2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

MPNet: Multiscale Compensated Probabilistic Adaptive Style Transfer Network for Underwater Image Enhancement

Shangyan Wang[1], Jianhua Yin[2], and Bingrong Xu[1,✉]

[1] School of Automation, Wuhan University of Technology, Wuhan 430070, China
bingrongxu@whut.edu.cn
[2] School of Aerospace Engineering, Beijing Institute of Technology, Beijing, 100081, China

Abstract. Underwater images often suffer from color distortion and low contrast due to complex degradation factors, severely limiting their utility in many fields, such as marine exploration. Although existing methods predominantly focus on establishing a deterministic mapping from the degraded images to the enhanced images, they frequently overlook underwater environmental diversity in water types and lighting conditions. Some studies have noticed this problem, but the proposed methods often cause overcorrection of image colors and low contrast. To address these limitations, we propose the MPNet, a novel deep learning network capable of restoring color details and improving contrast in underwater images. Specifically, our approach introduces a novel framework centered around a Probabilistic Adaptive Style Transfer (PAST) module that integrates depthwise separable convolutions for probabilistic enhancement to achieve more generalized color correction and contrast enhancement. Furthermore, a Multiscale Color-Texture Compensation (MCTC) module is developed through texture-color feedback utilizing parameter-shared SE-Res blocks and cross-layer fusion to mitigate detail loss and color bias in deep networks. Extensive experiments on the UIEB and the EUVP datasets demonstrate improvements in superiority over other advanced methods. Qualitative and visual results confirm its ability to effectively restore the color texture details and enhance contrast.

Keywords: Underwater Image Enhancement, Color Correction, Contrast Enhancement, Conditional Variational Autoencoder

1 Introduction

Underwater Image Enhancement (UIE) plays a pivotal role in marine exploration and robotics, where high-quality visual interpretation is paramount. However, aquatic environments introduce complex degradation through wavelength-selective attenuation, scattering effects, and non-uniform illumination, resulting in color distortion, low contrast, and haze artifacts [1]. These distortions severely impair vision systems, necessitating advanced enhancement techniques to restore perceptual fidelity.

Conventional UIE methods primarily include non-physical enhancement operators and physical model inversion strategies. Early non-physical methods employed pixellevel transformation, such as histogram-based techniques (HE/CLAHE [2]) and retinex-based methods [3]. While computationally efficient, these methods often produce artifacts or over-enhancement under non-uniform lighting or severe turbidity due to oversimplified assumptions [4]. Physical model-based approaches, particularly those implementing an underwater image formation model [5], attempt to reconstruct images by inverting degradation processes through physical principles. However, accurate parameter estimation remains challenging in complex environments, and their rigid mathematical formulations struggle to adapt to diverse optical conditions.

Deep learning has revolutionized UIE through data-driven degradation modeling. Early CNN architectures like WaterNet [1] established the superiority of end-to-end learning over traditional methods. Subsequent innovations include the introduction of frequency-domain processing, such as the wavelet-based dual-stream framework of [6] for frequency-separated reconstruction for color correction and detail enhancement, the introduction of physical priors, such as those of GUPDM [7] to improve model generalization, and the adoption of large-kernel convolution, such as the architecture of UIR-PolyKernel [8] enabling efficient global context modeling. Transformer-based models further enhanced long-range dependency modeling, such as Spectroformer [9], which combines spatial and frequency domain features to preserve detail. To address data scarcity, semi-supervised frameworks like Semi-UIR [10] and contrastive learning approaches like HCLR [11] have reduced reliance on paired training data.

Despite these advances, current approaches primarily attempt a deterministic mapping from the degraded images to enhanced images, which contradict the fundamental reality of underwater environmental diversity in water types and lighting conditions - a critical gap given the ambiguous ground truth in standard benchmarks like UIEB [1], which selects reference images from the results of multiple methods. In response to this challenge, SCNet [12] introduced a whitening module to learn desensitized features of the water type, and PUIE [13] pioneered the incorporation of probabilistic modeling into UIE, achieving robust enhancement by learning enhancement distributions rather than deterministic output. However, images generated by existing methods may exhibit overcorrection characterized by unnatural whitening and inadequate contrast levels, which significantly impair color-fidelity restoration in practical applications.

To address these challenges, this paper proposes a novel UIE framework characterized by a Probabilistic Adaptive Style Transfer (PAST) module and a Multiscale Color-Texture Compensation (MCTC) module, named "MPNet". The PAST module integrates depthwise separable convolutions, which can capture the local and cross-channel information of the image to realize local and global color correction and contrast enhancement for probabilistic enhancement to address the generalization limitations of deterministic methods in complex underwater degradations. Additionally, the MCTC module is developed through texture-color feedback utilizing three parameter-shared SE-Res blocks and multiscale cross-layer fusion, effectively mitigating detail loss and color bias in deep networks and preserving more raw information for the PAST module to realize the restoration of color and texture details. Extensive experiments demonstrate our superiority over other advanced methods and the ability to effectively restore both global and local color and texture details and enhance contrast, confirming enhanced robustness and generalization in diverse underwater conditions.

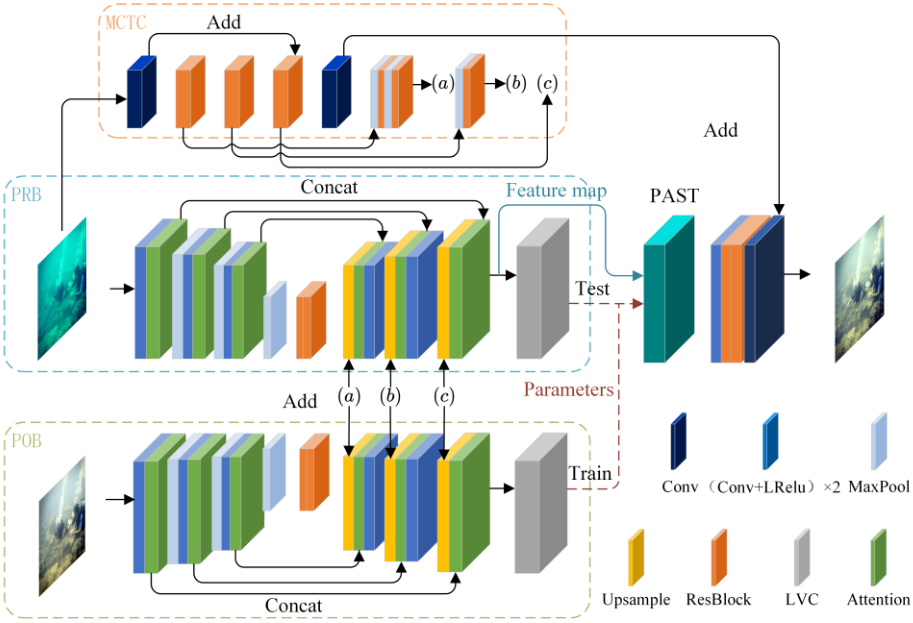

Fig. 1. The overall architecture of the proposed MPNet. The LVC module establishes probabilistic mappings between feature representations and parameter distributions. The PAST module achieves probabilistic adaptive style transfer by the feature map from PRB and the parameters sampling from either the prior distribution (during testing) of PRB or posterior distribution (during training) of POB to realize color correction and contrast enhancement. In the MCTC module, outputs from three parameter-shared SE-Res Blocks undergo dimensional alignment, which includes max pooling and SE-Res Block, before being summed with the corresponding upsampled features at three different scales (denoted as (a), (b), and (c) in the figure), enabling multiscale oringnal information feedback, which is useful for color and texture restoration of images.

2 Methodology

2.1 Preliminary Work

UIE primarily addresses two critical challenges: color bias and low contrast. While maintaining content fidelity, effective UIE requires simultaneous color correction and contrast enhancement, which can be realized by style transfer methodologies. Traditional deterministic approaches often lack flexibility in handling degradation diversity. Probabilistic style transfer offers a novel approach to image enhancement by modeling parameter probability distributions, moving beyond the limitations of fixed parameters.

To model the inherent uncertainty in UIE, this study incorporates a latent variable z to capture stochastic variations, drawing inspiration from probabilistic modeling approaches in image restoration [13]. Instead of deterministic transformation, the enhancement process is formulated through conditional probability:

$$p(y \mid x) \approx \mathbb{E}_{z \sim p(Z \mid x)}[p(y \mid z, x)] \tag{1}$$

where $p(z \mid x)$ denotes the distribution of the latent variable z given the input image x, and $p(y \mid z, x)$ represents the conditional probability of the enhanced image y given x and z. The expectation $\mathbb{E}_{z\sim p(z|x)}$ can be approximated by Maximum Probability (MP) selecting the highest probability sample for stable results, or Monte Carlo (MC) averaging sample likelihoods for diverse results, referring to PUIE [13].

Based on this assumption, we construct our model, the MPNet. The proposed method consists of several key components. These will be explained in detail later.

2.2 Feature Extraction

Fig. 1 illustrates the proposed MPNet architecture, which employs three parallel processing branches (dual U-Net-based [14] branch and the MCTC module) with distinct functional objectives.

Backbone: The feature extraction backbone consists of dual structurally similar U-Net-based Conditional Variational Autoencoder (CVAE) [15] networks. We modified the U-Net architecture by integrating SE-Res Block [16] and AHA attention module [11] to enhance channel dependency modeling and spatial feature localization, which are critical to the effective latent variables extraction. The Prior Branch (PRB) processes input images, and the Posterior Branch (POB) operates on target images. The feature maps extracted from PRB/POB are subsequently fed into the LVC module, which will be covered in the next section, to construct the prior distributions $\mathcal{N}_*^{\mathrm{pr}}$ and posterior distributions $\mathcal{N}_*^{\mathrm{po}}$ for the style transfer parameters. During training, both branches operate concurrently, with the style transfer parameters sampled from $\mathcal{N}_*^{\mathrm{po}}$ to enable target-guided feature adaptation through the PAST module, which will be covered in the next section. We optimize the network by minimizing the Kullback-Leibler (KL) divergence between $\mathcal{N}_*^{\mathrm{pr}}$ and $\mathcal{N}_*^{\mathrm{po}}$, effectively aligning the input image's style representation with that of the target image. During testing, only PRB remains active, where style transfer parameters are sampled from learned $\mathcal{N}_*^{\mathrm{pr}}$ and applied through the PAST module. Both phases employ identical processing pipelines for style-transformed features to generate final enhanced images, achieving reference-guided enhancement in training while maintaining input image-only processing during deployment.

MCTC Module: To address the common issue of color and texture degradation in deep neural networks, particularly the loss of chromatic fidelity in enhanced output, we develop the Multiscale Color-Texture Compensation (MCTC) module containing three sequentially connected parameter-shared SE-Res blocks, a residual connection, and two alignment structures. Each SE-Res block output undergoes dimensional alignment, ensuring compatibility with corresponding upsampling stages in PRB/POB before being progressively fused through multiscale feature additions with corresponding stages, while the final MCTC module output is summed with the backbone's result. Instead of just adding the input to the output of the backbone, the residual structure of the SE-Res Block can preserve useful original texture and color information, while the Squeeze-and-Excitation structure can filter out useless information through channel attention, and parameter sharing design can enhance feature representation without huge computational costs. The hierarchical fusion mechanism establishes persistent texture and color feedback across different network depths, providing more abundant information

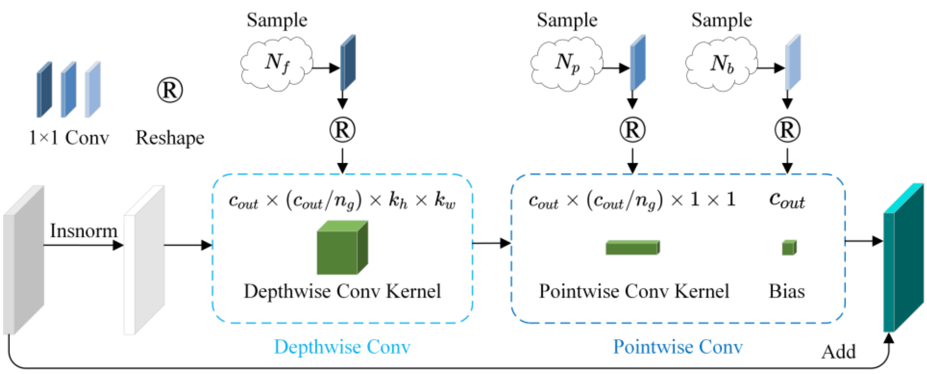

Fig. 2. The architecture of the PAST module.

for the feature extraction stage and image reconstruction stage, effectively alleviating the information loss caused by repeated downsampling-reconstruction processes.

In summary, via multiscale color-texture compensated parallel feature extraction, the network generates rich feature essential for probabilistic style transfer.

2.3 Probabilistic Style Transfer

PAST Module: The PAdaIN module proposed by PUIE [13] combines probability theory with Adaptive Instance Normalization to achieve style transfer through statistical alignment. Although this approach demonstrates effectiveness in adapting global style, its exclusive reliance on global statistical properties (mean and standard) introduces inherent limitations. Specifically, PAdaIN's neglect of capturing localized spatial structures results in compromised preservation of structural details during UIE, manifesting as both local and global perceptible contrast degradation in processed output.

To overcome these limitations, inspired by AdaConv [17], we propose the PAST module, extending adaptive normalization through depthwise separable convolution. As shown in Fig. 2, our architecture operates through two complementary operations:

$$\mathrm{PAST}(x) = v \odot \left[\left(u \circledast \frac{x-\mu_x}{\sigma_x}\right) + s\right] + x \tag{2}$$

where the depthwise convolution operation $\circledast$ with kernel $u \in \mathbb{R}^{c_{\mathrm{out}} \times (c_{\mathrm{out}}/n_g) \times k_h \times k_w}$ performs local spatial adaptation by applying $k_h \times k_w$ filters independently to each input channel. Subsequent pointwise convolution $\odot$ with kernel $v \in \mathbb{R}^{c_{\mathrm{out}} \times (c_{\mathrm{out}}/n_g) \times 1 \times 1}$ and bias $s \in \mathbb{R}^{c_{\mathrm{out}}}$ enables learning cross-channel correlation through linear combinations. The residual connection preserves the details of original feature and allow progressive refinement. We parameterize u, v, s through extracted Gaussian distributions:

$$\begin{array}{lll} u = \Gamma_u(z_u) & z_u \sim \mathcal{N}_f^{\mathrm{pr}}\left(\mu_f(\mathbf{x}), \sigma_f^2(\mathbf{x})\right) & z_u \sim \mathcal{N}_f^{\mathrm{po}}\left(\mu_f(\mathbf{x},\mathbf{y}), \sigma_f^2(\mathbf{x},\mathbf{y})\right) \\ v = \Gamma_v(z_v) & z_v \sim \mathcal{N}_p^{\mathrm{pr}}\left(\mu_p(\mathbf{x}), \sigma_p^2(\mathbf{x})\right) \text{ or } & z_v \sim \mathcal{N}_p^{\mathrm{po}}\left(\mu_p(\mathbf{x},\mathbf{y}), \sigma_p^2(\mathbf{x},\mathbf{y})\right) \\ s = \Gamma_s(z_s) & z_s \sim \mathcal{N}_b^{\mathrm{pr}}\left(\mu_b(\mathbf{x}), \sigma_b^2(\mathbf{x})\right) & z_s \sim \mathcal{N}_b^{\mathrm{po}}\left(\mu_b(\mathbf{x},\mathbf{y}), \sigma_b^2(\mathbf{x},\mathbf{y})\right) \end{array} \tag{3}$$

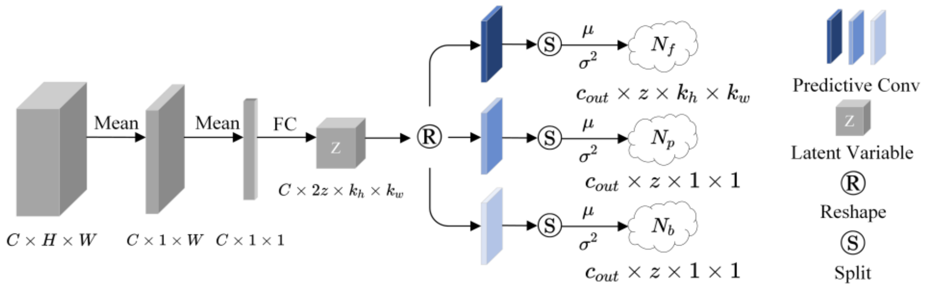

Fig. 3. The architecture of the LVC Module.

where Γ_* are 1×1 convolutions mapping sampled z_* to parameters u, v, s to align the parameter dimensions required for depthwise separable convolutions. $\mathcal{N}_*^{\mathrm{pr}}$ and $\mathcal{N}_*^{\mathrm{po}}$ are prior distributions conditioned on input **x** during training and testing and posterior distributions conditioned on input **x** and target **y** during only training, respectively.

LVC Module: To estimate these parameter distributions $\mathcal{N}_f$, $\mathcal{N}_p$, $\mathcal{N}_b$ which are required for the PAST module, we construct the LVC module as illustrated in Fig. 3. For an input feature map $x \in \mathbb{R}^{C \times H \times W}$, we first perform a spatial average pool along the dimensions of height H and width W to obtain channel statistics $\overline{x} \in \mathbb{R}^C$. These statistics are then projected through a fully connected layer to generate the latent variable:

$$Z = \mathrm{FC}(\overline{x}), \; Z \in \mathbb{R}^{C \times 2z \times k_h \times k_w} \tag{4}$$

The factor $2z$ in the latent dimension accounts for the separate estimation of mean (μ) and variance (σ^2) parameters for each distribution. Three parallel prediction heads subsequently transform Z into the parameters of our target distributions:

$$\begin{aligned} \mathcal{N}_f(\mu_f, \sigma_f^2) &= \mathrm{Conv}_f(Z) \\ \mathcal{N}_p(\mu_p, \sigma_p^2) &= \mathrm{Conv}_p(Z) \\ \mathcal{N}_b(\mu_b, \sigma_b^2) &= \mathrm{Conv}_b(Z) \end{aligned} \tag{5}$$

This design enables efficient learning of degradation patterns while maintaining spatial awareness through the preserved $k_h \times k_w$ dimensions. By operating on compressed channel statistics rather than raw spatial features, the LVC module captures global image characteristics essential for modeling parameter distributions of the PAST module and simultaneously imposes no strict size restrictions on the input images.

In summary, this design enables simultaneous global statistical and local structural alignment. Compared to PAdaIN's global normalization, PAST's spatial-adaptive filtering better preserves underwater image details and enhances contrast through both localized and globalized style transfer operations and residual connection.

2.4 Loss Function

Our loss function combines two components derived from the CVAE framework:

As shown in Equation (3), during training, PRB and POB generate parameter distributions $\mathcal{N}_*^{\text{pr}}$ and $\mathcal{N}_*^{\text{po}}$, respectively, and we minimize the KL divergence between them:

$$\mathcal{L}_{\text{KL}} = \sum_{k \in f,p,b} D_{\text{KL}}\left(\mathcal{N}_k^{\text{pr}}(\mathbf{x}) \parallel \mathcal{N}_k^{\text{po}}(\mathbf{x},\mathbf{y})\right) \tag{6}$$

The reconstruction loss combines three constraints:

$$\mathcal{L}_{\text{R}} = \mathcal{L}_{\text{MAE}} + \lambda\mathcal{L}_{\text{VGG16}} + \beta\mathcal{L}_{\text{MS-SSIM}} \tag{7}$$

where $\mathcal{L}_{\text{MAE}}$ is the mean absolute error loss, $\mathcal{L}_{\text{VGG16}}$ is the perceptual loss based on the VGG16 [18] network, $\mathcal{L}_{\text{MS-SSIM}}$ is the multiscale structural similarity [19] loss, and λ and β is the weight. So the total loss function of the MPNet is:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{KL}} + \gamma\mathcal{L}_{\text{R}} \tag{8}$$

where γ is the weight.

3 Experiments

3.1 Experimental Settings

We evaluated our method on two benchmark datasets: UIEB [1] and EUVP [20]. The UIEB dataset comprises 950 real-world underwater images, including 890 images with high-quality reference results generated through rigorous comparisons and 60 challenging unpaired samples. We randomly split the paired images into 800 training and 90 test samples. The EUVP dataset contains over 20,000 paired and unpaired underwater images captured under diverse conditions. We adopted its EUVP-I subset, consisting of 3,300 training and 400 test image pairs with varying distortion levels, which provides a standardized evaluation for underwater enhancement algorithms.

For quantitative evaluation, we employ both reference-based metrics (PSNR and SSIM) and non-reference metrics (UIQM [21] and UCIQE [22]). Our implementation uses PyTorch on a NVIDIA RTX 4060TI GPU. We preprocess images by resizing them to 256×256 pixels and applying data augmentation through random cropping, horizontal flipping, and rotation. The network is trained for 500 epochs using Adam optimizer (initial learning rate 1e-4, batch size 4) with loss weights $\lambda = 1$, $\beta = 1$, and $\gamma = 5$. The kernel size k_h & k_w of the PAST module are set to 5, and the latent dimension z is set to 128. In this section, we use MP estimation to generate stable results for comparison.

3.2 Comparisons with Other UIE Methods

We compare our proposed method with several advanced UIE approaches, spanning traditional methods (SMBL [23], MLLE [4], HLRP [24]) and deep learning methods (WaterNet [1], PhyNN [25], UWNet [26], SCNet [12], PUIE [13], FAPlus [27], DAUT [28], SFGNet [29]). As shown in Table 1, our method achieves outstanding performance with the highest PSNR and SSIM on both datasets while maintaining

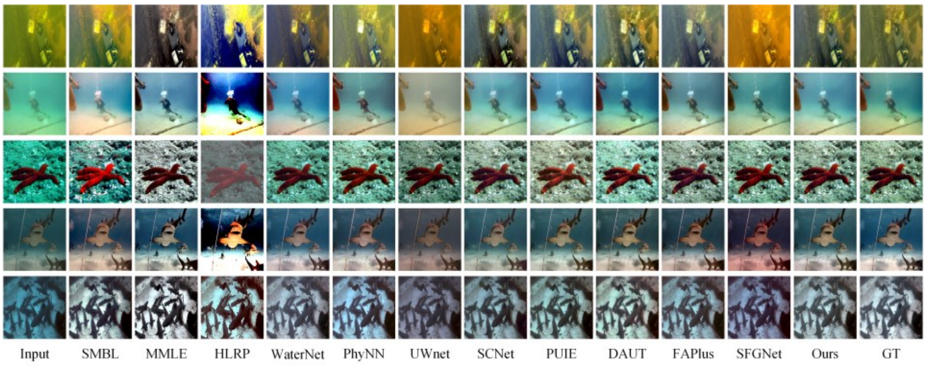

Fig. 4. Visual comparison of experimental results on the UIEB dataset.

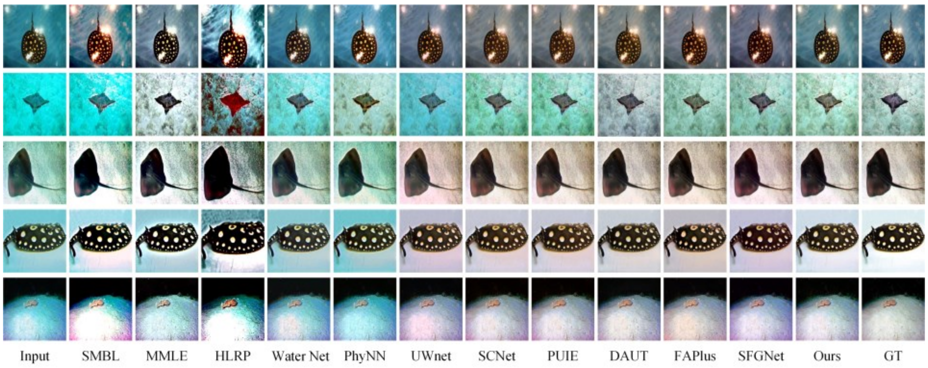

Fig. 5. Visual comparison of experimental results on the EUVP-I dataset.

competitive UIQM scores. The visual results on the UIEB dataset in Fig. 4 and on the EUVP-I dataset in Fig. 5 demonstrate that the proposed method achieves effective color calibration while preserving good contrast characteristics and maintaining enhanced color fidelity, thus ensuring images remain free from chromatic aberrations without compromising visual saturation. We conducted further experiments on a challenging image with severe color cast and low definition, as shown in Fig. 6. The experimental results show that our proposed method can effectively restore both the global and local color texture details and enhance contrast to the maximum extent.

3.3 Ablation Studies

To validate the effectiveness of each component in our proposed method, we performed ablation experiments on the UIEB dataset. The results are shown in Table 2. We start with a basic model and gradually add different components of our method, such as the PAST module, the MCTC module, etc. The results show that each component contributes to the overall performance. We further investigate the role of the MCTC module and the PAST module. As shown in Fig. 7, the MCTC module is helpful for image

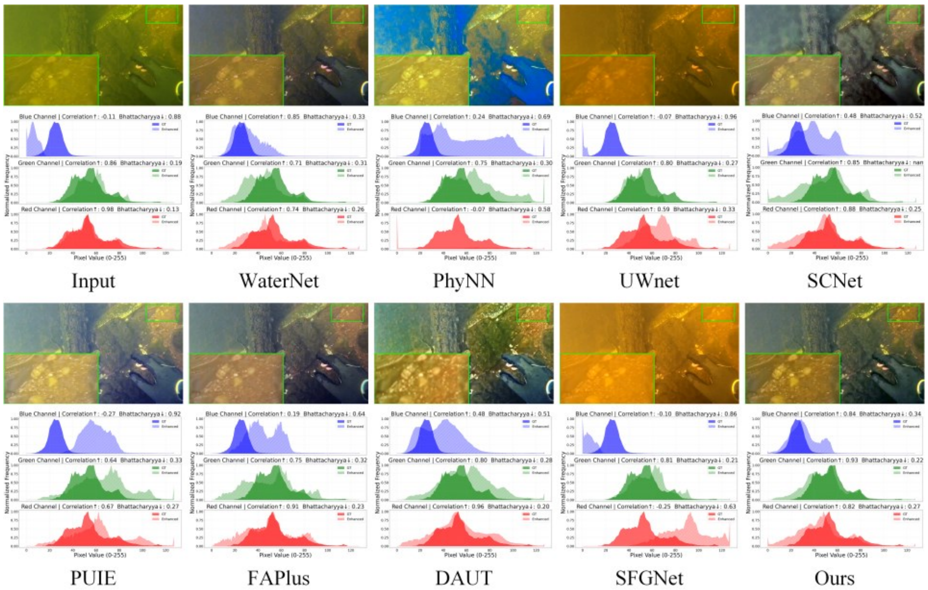

Fig. 6. A more detailed comparison on a selected challenging image. From both the macroscopic contrast and local magnification, it is evident that our proposed method retains the richest details and most accurate colors, without artifacts in SCNet and DAUT and incorrect color and low contrast in other methods. Furthermore, we analyzed the RGB histogram comparison between each enhanced image and the ground truth (GT) image. The results show that our proposed method has the highest histogram similarity to the GT image.

Table 1. Quantitative comparison on the UIEB and the EUVP-I datasets.

Methods	UIEB				EUVP-I			
	PSNR↑	SSIM↑	UIQM↑	UCIQE↑	PSNR↑	SSIM↑	UIQM↑	UCIQE↑
SMBL	16.647	0.764	2.050	0.631	16.083	0.651	2.195	0.617
MLLE	17.207	0.739	1.991	0.617	16.233	0.608	2.519	0.583
HLRP	16.519	0.718	2.463	0.667	15.217	0.562	2.499	0.628
WaterNet	21.198	0.858	2.978	0.612	23.673	0.832	2.991	0.562
PhyNN	18.386	0.802	2.785	0.594	23.583	0.814	2.983	0.553
UWnet	17.453	0.794	2.773	0.553	22.485	0.815	2.960	0.532
SCNet	20.775	0.877	2.987	0.586	24.535	0.822	3.042	0.556
PUIE	21.460	0.896	2.965	0.603	24.381	0.809	3.015	0.557
FAPlus	20.982	0.891	2.914	0.605	24.049	0.819	2.933	0.550
DAUT	21.415	0.875	3.003	0.614	24.709	0.835	3.011	0.553
SFGNet	19.103	0.850	2.698	0.607	24.372	0.835	2.967	0.562
Ours	21.991	0.902	2.962	0.602	25.364	0.841	2.997	0.558

↑ indicates that higher values correspond to better performance. The highest and second-highest values are highlighted in red and blue, respectively.

Table 2. Quantitative results of ablation study on the UIEB dataset.

baseline	attention +resblock	MCTC	PAST	PSNR↑	SSIM↑
√	-	-	-	19.220	0.874
√	√	-	-	21.426	0.891
√	√	-	√	21.819	0.900
√	√	√	-	21.627	0.901
√	√	√	√	**21.991**	**0.902**

↑ indicates that higher values correspond to better performance. The highest values are highlighted in **bold**.

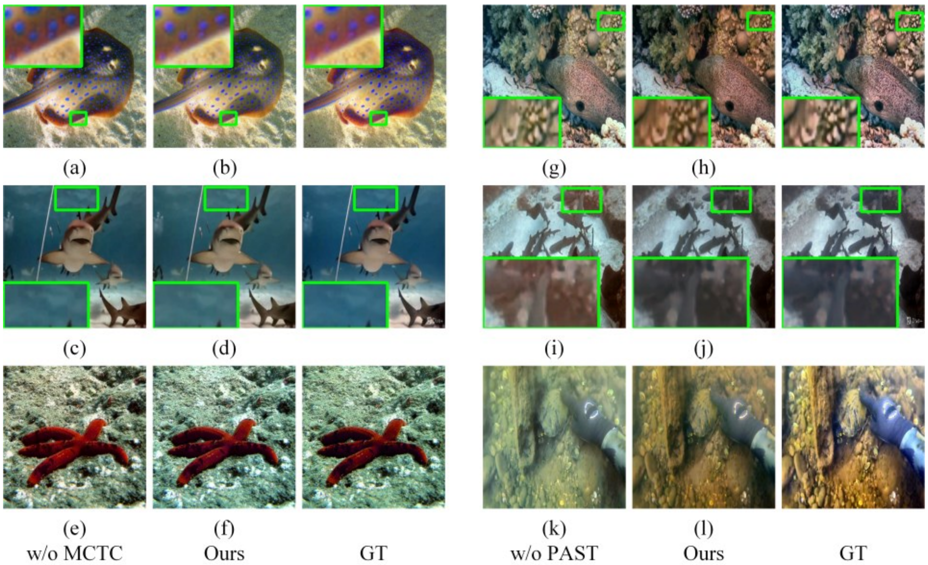

Fig. 7. Visual comparison of ablation experimental results. For the generated results without the MCTC, image (a) has an obvious circle of artifacts at the blue blob and more noise. Image (c) has obvious color fault and noise around the water environment. For image (e), the color is over-corrected, so the ground has lost its due color. However, images (b), (d), and (f) improve these problems and are closer to GT images, which illustrates the effect of the MCTC on the restoration of image texture details (such as artifacts, noise, etc.) and color. For the generated results without the PAST, the contrast of image (g) on the wall is low. Image (i) has a noticeable color shift in the shadows. For image (k), there is insufficient contrast in the whole image. However, images (h), (j), and (l) improve these problems and are closer to GT images, which illustrates the role of the PAST in improving local and global contrast of images and restoring image colors.

texture detail restoration, noise suppression, and color correction, while the PAST module is important for image local and global contrast enhancement and has the ability for color correction. This shows that the combination of the PAST module and the MCTC module is highly complementary. These ablation studies confirm that the combination of all components in our method is essential to achieve the best performance.

4 Conclusion

In this paper, we proposed the MPNet, a novel approach for probabilistic UIE. Our method effectively addresses the challenges of color and texture distortion and low contrast by developing a novel framework centered on the PAST module equipped with the MCTC module. Extensive experiments on multiple benchmark datasets demonstrated that our method achieves superior performance in both quantitative metrics and visual quality compared to other advanced UIE techniques. Future work will focus on addressing the limitations of our method and exploring real-time applications of UIE.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (62206204) and the Natural Science Foundation of Hubei Province, China (2023AFB705).

References

1. Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., Tao, D.: An underwater image enhancement benchmark dataset and beyond. IEEE transactions on image processing 29, 4376–4389 (2019)
2. Hitam, M.S., Awalludin, E.A., Yussof, W.N.J.H.W., Bachok, Z.: Mixture contrast limited adaptive histogram equalization for underwater image enhancement. In: 2013 International conference on computer applications technology (ICCAT). pp. 1–5. IEEE (2013)
3. Hassan, N., Ullah, S., Bhatti, N., Mahmood, H., Zia, M.: The retinex based improved underwater image enhancement. Multimedia Tools and Applications 80, 1839–1857 (2021)
4. Zhang, W., Zhuang, P., Sun, H.H., Li, G., Kwong, S., Li, C.: Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. IEEE Transactions on Image Processing 31, 3997–4010 (2022)
5. Akkaynak, D., Treibitz, T.: A revised underwater image formation model. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 6723–6732 (2018)
6. Ma, Z., Oh, C.: A wavelet-based dual-stream network for underwater image enhancement. In: ICASSP 2022-2022 IEEE international conference on acoustics, speech and signal processing (ICASSP). pp. 2769–2773. IEEE (2022)
7. Mu, P., Xu, H., Liu, Z., Wang, Z., Chan, S., Bai, C.: A generalized physical-knowledge-guided dynamic model for underwater image enhancement. In: Proceedings of the 31st ACM international conference on multimedia. pp. 7111–7120 (2023)
8. Guo, X., Dong, Y., Chen, X., Chen, W., Li, Z., Zheng, F., Pun, C.M.: Underwater image restoration via polymorphic large kernel cnns. arXiv preprint arXiv:2412.18459 (2024)
9. Khan, R., Mishra, P., Mehta, N., Phutke, S.S., Vipparthi, S.K., Nandi, S., Murala, S.: Spectroformer: Multi-domain query cascaded transformer network for underwater image enhancement. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 1454–1463 (2024)
10. Huang, S., Wang, K., Liu, H., Chen, J., Li, Y.: Contrastive semi-supervised learning for underwater image restoration via reliable bank. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 18145–18155 (2023)
11. Zhou, J., Sun, J., Li, C., Jiang, Q., Zhou, M., Lam, K.M., Zhang, W., Fu, X.: Hclr-net: hybrid contrastive learning regularization with locally randomized perturbation for underwater image enhancement. International Journal of Computer Vision 132(10), 4132–4156 (2024)

12. Fu, Z., Lin, X., Wang, W., Huang, Y., Ding, X.: Underwater image enhancement via learning water type desensitized representations. In: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 2764–2768. IEEE (2022)
13. Fu, Z., Wang, W., Huang, Y., Ding, X., Ma, K.K.: Uncertainty inspired underwater image enhancement. In: European conference on computer vision. pp. 465–482. Springer (2022)
14. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015. pp. 234–241. Springer (2015)
15. Rezende, D.J., Mohamed, S., Wierstra, D.: Stochastic backpropagation and approximate inference in deep generative models. In: International conference on machine learning. pp. 1278–1286. PMLR (2014)
16. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7132–7141 (2018)
17. Chandran, P., Zoss, G., Gotardo, P., Gross, M., Bradley, D.: Adaptive convolutions for structure-aware style transfer. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7972–7981 (2021)
18. Johnson, J., Alahi, A., Fei-Fei, L.: Perceptual losses for real-time style transfer and superresolution. In: Computer Vision–ECCV 2016. pp. 694–711. Springer (2016)
19. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. vol. 2, pp. 1398–1402. Ieee (2003）
20. Islam, M.J., Xia, Y., Sattar, J.: Fast underwater image enhancement for improved visual perception. IEEE Robotics and Automation Letters 5(2), 3227–3234 (2020)
21. Panetta, K., Gao, C., Agaian, S.: Human-visual-system-inspired underwater image quality measures. IEEE Journal of Oceanic Engineering 41(3), 541–551 (2015)
22. Yang, M., Sowmya, A.: An underwater color image quality evaluation metric. IEEE Transactions on Image Processing 24(12), 6062–6071 (2015)
23. Song, W., Wang, Y., Huang, D., Liotta, A., Perra, C.: Enhancement of underwater images with statistical model of background light and optimization of transmission map. IEEE Transactions on Broadcasting 66(1), 153–169 (2020)
24. Zhuang, P., Wu, J., Porikli, F., Li, C.: Underwater image enhancement with hyper-laplacian reflectance priors. IEEE Transactions on Image Processing 31, 5442–5455 (2022)
25. Chen, X., Zhang, P., Quan, L., Yi, C., Lu, C.: Underwater image enhancement based on deep learning and image formation model. arXiv preprint arXiv:2101.00991 (2021)
26. Naik, A., Swarnakar, A., Mittal, K.: Shallow-uwnet: Compressed model for underwater image enhancement (student abstract). In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 35, pp. 15853–15854 (2021)
27. Jiang, J., Ye, T., Chen, S., Chen, E., Liu, Y., Jun, S., Bai, J., Chai, W.: Five a+ network: You only need 9k parameters for underwater image enhancement. In: 34th British Machine Vision Conference 2023, BMVC 2023, Aberdeen, UK, November 20-24, 2023. BMVA (2023)
28. Badran, M., Torki, M.: Daut: Underwater image enhancement using depth aware u-shape transformer. In: 2023 IEEE International Conference on Image Processing (ICIP). pp. 1830–1834. IEEE (2023)
29. Zhao, C., Cai, W., Dong, C., Zeng, Z.: Toward sufficient spatial-frequency interaction for gradient-aware underwater image enhancement. In: ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 3220–3224. IEEE (2024)