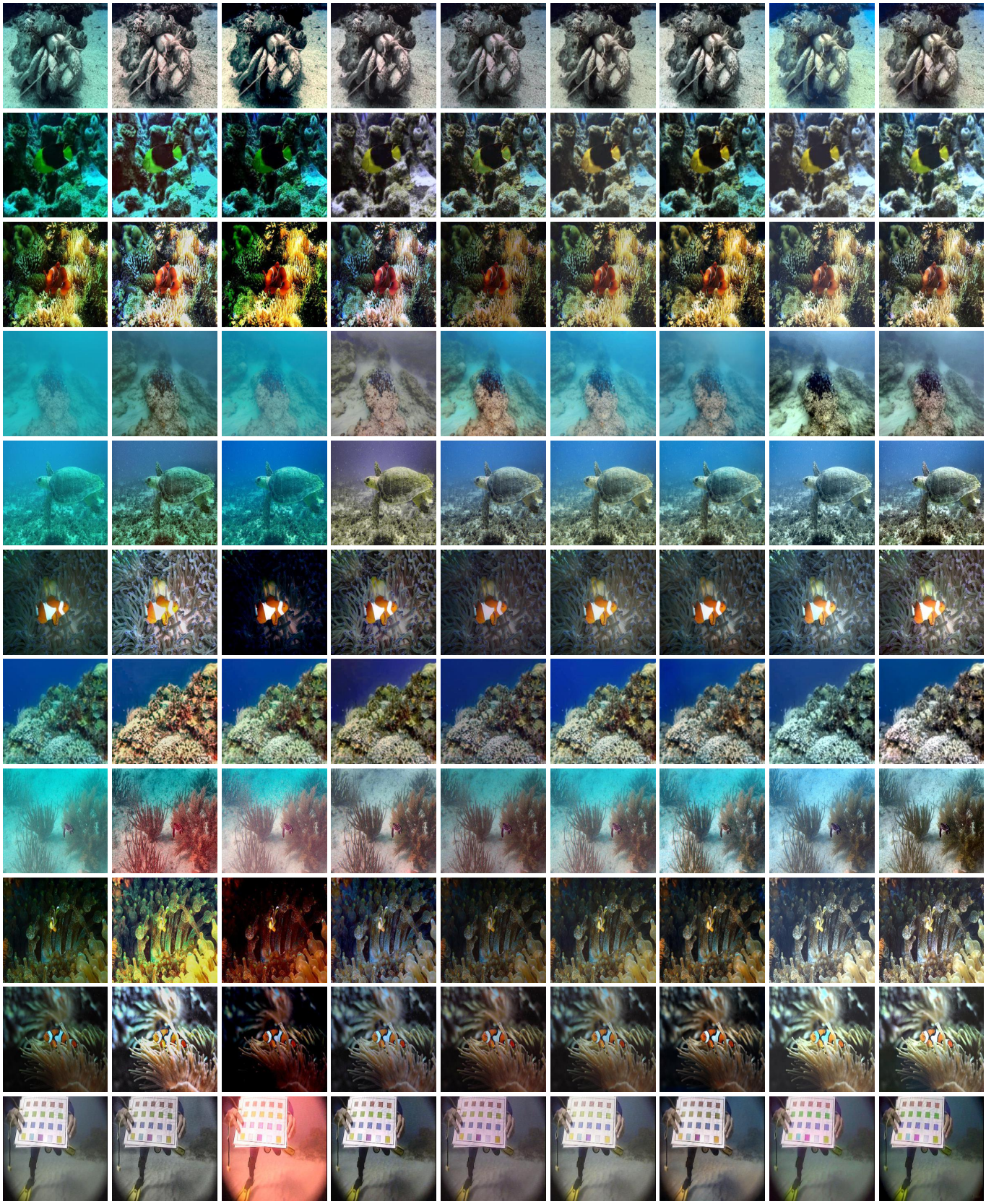
(a) Raw (b) TwoStep (c) ULAP (d) CBF (e) Water-Net (f) Ucolor (g) UICoE-Net (h) LCL-Net (i) TCTL-Net (j) Reference

Fig. 1. More visual comparison of enhancements on the images of the UIEB dataset and the UIEB challenging set [2]. From left to right, (a) the raw images, their enhancements with (b) TwoStep [3], (c) ULAP [4], (d) CBF [5] (e) Water-Net [2], (f) Ucolor [6], (g) UICoE-Net [7], (h) LCL-Net [9], (i) the proposed TCTL-Net, and (j) the given references are represented, respectively.



(a) Raw (b) TwoStep (c) ULAP (d) CBF (e) Water-Net (f) Ucolor (g) UICoE-Net (h) LCL-Net (i) TCTL-Net

Fig. 2. More visual comparison of enhancements on the images of the RUIE dataset [8]. From left to right, (a) the raw images, their enhancements with (b) TwoStep [3], (c) ULAP [4], (d) CBF [5] (e) Water-Net [2], (f) Ucolor [6], (g) UICoE-Net [7], (h) LCL-Net [9], and (i) the proposed TCTL-Net are presented, respectively.

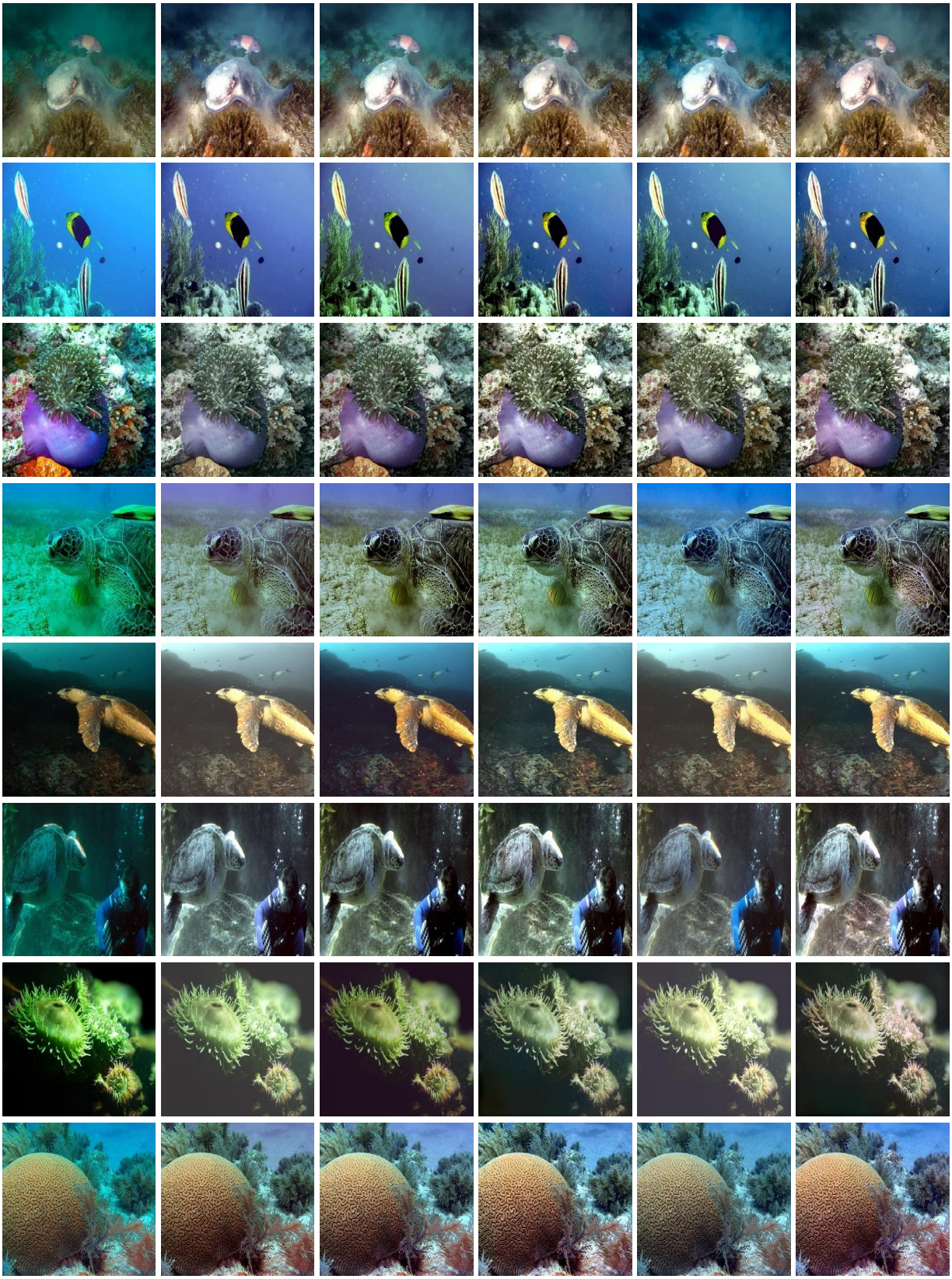


(a) Raw (b) TwoStep (c) ULAP (d) CBF (e) Water-Net (f) Ucolor (g) UICoE-Net (h) LCL-Net (i) TCTL-Net

Fig. 3. More visual comparison of enhancements on the images of the EUVP dataset [10]. From left to right, (a) the raw images, their enhancements with (b) TwoStep [3], (c) ULAP [4], (d) CBF [5] (e) Water-Net [2], (f) Ucolor [6], (g) UICoE-Net [7], (h) LCL-Net [9], and (i) the proposed TCTL-Net are presented, respectively.

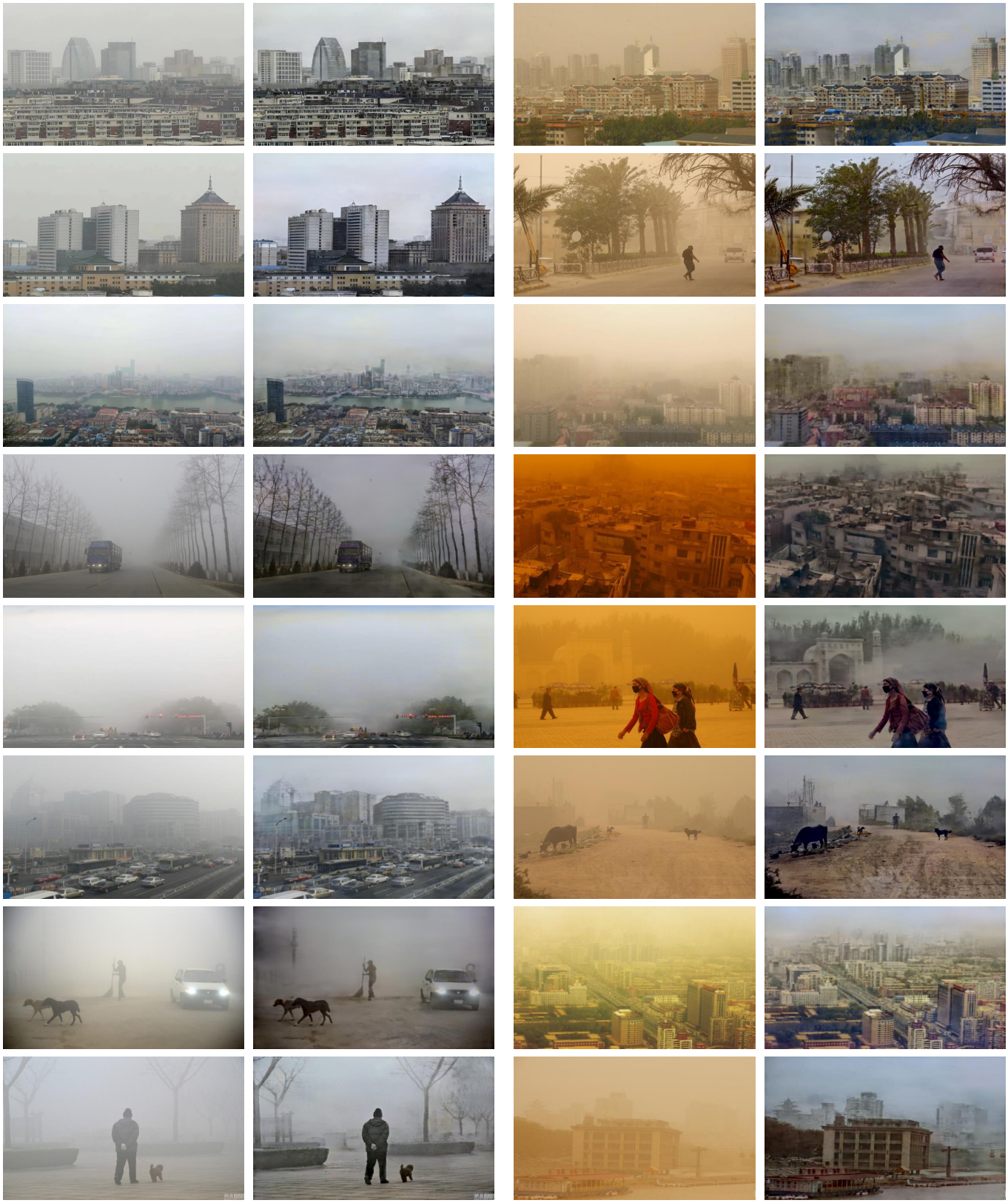
(a) Raw (b) Water-Net (c) Ucolor (d) UICoE-Net (e) LCL-Net (f) TCTL-Net

Fig. 4. The visual comparison of enhancements on the images of the SQUID dataset [11]. From left to right, (a) the raw images of SQUID from four dive sites, their enhancements with (b) Water-Net [2], (c) Ucolor [6], (d) UICoE-Net [7], (e) LCL-Net [9], and (f) the proposed TCTL-Net are presented, respectively.



(a) raw (b) -w/o-DPF (c) -w/o-DPP (d) -w/o-SA (e) LCL-Net (f) TCTL-Net

Fig. 5. The visual comparison of the enhancements on images that exhibit obvious uneven degradation. From left to right, (a) the raw images, enhancements achieved by ablated models (b) -w/o-DPF, (c) -w/o-DPP and (d) -w/o-SA, (e) LCL-Net, and (f) the full TCTL-Net model are presented, respectively.



(a) (b) (c) (d)

Fig. 6. The visual comparison of enhancements on the hazy and sand-dust images. (a) and (c) are the raw hazy and sand-dust images, (b) and (d) are their enhancements with the proposed TCTL-Net, respectively.