

LIGHTWEIGHT UNDERWATER IMAGE ENHANCEMENT VIA IMPULSE RESPONSE OF LOW-PASS FILTER BASED ATTENTION NETWORK

May Thet Tun, Yosuke Sugiura, Tetsuya Shimamura

Graduate School of Science and Engineering, Saitama University
Shimo-Okubo 255, Sakura-ku, Saitama 338-8570, Japan

may.t.t.218@ms.saitama-u.ac.jp, ysugiura@mail.saitama-u.ac.jp, shima@mail.saitama-u.ac.jp

ABSTRACT

In this paper, we propose an improved model of Shallow-UWnet for underwater image enhancement. In the proposed method, we enhance the learning process and solve the vanishing gradient problem by a skip connection, which concatenates the raw underwater image and the impulse response of low-pass filter (LPF) into Shallow-UWnet. Additionally, we integrate the simple, parameter-free attention module (SimAM) into each Convolution Block to enhance the visual quality of images. Performance evaluations with state-of-the-art methods show that the proposed method has comparable results on EUVP-Dark, UFO-120, and UIEB datasets. Moreover, the proposed model has fewer trainable parameters and the resulting faster testing time is suitable for real-time processing in underwater image enhancement, which is particularly for resource-constrained underwater robots.

Index Terms— Underwater image enhancement, CNN, Shallow-UWnet, impulse response of low-pass filter, SimAM attention module

1. INTRODUCTION

The decline in underwater image quality has constrained accurate visuals for diverse ocean engineering and scientific research, such as underwater object classification, saliency detection, monitoring the marine environment, and detecting underwater targets. While there has been some advancement in enhancing underwater image quality, serious degradation problems remain, and unfortunately, research in this area is infrequent [1]. When light travels through water, the red channel has more significant attenuation. Thus, the absorption rate of the red light is higher compared with the blue and green channels, resulting in underwater images with blue-green tones. Then, scattering and absorption are the main causes of underwater image deterioration. Light absorption deteriorates the color and contrast of images, while forward and backward scattering causes image blurring. Moreover, the scattering of light by tiny underwater particles, including suspended solids, plankton, and waterborne dust, reduces

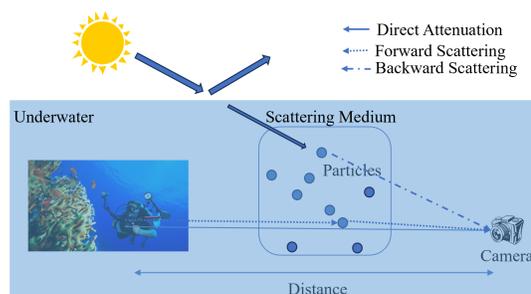


Fig. 1 Schematic diagram of underwater imaging

the image intensity. In an underwater environment, light encounters suspended particles within the water medium before reaching the camera. This can alter the direction of the light from the camera, leading to image blurring, resulting in low contrast and a fog-like effect [1]. Then, underwater images suffer from noise and distortion caused by water turbidity and low light conditions. Due to the impact of light scattering, absorption, and noise, underwater images have reduced contrast, color distortion, and blurred details. Fig. 1 illustrates the schematic representation of the underwater imaging. In accordance with the Jaffe-McGlamery model [2], underwater images can be modeled as

$$U_T = U_d + U_{fs} + U_{bs} \quad (1)$$

where U_T represents the final underwater image captured by the camera, U_d is the direct attenuation (light reflected by the object without scattering), and U_{fs} and U_{bs} correspond to the forward and backward scattering components, respectively [2].

2. EXISTING RESEARCH

Various approaches are utilized to enhance the visual clarity of underwater images, and they can be classified into three main groups: non-physical model-based methods, physical model-based methods, and data-driven methods. Non-physical model-based methods such as histogram equalization and filtering methods enhance the visual effect by

modifying the image pixels. However, these methods only concentrate on color correction and contrast. Thus, they excessively enhance the underwater images, resulting in structural detail damage and undesirable artifacts [3]. The physical model-based method utilizes prior knowledge to estimate the model parameters for achieving image enhancement. Dark channel prior [4], one line of research on physical model-based methods, incorporates wavelength-dependent algorithms. It utilizes the dark channel for estimating the transmission map in underwater images. However, these methods face challenges such as lack of adaptability, color distortion, and high computational cost, resulting in unsatisfactory outcomes when there is a mismatch between prior and target scenes.

Data-driven methods, trained on synthetic pairs of degraded and corresponding high-quality images, are successful in low-level vision tasks because they can adapt to various image conditions and capture complex patterns. WaterNet [5] constructs the underwater image enhancement benchmark (UIEB) dataset with real underwater images and corresponding reference images and encourages deep learning-based underwater image enhancement. Then, experimental results show that WaterNet outperforms state-of-the-art (SOTA) methods in non-physical and physical-based methods. Nevertheless, the complete removal of backscatter remains in some conditions. The generative adversarial network (GAN)-based underwater image enhancement model, FUnIE-GAN [6], is adversarially trained on the enhancement of underwater visual perception (EUVP) dataset using a non-linear mapping between the degraded and enhanced images. It improves underwater image quality and enhances image details. However, the resulting images have undesirable oversaturation and undersaturation of color. Furthermore, the inability to precisely restore pixel intensity often results in poor texture and color recovery. Fabbri et al. proposed the underwater generative adversarial network (UGAN) [7] to enhance the color in underwater images. However, the generation of paired distorted images by CycleGAN within the underwater dataset causes inaccuracies in the underwater enhancement learning model, which is common for GAN-based approaches. Then, the common issue, carefully tuning hyperparameters in GANs, remains. Additionally, underwater images face a challenge due to the lack of standardized and high-quality reference images. To solve this problem, simultaneous improvement and super-resolution (Deep SESR) [8], is introduced to the UFO-120 dataset, the first dataset for super-resolution and enhancement of underwater imagery. Deep SESR employs a feature extraction network that includes residual dense blocks and an auxiliary attention network. While Deep SESR performs better than existing methods, the large number of parameters makes it unsuitable for resource-constrained underwater vehicles.

In 2023, underwater image enhancement and turbidity removal (iDehaze) [9] is proposed using a two-step deep

learning approach. In iDehaze, input underwater images are first dehazed using a dehazing model and subsequently transformed into the color model. iDehaze outperforms underwater image quality measure (UIQM) scores; however, peak-signal-to-noise ratio (PSNR) scores are less favorable. Then, the visual results of iDehaze have an excessively processed appearance when observed by the human eye. The multichannel deep convolutional neural network linked to the VGG network (MDCNN-VGG) [10] improves domain adaptation for multi-domain underwater images. It consists of two network streams, followed by three fully connected layers, and passes the information to the VGG network to enhance underwater images. Thus, MDCNN-VGG has a complex deep architecture but still encounters problems in preserving fine details in blurred images, color bias and overexposure. Nowadays, autonomous underwater vehicles (AUVs) and remotely operated vehicles (ROVs) play a significant role in underwater explorations to capture real-time underwater images. However, the existing SOTA methods are computationally demanding, require significant memory resources, and are unsuitable for portable AUVs. Shallow-UWnet [11], a lightweight and compressed model for underwater image enhancement, emerges as an option for AUVs. Shallow-UWnet has 18 times fewer trainable parameters and 10 times faster computational time compared with SOTA methods. With the lowest testing time, Shallow-UWnet has better-enhanced underwater image ability and color correction. Despite improvements, Shallow-UWnet has some limitations, such as noise artifacts and color spots with a reddish hue in heavy hazy areas. Then, it struggles to distinguish between image information and noise in underwater images. The improved version of Shallow-UWnet [12] incorporates batch normalization and replaces the Leaky-ReLU function with the rectified linear unit (ReLU) function in Shallow-UWnet. Additionally, an extra structural similarity (SSIM) loss is added, which increases computational resources and training time. Consequently, the improved Shallow-UWnet has a larger number of parameters compared with the conventional Shallow-UWnet. Then, it does not have an improvement in the UFO-120 dataset when compared with the conventional Shallow-UWnet. In 2023, Zhou et al. proposed a Shallow-RepNet [13] by incorporating RepBlocks and Shallow-UWnet. Furthermore, Shallow-RepNet also introduced an area contrast distribution loss to optimize the difference between the output and reference image. However, Shallow-RepNet has lower PSNR and SSIM scores compared with Shallow-UWnet.

3. PROPOSED METHOD

3.1. Architecture and learning

Fig. 2 illustrates the architecture of the conventional Shallow-UWnet and the proposed method. In Shallow-UWnet, the first convolutional (Conv) layer utilizes 256×256 RGB underwa-

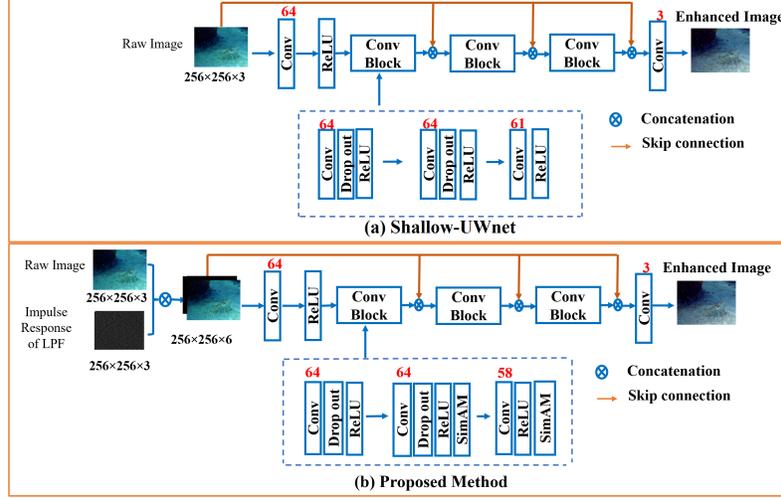


Fig. 2 Architectures of (a) conventional Shallow-UWnet and (b) proposed method

ter images with 64 feature maps and a kernel size of 3×3 . Subsequently, a ReLU activation is applied, and three successive convolutional blocks (ConvBlocks) are connected to the raw image via a skip connection. Each ConvBlock comprises a Conv layer and ReLU activation function paired with a drop-out regularization technique. Finally, it passes through the final Conv layer and generates the enhanced underwater image as depicted in Fig. 2 (a). The mean square error (MSE) loss and VGG perceptual loss functions are combined to preserve the sharpness of edges and structural similarity. The MSE loss is computed as

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (I_{GT} - I_{EN})^2 \quad (2)$$

The VGG perceptual loss is generated from 19 layers of the pre-trained VGG model, which measures the Euclidean distance between the ground truth and enhanced images as

$$L_{VGG} = \frac{1}{n} \sum_{i=1}^n \|\phi(I_{GT})_i - \phi(I_{EN})_i\|^2 \quad (3)$$

where ϕ refers to the feature representation function of the VGG model, n represents the number of pixels, and I_{GT} and I_{EN} refer to the ground truth image and its respective enhanced image, respectively [11]. Finally, the total loss function is computed as

$$L_{total} = L_{MSE} + L_{VGG} \quad (4)$$

3.2. Impulse response of LPF

The sequential input of a noisy image and the impulse response of LPF in a convolutional neural network (CNN) improved image denoising performance in [14] [15]. Motivated by this fact, the proposed method sequentially inputs the raw underwater image and the impulse response of

LPF via skip connection, aiming at enhancing the underwater image's visual quality. The impulse response of the sparsity-based LPF (SLPF) is created by the following steps. Initially, the power spectrum value $P(\omega_1, \omega_2)$ is determined by $P(\omega_1, \omega_2) = |X(\omega_1, \omega_2)|^2$, where $X(\omega_1, \omega_2)$ represents the Fourier transform of the underwater image. Subsequently, the power spectrum sparsity S is calculated as

$$S = \frac{P_a}{P_h + P_v} \quad (5)$$

where P_a represents the overall power spectrum value, while P_h and P_v indicate the horizontal and vertical power spectrum values at the center, respectively, as shown in Fig. 3. Then, the threshold value γ is formulated as

$$\gamma = \lambda S \quad (6)$$

The scaling parameter λ is specified based on the image size. Following this, the predominant portion of the image, focused at the center, is assigned to the value of one according to the γ value, while the remaining portion is set to zero. Then, the frequency response of the SLPF, $H_S(\omega_1, \omega_2)$, is designed by

$$H_S(\omega_1, \omega_2) = \begin{cases} 1, & \text{if } P(\omega_1, \omega_2) \leq \gamma \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

Subsequently, the frequency axis is reverted to the four corners. Finally, the impulse response of SLPF is obtained by the inverse Fourier transform of the shifted $H_S(\omega_1, \omega_2)$ version, thereby converting it to a spatial domain image.

Following this, the proposed method employs three well-known impulse response of LPFs such as direct LPF (DLPF), Gaussian LPF (GLPF), and Butterworth LPF (BLPF) for comparison purposes. Firstly, specify the size of the impulse response of LPF based on the image size. Subsequently, the corresponding frequency response of the LPF is applied.

For the DLPF,

$$H_D(\omega_1, \omega_2) = \begin{cases} 1, & \sqrt{(\omega_1^2 + \omega_2^2)} \leq \omega_c \\ 0, & \sqrt{(\omega_1^2 + \omega_2^2)} > \omega_c \end{cases} \quad (8)$$

For the GLPF,

$$H_G(\omega_1, \omega_2) = e^{-(\omega_1^2 + \omega_2^2)/2\omega_c^2} \quad (9)$$

For the BLPF,

$$H_B(\omega_1, \omega_2) = \frac{1}{1 + \left[\sqrt{(\omega_1^2 + \omega_2^2)}/\omega_c \right]^{2k}} \quad (10)$$

where ω_c is cut-off angular frequency and set to $\pi/2$, and k denotes the filter order with a value of 1. Finally, the impulse response of the LPF is determined by the inverse Fourier transform, which converts it into a spatial domain image for each scenario. After creating the impulse response of LPF, a skip connection is utilized by fusing the raw underwater image and the impulse response of LPF within each ConvBlock. Thus, the difference between Shallow-UWnet and the proposed method is the number of features in the last Conv layer of each ConvBlock. In detail, Shallow-UWnet employs 61 features, whereas the proposed method uses 58 features. SimAM with limited image datasets impacts the generalization ability. As shown in Fig. 2(b), the proposed method also incorporates the SimAM into each of the ConvBlocks, thereby providing a powerful feature extraction ability.

3.3. SimAM attention module

Currently, the attention module can easily enhance the backbone network's ability to highlight valuable information while suppressing irrelevant information. Conventional attention modules such as squeeze and excitation (SE), coordinate attention (CA) and efficient channel attention (ECA) require an additional subnetwork structure. Furthermore, current attention modules are limited in either channel or spatial dimensions and built with complex factors such as pooling, which can increase the network's computational complexity. To tackle these issues, SimAM [16], a non-parametric attention mechanism, was introduced in 2021. SimAM is an energy-based attention mechanism capable of generating 3D weights. According to neuroscience theory, the lower energy neuron, which is the most separable from neighboring neurons, is the most significant. The minimum energy neuron ϵ_T is calculated as

$$\epsilon_T = \frac{4(\rho^2 + \alpha)}{(T - \eta)^2 + 2\rho^2 + 2\alpha} \quad (11)$$

where T is the target neuron and ϵ_T represents the lower energy of neuron T , η is determined as the mean of neurons, expressed as $\eta = \frac{1}{N} \sum_{i=1}^N y_i$, and ρ^2 is the variance of neurons, computed as $\rho^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \eta)^2$. The y_i denotes

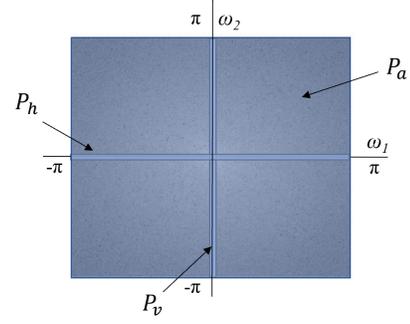


Fig. 3 Power spectrum sparsity of image

other neurons within a single channel. Then, i denotes the spatial dimension, N represents the total number of neurons, and α is a hyper-parameter coefficient set to 0.0001 [16]. Finally, the SimAM can be compromised as

$$\tilde{Y} = \text{sigmoid}\left(\frac{1}{E}\right) \odot Y \quad (12)$$

where the total energy function of ϵ_T is denoted as E and \odot indicates dot product operation and \tilde{Y} and Y represent the refined and original features, respectively [16].

4. EXPERIMENTAL RESULTS

4.1. Implementation details

The proposed model is trained using the ADAM optimizer with a learning rate of 0.0002 and layer dropout of 0.2, a batch size of 1, and 50 epochs, which is conducted in the same configuration as Shallow-UWnet. The input images are resized to 256×256 . The model employs the PyTorch framework and is trained on an Intel Core i9 CPU, Nvidia GeForce RTX 4070, and 32GB of RAM.

4.2. Datasets

Currently, there are available paired real-world underwater image datasets [17][18] like EUVP [6], UFO-120 [8], and UIEB [5]. The images in the EUVP dataset were taken with seven distinct cameras during deep-sea exploration and human-robot studies [19]. In this experiment, we utilized 3500 pairs of images from EUVP-ImageNet [6] for the training, while the remaining 200 image pairs were used for validation. The testing datasets are as follows. EUVP-Dark [6] dataset includes 5,500 paired images that capture dark-hazed underwater scenes. For testing, 1000 images were used in accordance with Shallow-UWnet model. UFO-120 dataset captures high-quality images during oceanic explorations. Distorted images in UFO-120 dataset were created using style transfer, and 120 paired images were utilized as a benchmark

Table 1 Quantitative comparisons (PSNR, SSIM, and UIQM) on EUVP-Dark, UFO-120, and UIEB datasets [Bold: Best, Underline: Second best]

Method	Dataset								
	EUVP-Dark			UFO-120			UIEB		
	PSNR	SSIM	UIQM	PSNR	SSIM	UIQM	PSNR	SSIM	UIQM
WaterNet[5]	24.43±4.6	0.82±0.08	2.97±0.32	23.12±3.3	0.73±0.07	2.94±0.38	19.11±3.7	<u>0.79±0.09</u>	<u>3.02±0.34</u>
FUnIE-GAN [6]	26.19±2.9	0.82±0.08	2.84±0.45	24.72±2.6	<u>0.74±0.06</u>	2.88±0.41	19.13±3.9	0.73±0.11	2.99±0.39
UGAN [7]	26.53±3.1	0.80±0.05	2.89±0.43	24.23±3.0	0.69±0.07	2.54±0.45	-	-	-
DeepSESR [8]	25.30±2.6	0.81±0.07	2.95±0.32	26.46±3.1	0.78±0.07	<u>2.98±0.37</u>	19.26±3.6	0.73±0.11	2.95±0.39
iDehaze [9]	23.01±2.0	0.84±0.09	3.11±0.36	17.55±1.9	0.72±0.07	3.29±0.26	17.96±2.8	0.80±0.07	3.28±0.33
MDCNN-VGG [10]	27.49	0.82	<u>3.0</u>	<u>25.27</u>	0.74	2.88	19.09	0.75	2.80
Xing et.al [12]	33.45±4.2	0.89±0.09	2.98±0.37	24.35±3.0	0.72±0.08	2.85±0.37	<u>19.71±4.0</u>	0.71±0.13	2.71±0.45
Shallow-RepNet [13]	24.49±2.5	0.79±0.06	2.82±0.29	22.32±2.4	0.72±0.07	2.98±0.33	19.80±2.8	0.77±0.08	2.79±0.32
Shallow-UWnet [11]	27.86±3.1	<u>0.85±0.04</u>	2.93±0.40	25.07±2.9	<u>0.74±0.08</u>	2.87±0.39	19.01±3.6	0.68±0.14	2.79±0.44
Proposed (SLPF)	27.87±3.0	0.84±0.05	2.96±0.36	25.27±2.8	0.73±0.08	2.90±0.36	19.14±3.7	0.69±0.13	2.84±0.41
Proposed (DLPF)	<u>27.89±3.1</u>	0.84±0.05	2.98±0.35	25.23±2.9	0.73±0.08	2.91±0.36	19.17±3.6	0.69±0.13	2.85±0.41
Proposed (GLPF)	27.87±3.0	<u>0.85±0.05</u>	2.95±0.37	25.25±2.9	<u>0.74±0.08</u>	2.89±0.37	19.08±3.6	0.69±0.13	2.82±0.42
Proposed (BLPF)	27.77±3.0	0.84±0.05	2.96±0.35	25.22±2.9	0.73±0.08	2.90±0.36	19.10±3.6	0.68±0.13	2.83±0.41

Table 2 Performance metrics of model lightweight [Bold: Best, Underline: Second best]

Metrics	Number of parameters	Testing per image (sec)
WaterNet [5]	1,090,688	0.5
FUnIE-GAN [6]	4,212,707	0.18
Deep SESR [8]	2,454,023	0.16
Xing et.al [12]	219,840	0.02
Shallow-UWnet [11]	<u>219,456</u>	<u>0.04</u>
Proposed (SLPF)	216,000	0.05
Proposed (DLPF)	216,000	0.2
Proposed (GLPF)	216,000	0.3
Proposed (BLPF)	216,000	0.3

to evaluate testing datasets. The UIEB dataset includes a collection of 890 real underwater images. The dataset comprises a variety of distortion levels and different light conditions, with a range of colors and contrast levels [20]. The reference images in UIEB dataset are free from color casts and display accurate colors.

4.3. Evaluation metrics

PSNR and SSIM evaluate the difference between underwater and reference images. Since obtaining perfect and ideal underwater images for evaluating underwater images is impractical, a non-reference image quality measure becomes necessary. The UIQM [17] is a non-reference underwater image quality measure inspired by the human visual system, which is defined by

$$UIQM = c_1 \times UICM + c_2 \times UISM + c_3 \times UICoM \quad (13)$$

where $UICM$, $UISM$, and $UICoM$ represent image colorfulness, sharpness, and contrast, respectively, and the pa-

rameters are set to $c_1 = 0.028$, $c_2 = 0.2953$, and $c_3 = 3.5753$ [17] [21]. A higher UIQM value indicates that the resulting images have high color saturation and contrast and are closely similar to human visual perception [22][23].

4.4. Experiment evaluation

The performance of the proposed method is compared with that of the conventional Shallow-UWnet [11] and SOTA models such as WaterNet [5], FUnIE-GAN [6], UGAN [7], DeepSESR [8], iDehaze [9], MDCNN-VGG [10], Xing et.al [12] and Shallow-RepNet [13], as shown in Table 1. It is observed that, in comparison with Shallow-UWnet, the proposed method achieves superior PSNR and UIQM results. This indicates that the proposed method effectively distinguishes image information from noise in underwater images compared with Shallow-UWnet. Consequently, the proposed method has improved image clarity and provided more accurate colors than the Shallow-UWnet. In comparison to Xing et al. [12], the proposed method outperforms in terms of UIQM on UFO-120 and UIEB datasets.

In more detail in Table 1, Xing et al. [12], Shallow-UWnet [11], and the proposed method, when trained on the EUVP dataset, have superior performance compared with other models on the EUVP-Dark dataset. Similar patterns are observed in DeepSESR [8], when trained on the UFO-120 dataset, outperforms the UFO-120 dataset case, and WaterNet [5], trained on the UIEB dataset, has better performance on the UIEB dataset. However, it is observed that the proposed method achieves comparable performance results on the UFO-120 and UIEB datasets, even though it was not trained on them. Furthermore, it is the second-best option in terms of PSNR and SSIM results on the EUVP-Dark and UFO-120 datasets. Then, the proposed method performs better than the GAN-based models such as FUnIE-GAN [6]

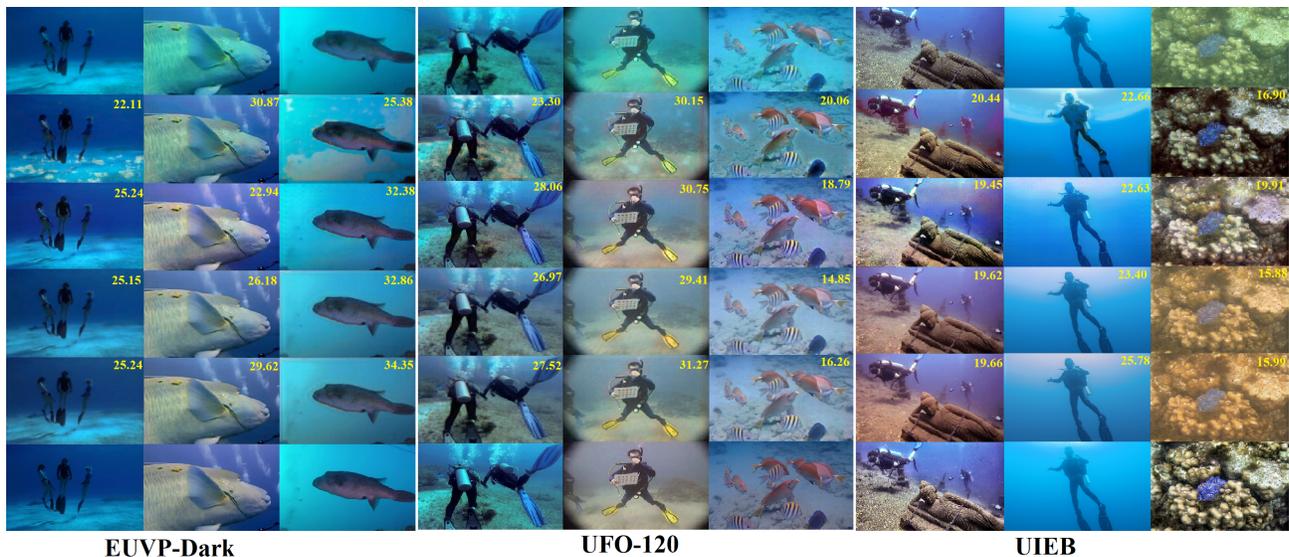


Fig. 4 Comparison of different methods on the EUVP-Dark, UFO-120, and UIEB datasets [from top to bottom] Raw Input Image, WaterNet, FUNIE-GAN, Shallow-UWnet, Proposed method (SLPF), and Ground Truth

and UGAN [7] on the EUVP-Dark and UFO-120 datasets. The two-step pipeline model, iDehaze [9], achieves superior UIQM scores, but it has the lowest PSNR results in comparison with others. Finally, although the proposed method does not perform superior to others, it achieves comparable results across all three datasets, as shown in Table 1. As shown in Table 2, the proposed method has fewer trainable parameters when evaluating the model's performance than others. Consequently, the proposed method with a smaller number of trainable parameters is suitable for underwater AUV and ROV devices with limited computing power. Other SOTA methods, such as iDehaze [9] and MDCNN-VGG [10], do not highlight their efficiency in terms of being lightweight. Undoubtedly, the deeper architectures in iDehaze and MDCNN-VGG result in a larger number of parameters compared with others. Regarding testing time, Xing et al. [12] outperforms other methods. Nevertheless, the proposed method, utilizing the impulse response of SLPF, is also suitable for real-time underwater image enhancement.

In Fig. 4, a visual comparison of the proposed method with the impulse response of SLPF, WaterNet [5], FUNIE-GAN [6], and Shallow-UWnet [11] is shown where in each image the resulting PSNR is also shown. In each dataset, the raw input images and reference images for ground truth are provided for visual comparison. We have the following observations.

(1) EUVP-Dark Dataset : WaterNet has artificial colors and noise artifacts in both blue and green-hued images compared with the proposed method. In comparison with the FUNIE-GAN, there are incorrect color corrections, especially noticeable in the fin of the fish image. Conversely, Shallow-UWnet and the proposed method have a noticeable effect on

restoring color, contrast improvement, and image sharpening.

(2) UFO-120 Dataset: WaterNet suffers from pixel information loss and the inability to rectify the green-colored hue in images, whereas FUNIE-GAN encounters oversaturation issues. In contrast, Shallow-UWnet and the proposed method perform well in enhancing images compared with others.

(3) UIEB Dataset: The results from WaterNet have noise artifacts and gray-tone color images on the UIEB dataset. While FUNIE-GAN has some effectiveness in color restoration, noise artifacts from scattering problems remain. In both Shallow-UWnet and the proposed method, there is a correction of the color cast to the ground truth image, resulting in superior visual effects compared with others, except for hazy images. In Shallow-UWnet and the proposed method, there is overcontrast with a reddish hue in heavy hazy regions. Nevertheless, the proposed method enhances color and higher PSNR values compared with Shallow-UWnet.

5. CONCLUSION

In this paper, we proposed a lightweight and compressed model for underwater image enhancement. The proposed method integrates the SimAM and the skip connection, which incorporates both the raw underwater image and the impulse response of LPF in the conventional Shallow-UWnet. It achieves better generalization ability for unseen features and enhances performance compared to using either approach separately. Subsequently, the proposed method outperforms the conventional Shallow-UWnet in terms of PSNR and UIQM metrics, with fewer trainable parameters and faster testing time. It could be employed in real-time applications, specifically for resource-constrained underwater robots.

6. REFERENCES

- [1] W. Zhang, L. Dong, X. Pan, P. Zou, L. Qin, and W. Xu, "A survey of restoration and enhancement for underwater images," *IEEE Access*, vol. 7, pp. 182259–182279, 2019.
- [2] J.S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE J. Ocean. Eng.*, vol. 15, no. 2, pp. 101–111, 1990.
- [3] J. Zhou, D. Zhang, and W. Zhang, "Underwater image enhancement method via multi-feature prior fusion," *Applied Intelligence*, vol. 52, no. 14, pp. 16435–16457, 2022.
- [4] Y. Ueki and M. Ikehara, "Underwater image enhancement based on the iteration of a generalization of dark channel prior," *Proc. IEEE Visual Communications and Image Processing*, 2019, vol. 52, pp. 1–4.
- [5] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao, "An underwater image enhancement benchmark dataset and beyond," *IEEE Trans. Image Processing*, vol. 29, pp. 4376–4389, 2019.
- [6] 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.
- [7] C. Fabbri, M.J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," *Proc. IEEE International Conference on Robotics and Automation*, 2018, pp. 7159–7165.
- [8] M.J. Islam, Y. Xia, and J. Sattar, "Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception," *Proc. of Robotics: Science and Systems*, 2020.
- [9] M. Mousavi, R. Estrada, and A. Ashok, "iDehaze: Supervised underwater image enhancement and dehazing via physically accurate photorealistic simulations," *Electronics*, vol. 12, no. 11, pp. 1–13, 2023.
- [10] L. Zhao, S.W. Lee, "Multi-domain rapid enhancement networks for underwater images," *Sensors*, vol. 23, no. 21, pp. 1–18, 2023.
- [11] A. Naik, A. Swarnakar, and K. Mittal, "Shallow-UWnet: Compressed model for underwater image enhancement," *Proc. AAAI Conference on Artificial Intelligence*, 2021, vol. 35, no. 18, pp. 15853–15854.
- [12] Z. Xing, M. Cai, and J. Li, "Improved Shallow-UWnet for underwater image enhancement," *Proc. IEEE International Conference on Unmanned Systems*, 2022, pp. 1191–1196.
- [13] J. Zhou, J. Zhuang, Y. Zheng, and J. Li, "Area contrast distribution loss for underwater image enhancement," *J. Marine Science and Engineering*, vol. 11, no. 5, 2023.
- [14] M.T. Tun, Y. Sugiura, and T. Shimamura, "Image denoising by incorporating noisy image patch and impulse response of low-pass filter in CNN learning," *Proc. IEEE 12th Global Conference on Consumer Electronics*, 2023, pp. 721–722.
- [15] M.T. Tun, Y. Sugiura, and T. Shimamura, "Joint training of noisy image patch and impulse response of low-pass filter in CNN for image denoising," *Journal of Signal Processing*, vol. 28, no. 1, pp. 1–17, 2024.
- [16] L. Yang, R. Y. Zhang, L. Li, and X. Xie, "Simam: A simple, parameter-free attention module for convolutional neural networks," *Proc. International Conference on Machine Learning*, 2021, pp. 11863–11874.
- [17] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE J. Ocean. Eng.*, vol. 41, no. 3, pp. 541–551, 2015.
- [18] C. Li, S. Anwar, and F. Porikli, "Underwater scene prior inspired deep underwater image and video enhancement," *Pattern Recognition*, vol. 98, pp. 1–11, 2020.
- [19] H. Zhao, and H. Yuan, "Residual dense blocks and contrastive regularization integrated underwater image enhancement network," *IEEE Access*, 2023.
- [20] G. Hou, X. Zhao, Z. Pan, H. Yang, L. Tan, and J. Li, "Benchmarking underwater image enhancement and restoration, and beyond," *IEEE Access*, vol. 8, pp. 122078–122091, 2020.
- [21] Y. Guo, H. Li, and P. Zhuang, "Underwater image enhancement using a multiscale dense generative adversarial network," *IEEE J. of Ocean. Eng.*, vol. 45, no. 3, pp. 862–870, 2019.
- [22] P. Zhuang, J. Wu, F. Porikli, and C. Li, "Underwater image enhancement with hyper-laplacian reflectance priors," *IEEE Trans. on Image Processing*, vol. 31, pp. 5442–5455, 2022.
- [23] S. Huang, K. Wang, H. Liu, J. Chen, and Y. Li, "Contrastive semi-supervised learning for underwater image restoration via reliable bank," *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition*, 2023, pp. 18145–18155.