PhyIC-Net: A Robust Unsupervised Framework for Underwater Image Restoration via Intrinsic Consistency

1st Hongxun Gao
Jiangnan University
Wuxi, China
1831536599@qq.com

2nd Gaoe Qin
Shanghai Jiao Tong University
Shanghai, China
qingaoe@hwaintek.com

3rd Tao Zhang*
Jiangnan Univ. & Central South Univ.
Wuxi & Changsha, China
taozhang@csu.edu.cn

***Abstract*—Underwater image restoration remains a challenging task due to complex optical degradation and the scarcity of paired training data. Existing learning-based methods often rely on paired training data or simplified physical models, which limits their robustness and generalization in real-world scenarios. To address these issues, this paper proposes PhyIC-Net, a robust unsupervised framework that effectively bridges physical modeling with deep learning. To mitigate the overfitting issue common in unsupervised training, we design a Spatial-Channel Decoupled Network (SCD-Net) as the backbone, which explicitly disentangles spatial blurring from color attenuation using depthwise separable convolutions. Furthermore, to constrain the ill-posed solution space of the physical model inversion, we introduce a Physics-Aware Perturbation (PAP) strategy. This strategy simulates diverse water turbidity conditions to enforce an intrinsic consistency constraint, compelling the network to learn invariant scene representations without relying on ground truth data. Finally, a Multi-Prior Guided (MPG) loss is incorporated to regulate color balance and structural smoothness. Extensive experiments on multiple underwater image datasets demonstrate that PhyIC-Net consistently outperforms existing unsupervised and weakly supervised methods in terms of quantitative metrics and visual quality, exhibiting strong generalization across diverse underwater conditions.**

***Index Terms*—Underwater image restoration, unsupervised learning, physical model, intrinsic consistency.**

I. Introduction

Underwater optical imaging plays an irreplaceable role in marine resource exploration and robotic navigation [1]. However, the acquisition of high-fidelity images remains a formidable challenge due to the complex interaction between light and the water medium [2]. As light propagates through water, it suffers from wavelength-dependent absorption and scattering by suspended particles, leading to severe degradation characterized by color distortion, low contrast, and blurred details [3]. These phenomena significantly hinder downstream vision tasks such as object detection [4], [5], making the recovery of clear scenes a critical prerequisite.

Early restoration methods relied on physical priors like the Dark Channel Prior (DCP) [6] and its variants [7]. While computationally simple, these hand-crafted priors lack robustness in varying water conditions. With the advent of deep learning, supervised methods [8], [9] have demonstrated superior performance. However, since acquiring paired "degraded-clean" data in real-world scenarios is impractical, these methods heavily rely on synthetic datasets, inevitably leading to significant domain gaps in real-world deployments [10].

To circumvent paired data dependence, unsupervised learning has emerged as a promising direction. Generative Adversarial Networks (GANs) [11]–[13] attempt to bridge the domain gap but often prioritize perceptual style transfer over physical consistency, introducing geometric artifacts [14]. Alternatively, physics-guided networks [16], [17] integrate optical models into neural architectures. Nevertheless, without sufficient constraints, these architectures often struggle to effectively disentangle scene radiance from degradation components, leading to trivial solutions or underfitting.

To address these challenges, we propose the Physics-aware Intrinsic Consistency Network (PhyIC-Net), a robust unsupervised underwater image restoration framework. The main contributions are summarized as follows:

- We design a Spatial-Channel Decoupled Network (SCD-Net) as the backbone, which explicitly disentangles spatial blurring from color attenuation via depthwise separable convolutions to prevent overfitting in unsupervised training.
- We introduce a Physics-Aware Perturbation (PAP) strategy that simulates diverse turbidity conditions to enforce intrinsic consistency, effectively eliminating degenerate solutions without paired data.
- We propose a Multi-Prior Guided (MPG) loss that synergizes optical and statistical priors to regularize the ill-posed solution space, ensuring optimal color fidelity and structural smoothness.

The source code is available at https://github.com/MaioMao/PhyIC-Net.

This research is supported in part by the National Science and Technology Major Project (2025ZD1009908).

*Corresponding author.

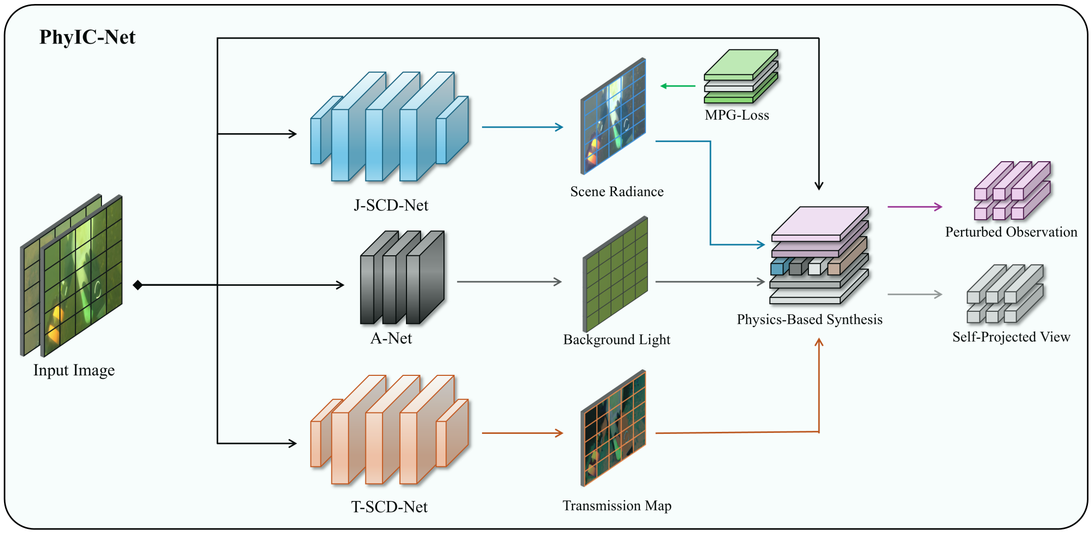

Fig. 1. The overall architecture of the proposed PhyIC-Net. The input image is first disentangled via three parallel branches. The disentangled components are then used to synthesize the Self-Projected View and Perturbed Observation to effectively constrain the unsupervised learning process. Finally, the MPG loss is incorporated to synergize optical and statistical priors, ensuring optimal color fidelity and structural smoothness.

II. Related Work

A. Physical Prior-based Methods

Traditional restoration approaches primarily rely on inverting the underwater optical model using hand-crafted priors. The DCP [6], originally for atmospheric dehazing, often fails in underwater scenes due to the rapid absorption of red light. To address this, variants like the Underwater DCP (UDCP) [18] and Minimum Information Loss Prior [20] were developed to refine transmission estimation. Beyond model inversion, fusion-based strategies [21] and hybrid methods like WWPF [22] combine color correction with contrast enhancement to synthesize high-quality outputs. However, these non-learning methods often rely on rigid parameters, lacking robustness in complex, dynamic water environments.

B. Unsupervised Deep Learning-based Methods

To overcome the domain shift inherent in supervised learning, unsupervised paradigms have gained significant attention. UDnet [28] adapts to uncertainty distributions by incorporating a conditional variational autoencoder and probabilistic adaptive instance normalization to achieve unsupervised enhancement. GAN-based methods, such as WaterGAN [12] and FUnIE-GAN [13], perform domain translation between distorted and clear domains. While visually appealing, these generative approaches may hallucinate artifacts or lose structural details [14]. Conversely, physics-guided frameworks integrate optical models into neural networks. Methods like HybrUR [24] and USUIR [16] attempt to disentangle the input into scene radiance and physical parameters without paired data. Despite their theoretical soundness, these blind decomposition methods face ill-posedness challenges, often risking convergence to trivial identity mappings. Our PhyIC-Net advances this direction by introducing intrinsic consistency constraints to robustly regularize the physical disentanglement process.

III. Method

A. Overall Architecture

As illustrated in Fig. 1, PhyIC-Net is founded on the revised Koschmieder model, treating unsupervised restoration as a physical parameter disentanglement process. Given a raw image I, the network utilizes three parallel branches to estimate the latent components: the J-SCD-Net for clean scene radiance J, the T-SCD-Net for transmission map t, and the non-parametric A-Net for global background light A. These components are then used to synthesize two distinct views to regularize training: 1) A Self-Projected View $I_{proj} = J \cdot t + A \cdot (1 - t)$ is generated via the physical model inversion. A reconstruction loss $\mathcal{L}_{rec} = ||I - I_{proj}||_1$ enforces the fidelity of this decomposition. 2) A Perturbed Observation is synthesized to enforce intrinsic consistency, as detailed in Sec. III-C.

B. Spatial-Channel Decoupled Network

In unsupervised restoration tasks, standard convolutional networks often suffer from overfitting to degradation noise or feature aliasing due to the simultaneous processing of spatial and channel information. To mitigate these issues, we design the SCD-Net as the backbone for J-Net and T-Net. As shown in Fig. 2(a), SCD-Net consists of a Feature Extraction Head, a Spatial-Channel Decoupled Body, and a Physical Mapping Tail. The core component, the SCD-Block, leverages Depthwise Separable Convolutions as a structural regularization mechanism to explicitly decouple the processing of spatial haze and chromatic attenuation. As illustrated in Fig. 2(b), the SCD-Block processes the input features $F_{in} \in \mathbb{R}^{H \times W \times C}$ through four sequential stages.

To begin with, in order to maintain the spatial resolution of features and eliminate boundary artifacts during convolution, we apply a boundary padding operation to the input features.

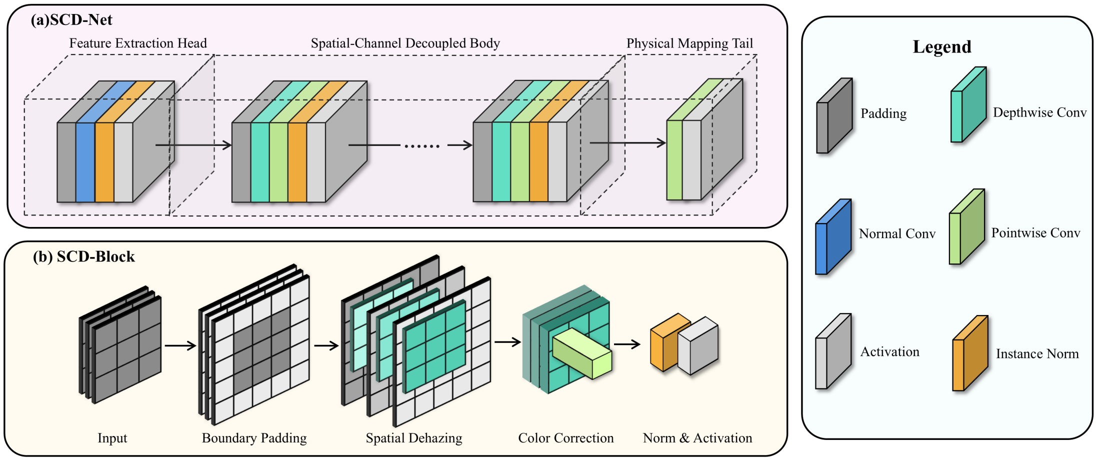

Fig. 2. The architecture of the proposed SCD-Net. (a) The overall pipeline of the network. (b) The detailed structure of the SCD-Block, which utilizes depthwise separable convolutions to decouple spatial and channel features.

This preprocessing step ensures that the subsequent spatial filtering does not introduce edge noise, which is critical for preserving the continuity of the underwater scene. Subsequently, a Depthwise Convolution is performed to filter spatial information independently for each channel. Since the haze effect caused by scattering primarily manifests as spatial blurring, this operation focuses on spatial neighbors to remove haze-induced blur and preserve texture details without mixing channel information:

$$F_{DW}^{(c)} = K_{DW}^{(c)} * \text{Pad}(F_{in})^{(c)}, \quad c = 1, 2, \ldots, C \tag{1}$$

where $K_{DW}^{(c)}$ represents the spatial kernel for the c-th channel, and Pad$(\cdot)$ denotes the boundary padding.

Following the spatial filtering, a Pointwise Convolution is employed to perform linear combinations across channels. This step models the correlations between spectral channels to correct color shifts caused by wavelength-selective absorption. Crucially, this operation changes the channel interactions without altering the spatial receptive field, thereby achieving the explicit decoupling of color correction from spatial dehazing:

$$F_{PW} = W_{PW} * F_{DW} + b_{PW} \tag{2}$$

where W_{PW} and b_{PW} denote the pointwise weights and bias, respectively. Finally, the features pass through an Instance Normalization (IN) layer and a LeakyReLU activation function $\phi(\cdot)$ to unify the feature distribution and introduce non-linearity:

$$F_{out} = \phi(\text{IN}(F_{PW})) \tag{3}$$

By stacking multiple SCD-Blocks, the network effectively disentangles the orthogonal degradation factors of spatial blurring and color casting while maintaining a lightweight structure, preventing the unsupervised model from overfitting to trivial solutions.

C. Physics-Aware Perturbation Strategy

Single-image blind restoration is inherently an ill-posed problem with infinite feasible solutions. Although the physical reconstruction loop ensures that the decomposed components are mathematically self-consistent, it does not guarantee that they are physically correct, often leading the unsupervised network to converge towards degenerate solutions such as identity mapping. To further constrain the solution space and enforce the network to learn invariant scene properties, we propose the PAP training strategy. Distinct from conventional data augmentation techniques that add random noise, PAP generates "Perturbed Observations" that strictly adhere to the optical laws of underwater imaging. Specifically, utilizing the estimated clean radiance J, background light A, and the raw input I, we synthesize a new degraded image I_{pert} through a convex combination:

$$I_{pert} = \alpha \cdot I + \beta \cdot J + \gamma \cdot A \tag{4}$$

where the perturbation coefficients satisfy the constraint $\alpha + \beta + \gamma = 1$. Mathematical derivation proves that this linear mixing is strictly equivalent to maintaining the intrinsic scene properties (J and A) invariant while linearly transforming the transmission map t to a new state $t_{pert} = \alpha t + \beta$. Consequently, this operation simulates a physically plausible re-degradation process under varying water turbidity levels rather than simple image enhancement.

To simulate the complex and dynamic turbidity of real-world underwater environments, the coefficients (α, β, γ) are sampled from a multivariate Dirichlet distribution, denoted as $(\alpha, \beta, \gamma) \sim Dir(\eta)$. This non-uniform sampling strategy allows for the generation of diverse degradation scenarios, ranging from slight haze to severe turbidity, thereby preventing the network from overfitting to a specific degradation pattern. Based on the generated perturbed observation I_{pert}, we construct an Intrinsic Consistency constraint. The core hypothesis is that regardless of how the external water medium changes, the intrinsic color and structure of the object should remain constant. Therefore, we feed I_{pert} back into the shared-weight J-SCD-Net to predict a secondary scene radiance J_{pert}. The

TABLE I
QUANTITATIVE COMPARISON ON THE EUVP DATASET. BEST: **BOLD**, SECOND: UNDERLINED.

Method	**PSNR↑**	**SSIM↑**	**MSE($\times 10^3$) ↓**	**UIQM↑**	**UICM↑**	**UISM↑**	**UIConM↑**	**UCIQE↑**	**SigmaC↑**	**ConL↑**	**MuS↑**
IBLA [29]	17.21	0.713	2.098	7.1822	8.4585	5.8611	1.4580	0.4290	0.1025	0.8677	0.5545
WaterFusion [21]	17.58	0.744	1.341	6.0914	7.6862	5.8779	1.1576	0.4144	0.0951	0.9218	0.4538
HE Prior [30]	15.14	0.636	2.444	6.2514	**9.5888**	4.9411	1.2647	0.4238	0.1042	0.9447	0.4493
Water-Net [31]	20.99	0.796	0.637	4.9480	5.5713	5.9721	0.8468	0.3503	0.0806	0.7715	0.3911
USUIR [16]	18.00	0.737	1.273	5.1182	7.5440	6.0073	0.8759	0.3737	0.0761	0.8640	0.3918
UUUIR [26]	14.60	0.656	2.588	6.7850	8.9895	5.5807	1.3659	0.4052	0.0916	0.9219	0.4243
HFM [23]	17.02	0.705	1.575	8.5766	6.2513	5.5614	1.8902	0.3882	0.0888	0.9219	0.3635
WWPF [22]	15.82	0.640	1.940	8.5393	5.4694	5.3694	1.9018	0.4033	0.0773	0.9757	0.3854
UDnet [28]	19.55	0.777	0.863	4.2092	4.8427	6.1506	0.6311	0.3206	0.0571	0.7350	0.3575
KMUUIR [27]	18.86	0.752	1.033	4.9932	7.0961	5.9837	0.8464	0.3665	0.0754	0.8475	0.3827
Ours	**21.41**	**0.813**	**0.607**	**10.1760**	8.8137	**8.8456**	**2.0461**	**0.4651**	**0.1084**	**0.9805**	**0.5641**

network is optimized by minimizing the consistency loss $\mathcal{L}_{cons}$ between the secondary estimate and the primary estimate:

$$\mathcal{L}_{cons} = ||\mathrm{sg}(J) - J_{pert}||_2^2 \tag{5}$$

where $\mathrm{sg}(\cdot)$ represents the stop-gradient operation. By treating the primary estimation J as a pseudo-label, this mechanism effectively eliminates trivial solutions and forces the network to extract robust features that are invariant to medium perturbations, significantly enhancing the generalization capability of the model.

D. Multi-Prior Guided Loss and Total Objective

To regularize the ill-posed unsupervised solution space, we incorporate a MPG loss, denoted as $\mathcal{L}_{mpg}$, which synergizes four statistical constraints. First, to correct the blue-green dominance, we employ a Color Balance constraint ($\mathcal{L}_{bal}$) based on the Gray-World assumption. Unlike standard methods, we adopt a relative balance strategy that enforces the convergence of global channel means without fixing them to a specific value. Second, derived from Retinex theory, a Max-RGB constraint ($\mathcal{L}_{max}$) is applied to stretch the dynamic range and correct illumination by maximizing local intensities. Third, to remove haze while preserving details, we propose a Weighted Underwater DCP constraint ($\mathcal{L}_{wudcp}$). This term improves upon the standard Dark Channel Prior by assigning lower weights to the red channel, effectively avoiding the over-enhancement caused by red light absorption. Finally, a Total Variation regularization ($\mathcal{L}_{tv}$) is utilized to suppress high-frequency artifacts and ensure spatial smoothness.

Formally, the MPG loss is defined as the weighted sum of these four components:

$$\mathcal{L}_{mpg} = \lambda_b \mathcal{L}_{bal} + \lambda_m \mathcal{L}_{max} + \lambda_u \mathcal{L}_{wudcp} + \lambda_t \mathcal{L}_{tv} \tag{6}$$

where λ denotes the balancing weights. Consequently, the total objective function is formulated as:

$$\mathcal{L}_{total} = \lambda_{rec} \mathcal{L}_{rec} + \lambda_{cons} \mathcal{L}_{cons} + \lambda_{mpg} \mathcal{L}_{mpg} \tag{7}$$

By jointly optimizing these terms, PhyIC-Net effectively balances physical fidelity, intrinsic consistency, and visual perceptual quality.

IV. EXPERIMENT

A. Experimental Settings

We implemented the proposed PhyIC-Net using the PyTorch framework. All experiments were conducted on a workstation equipped with an Intel Core i5-14600 CPU and an NVIDIA GeForce RTX 4070 Ti GPU. The network was optimized using the Adam optimizer with an initial learning rate of 1×10^{-4}. The training process spanned 150 epochs with a batch size of 1. For the hyper-parameters in the total objective function, we empirically set the weights of the physical reconstruction loss and the intrinsic consistency loss to $\lambda_{rec} = 1$ and $\lambda_{cons} = 1$, respectively. The weight for the MPG loss was set to $\lambda_{mpg} = 1 \times 10^{-3}$. Furthermore, the internal balancing weights for the MPG-Loss terms $(\lambda_b, \lambda_m, \lambda_u, \lambda_t)$ were all set to 1.

B. Comparative Experiments

We evaluated PhyIC-Net against ten state-of-the-art methods on the EUVP [32], UIEB [31], and RUIE [33] datasets. To measure reconstruction accuracy, we employed full-reference metrics including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM), and Mean Squared Error (MSE). Furthermore, we assessed perceptual quality using non-reference metrics: Underwater Image Quality Measure (UIQM) [34] with its components (UICM, UISM, UIConM), Underwater Color Image Quality Evaluation (UCIQE) [35], and statistical indicators including Chroma spread (SigmaC), Local contrast (ConL), and Saturation mean (MuS).

Table I and Table II present the quantitative comparisons on paired datasets. On the EUVP dataset, our method achieves the best performance with a PSNR of 21.41 dB and an SSIM of 0.813. For the challenging UIEB dataset containing real-world degradation, PhyIC-Net maintains superior performance, reaching 22.47 dB in PSNR and 0.891 in SSIM. This result surpasses the recent unsupervised method KMUUIR by a margin of 1.33 dB in PSNR. In the no-reference evaluation on the RUIE dataset shown in Table III, our approach secures the highest scores in UIQM and UCIQE, recorded at 9.0242 and 0.4842 respectively. These metrics indicate that our restored images possess higher contrast and better perceptual quality compared to the competitors.

TABLE II

QUANTITATIVE COMPARISON ON THE UIEB DATASET. BEST: **BOLD**, SECOND: <u>UNDERLINED</u>.

Method	PSNR↑	SSIM↑	MSE($\times 10^3$)↓	UIQM↑	UICM↑	UISM↑	UIConM↑	UCIQE↑	SigmaC↑	ConL↑	MuS↑
IBLA [29]	17.85	0.767	1.832	7.6447	8.1731	8.4773	1.3735	0.4233	<u>0.0981</u>	0.8463	<u>0.5632</u>
WaterFusion [21]	<u>22.35</u>	0.868	0.936	5.3037	8.5140	8.3512	0.7265	0.4168	0.0903	0.9109	0.4834
HE Prior [30]	18.31	0.792	1.745	6.2800	<u>10.1555</u>	8.4937	0.9749	<u>0.4266</u>	**0.1169**	0.9300	0.4527
Water-Net [31]	18.61	0.848	1.225	4.4830	5.2630	8.1794	0.5368	0.3347	0.0695	0.7165	0.4097
USUIR [16]	20.21	0.859	0.766	4.8547	7.9476	7.9931	0.6350	0.3797	0.0758	0.8768	0.4019
UUUIR [26]	14.67	0.500	2.682	**8.6165**	8.5287	7.1782	**1.7499**	0.3887	0.0890	0.8686	0.4217
HFM [23]	17.17	0.811	1.562	7.4214	6.3070	<u>8.7077</u>	1.3068	0.3794	0.0910	0.9020	0.3465
WWPF [22]	18.42	0.803	1.193	6.8567	5.4764	7.5055	1.2547	0.3882	0.0695	<u>0.9682</u>	0.3489
UDnet [28]	22.09	<u>0.872</u>	<u>0.739</u>	4.0554	4.9377	8.3480	0.4058	0.3199	0.0541	0.7181	0.3785
KMUUIR [27]	21.14	0.848	0.808	5.7313	9.2739	7.5740	0.9043	0.4006	0.0903	0.9405	0.4120
Ours	**22.47**	**0.891**	**0.612**	<u>7.9262</u>	**10.3256**	**8.9221**	<u>1.3986</u>	**0.4527**	0.0818	**0.9719**	**0.5733**

TABLE III

QUANTITATIVE COMPARISON ON THE RUIE DATASET. BEST: **BOLD**, SECOND: <u>UNDERLINED</u>.

Method	UIQM↑	UICM↑	UISM↑	UIConM↑	UCIQE↑	SigmaC↑	ConL↑	MuS↑
IBLA [29]	4.3102	4.6366	7.6568	0.5366	0.3444	0.0684	0.6385	<u>0.5325</u>
WaterFusion [21]	4.8721	<u>8.8103</u>	6.3335	0.7701	0.3792	0.0876	0.8834	0.3716
HE Prior [30]	6.1075	**11.2563**	6.5109	1.0817	<u>0.4244</u>	<u>0.1098</u>	0.9344	0.4524
Water-Net [31]	3.9433	4.5714	6.8923	0.4976	0.3071	0.0532	0.7856	0.2582
USUIR [16]	4.5112	7.3175	6.9333	0.6314	0.3701	0.0768	0.8811	0.3584
UUUIR [26]	3.9834	5.6067	6.9714	0.4941	0.2687	0.0675	0.6346	0.2444
HFM [23]	<u>6.7523</u>	3.0242	7.0919	1.2790	0.3282	0.0474	0.8959	0.2331
WWPF [22]	6.7287	3.7426	6.1710	<u>1.3428</u>	0.3499	0.0485	<u>0.9513</u>	0.2566
UDnet [28]	3.7573	4.4905	<u>7.7407</u>	0.3762	0.2626	0.0550	0.6146	0.2644
KMUUIR [27]	4.3604	6.7382	6.9062	0.5960	0.3558	0.0714	0.8628	0.3321
Ours	**9.0242**	8.4764	**8.5055**	**1.7547**	**0.4842**	**0.1173**	**1.0573**	**0.5400**

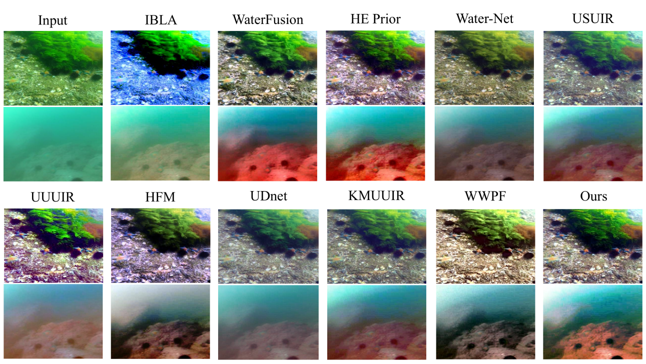

Fig. 3. Visual comparison with state-of-the-art methods.

Fig. 3 illustrates the qualitative comparison results. Compared with other baselines, our PhyIC-Net recovers more natural colors and clearer details, effectively suppressing underwater degradation.

C. Ablation Study

To verify the contribution of each component in PhyIC-Net, we conducted ablation studies on the UIEB dataset using three variants: 1) w/o SCD, which replaces the SCD-Net with standard convolutions; 2) w/o Consistency, which removes the intrinsic consistency constraint; and 3) w/o MPG, which trains the network without the MPG loss.

Table IV details the quantitative results. The Full Model achieves the best performance with a PSNR of 22.47 dB and an SSIM of 0.891. Removing the SCD module results in a performance drop to 21.36 dB in PSNR, indicating that standard convolutions fail to effectively disentangle spatial haze from channel-wise attenuation. The most significant degradation occurs in the w/o Consistency variant where PSNR drops to 20.26 dB, demonstrating that the PAP strategy is critical for regularizing the unsupervised learning process and preventing unstable optimization. Furthermore, the exclusion of the MPG loss leads to a PSNR decrease to 21.77 dB, confirming its necessity for guiding the network towards visually pleasing results with correct color balance.

TABLE IV

ABLATION STUDY OF PHYIC-NET ON THE UIEB DATASET. THE BEST RESULTS ARE HIGHLIGHTED IN BOLD.

Configuration	PSNR ↑	SSIM ↑
w/o SCD	21.36	0.862
w/o Consistency	20.26	0.855
w/o MPG	21.77	0.884
Full Model	**22.47**	**0.891**

V. CONCLUSION

In this paper, we proposed PhyIC-Net, a robust unsupervised framework for underwater image restoration. By synergizing the SCD-Net with a PAP strategy, our method effectively prevents overfitting to noise and constrains the solution space through intrinsic consistency, eliminating the need for paired training data. Experimental results on the EUVP, UIEB, and RUIE datasets demonstrate that PhyIC-Net outperforms state-of-the-art approaches in terms of color correction, detail recovery, and quantitative metrics. However, since our approach relies on the simplified atmospheric scattering model, it may exhibit limitations in handling scenes with complex artificial lighting or highly non-uniform optical fields. Future work will focus on extending this physics-consistent paradigm to underwater video enhancement and optimizing the architecture for real-time deployment on autonomous marine vehicles.

References

[1] M. Jian, X. Liu, H. Luo, X. Lu, H. Yu, and J. Dong, "Underwater image processing and analysis: A review," *Signal Processing: Image Communication*, vol. 91, p. 116088, 2021.

[2] Y. M. Alsakar, N. A. Sakr, S. El-Sappagh, T. Abuhmed, and M. Elmogy, "Underwater image restoration and enhancement: a comprehensive review of recent trends, challenges, and applications," *The Visual Computer*, vol. 41, no. 6, pp. 3735–3783, 2025.

[3] L. Chen, Y. Huang, J. Dong, Q. Xu, S. Kwong, H. Lu, H. Lu, and C. Li, "Underwater optical object detection in the era of artificial intelligence: current, challenge, and future," *ACM Computing Surveys*, vol. 58, no. 3, pp. 1–34, 2025.

[4] C. Fu, R. Liu, X. Fan, P. Chen, H. Fu, W. Yuan, M. Zhu, and Z. Luo, "Rethinking general underwater object detection: Datasets, challenges, and solutions," *Neurocomputing*, vol. 517, pp. 243–256, 2023.

[5] S. Mittal, S. Srivastava, and J. P. Jayanth, "A survey of deep learning techniques for underwater image classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 10, pp. 6968–6982, 2022.

[6] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010.

[7] Y.-T. Peng, K. Cao, and P. C. Cosman, "Generalization of the dark channel prior for single image restoration," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2856–2868, 2018.

[8] Y. Xie, L. Kong, K. Chen, Z. Zheng, X. Yu, Z. Yu, and B. Zheng, "Uveb: A large-scale benchmark and baseline towards real-world underwater video enhancement," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2024, Conference Proceedings, pp. 22 358–22 367.

[9] C. Zhao, W. Cai, C. Dong, and C. Hu, "Wavelet-based fourier information interaction with frequency diffusion adjustment for underwater image restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, Conference Proceedings, pp. 8281–8291.

[10] T. Ye, S. Chen, Y. Liu, Y. Ye, E. Chen, and Y. Li, "Underwater light field retention: Neural rendering for underwater imaging," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, Conference Proceedings, pp. 488–497.

[11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[12] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robotics and Automation letters*, vol. 3, no. 1, pp. 387–394, 2017.

[13] 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.

[14] K. Kishan Babu, A. Tabassum, B. Navaneeth, T. Jahnavi, and Y. Akshaya, "Underwater image enhancement using generative adversarial networks: a survey," *International Journal of Computers and Applications*, vol. 47, no. 4, pp. 356–372, 2025.

[15] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, Conference Proceedings, pp. 9446–9454.

[16] Z. Fu, H. Lin, Y. Yang, S. Chai, L. Sun, Y. Huang, and X. Ding, "Unsupervised underwater image restoration: From a homology perspective," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 36, 2022, Conference Proceedings, pp. 643–651.

[17] N. Varghese, A. Kumar, and A. Rajagopalan, "Self-supervised monocular underwater depth recovery, image restoration, and a real-sea video dataset," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, Conference Proceedings, pp. 12 248–12 258.

[18] P. Drews Jr, E. do Nascimento, F. Moraes, S. Botelho, and M. Campos, "Transmission estimation in underwater single images," in *Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops*, 2013, Conference Proceedings, pp. 825–830.

[19] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," *Journal of Visual Communication and Image Representation*, vol. 26, pp. 132–145, 2015.

[20] C.-Y. Li, J.-C. Guo, R.-M. Cong, Y.-W. Pang, and B. Wang, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5664–5677, 2016.

[21] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *2012 IEEE conference on computer vision and pattern recognition*. IEEE, 2012, Conference Proceedings, pp. 81–88.

[22] W. Zhang, L. Zhou, P. Zhuang, G. Li, X. Pan, W. Zhao, and C. Li, "Underwater image enhancement via weighted wavelet visual perception fusion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 4, pp. 2469–2483, 2023.

[23] S. An, L. Xu, I. Senior Member, Z. Deng, and H. Zhang, "Hfm: A hybrid fusion method for underwater image enhancement," *Engineering Applications of Artificial Intelligence*, vol. 127, p. 107219, 2024.

[24] S. Yan, X. Chen, Z. Wu, M. Tan, and J. Yu, "Hybrur: A hybrid physical-neural solution for unsupervised underwater image restoration," *IEEE Transactions on image processing*, vol. 32, pp. 5004–5016, 2023.

[25] A. Kar, S. K. Dhara, D. Sen, and P. K. Biswas, "Zero-shot single image restoration through controlled perturbation of koschmieder's model," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, Conference Proceedings, pp. 16 205–16 215.

[26] S. Chai, Z. Fu, Y. Huang, X. Tu, and X. Ding, "Unsupervised and untrained underwater image restoration based on physical image formation model," in *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, Conference Proceedings, pp. 2774–2778.

[27] S. Zhang, D. An, D. Li, and R. Zhao, "Unsupervised underwater image restoration via koschmieder model disentanglement," *Expert Systems with Applications*, vol. 266, p. 126075, 2025.

[28] A. Saleh, M. Sheaves, D. Jerry, and M. R. Azghadi, "Adaptive deep learning framework for robust unsupervised underwater image enhancement," *Expert Systems with Applications*, vol. 268, p. 126314, 2025.

[29] Y.-T. Peng and P. C. Cosman, "Underwater image restoration based on image blurriness and light absorption," *IEEE transactions on image processing*, vol. 26, no. 4, pp. 1579–1594, 2017.

[30] C.-Y. Li, J.-C. Guo, R.-M. Cong, Y.-W. Pang, and B. Wang, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," *IEEE Transactions on Image Processing*, vol. 25, no. 12, pp. 5664–5677, 2016.

[31] 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, 2019.

[32] 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.

[33] 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.

[34] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 3, pp. 541–551, 2015.

[35] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 6062–6071, 2015.