Deep2.6GNet: A Lightweight 2.6 G FLOPs Architecture Model for Underwater Image Enhancement

May Thet Tun [1,*] and Tetsuya Shimamura [2]

[1] Department of Engineering, Computer University (Pathein), Pathein, Myanmar
[2] Department of Information and Computer Sciences, Saitama University, Saitama, Japan
Email: maythettun23@gmail.com (M.T.T.); shima@mail.saitama-u.ac.jp (T.S.)
*Corresponding author

Abstract—**High-resolution underwater imaging with rich color details and real time performance in underwater robots is crucial in Underwater Image Enhancement (UIE). However, light absorption, deep-sea, haziness, scattering, and suspended particles degrade underwater image quality. It is leading to low contrast and color distortion, which negatively affect visual tasks. In UIE, deploying models on mobile platforms and embedded devices has challenges due to their limited computational capacity. Therefore, a lightweight and computationally efficient model is required to enhance underwater images in resource-constrained underwater robots. To overcome this issue, this paper introduces Deep2.6GNet, a lightweight U-shaped encoder-decoder architecture model. Unlike existing lightweight models, Deep2.6GNet proposes a novel integration of Depthwise Separable Convolution (DWSConv) with the simple Attention Module (SimAM). It significantly reduces Floating-point Operations Per second (FLOPs) and the number of parameters and computational demands in UIE. Moreover, experimental evaluations on the high Structural Similarity Index Measure (SSIM) and Underwater Image Quality Measure (UIQM) scores and competitive Peak-Signal-to-Noise (PSNR) values, confirmed that Deep2.6GNet offers an excellent balance between computational efficiency and enhancement quality. Deep2.6GNet outperforms other State-of-the-Art (SOTA) methods with only 2.6G FLOPs, 0.067M parameters and 0.01s for runtime on a computer with an NVIDIA GeForce MX230 GPU, Intel® Core™ i7 processor, and 24 GB of RAM. It is well-suited for underwater robots, particularly for tasks such as mapping, target detection, and navigation.**

Keywords—**underwater image enhancement, lightweight, computational efficiency, U-shaped encoder-decoder architecture, depthwise separable convolution, simple Attention Module (SimAM)**

I. Introduction

Underwater Image Enhancement (UIE) serves as an essential preprocessing stage for tasks such as detection of objects and estimation of depth in underwater environments. Underwater robots, such as Remotely Operated Vehicles (ROVs) and Autonomous Underwater Vehicles (AUVs), are capable of capturing images in dangerous deep-sea areas that are difficult for humans to access. However, underwater robots operate under strict power limitations due to their reliance on onboard batteries. In addition, these devices are constrained by both space and computational resources. It can be challenging to deploy Deep Learning (DL) models on these devices. Hence, optimizing computational efficiency is essential in UIE. In DL, Floating-point Operations Per second (FLOPs) are commonly used to evaluate the computational cost of a network. Minimizing the computational load can help prolong battery life in underwater robots. Models with higher FLOPs require more computational resources, which lead to greater power consumption and increased heat generation. Conversely, models with low FLOPs consume less power, can run on more affordable hardware, and produce less heat. Therefore, models with low FLOPs are well-suited for real-time UIE on underwater robotic platforms such as mapping, target detection and navigation.

With the rapid progress of DL, neural network models have a strong generalization capability in handling complex processes in a data-driven manner, achieving remarkable results. However, their heavy dependence on large datasets, extensive parameters and high FLOPs leads to high memory consumption. It makes unsuitable for real-time UIE tasks, particularly on ROVs and AUVs. Although advanced models such as Generative Adversarial Networks (GANs) and transformers can deal with image distortions including low contrast, color changes and blurring, they are computationally expensive and impractical for deployment in real-time scenarios. State-of-the-art GANs and transformers have high FLOPs and millions of parameters, and they are unsuitable for edge devices and embedded underwater robots. Their high FLOPs result in slow inference, which restricts them in real-time applications. Moreover, the high FLOPs of these models lead to extreme energy consumption, which is challenging for resource-constrained devices.

Manuscript received November 17, 2025; revised December 12, 2025; accepted January 22, 2025; published July 16, 2026.

doi: 10.18178/joig.14.4.543-550

The attention mechanism of transformers and quadratic complexity can increase FLOPs and the number of parameters and delay image preprocessing. Consequently, it can consume overhead memory and time when processing high-resolution underwater images. Such delays can negatively affect underwater object detection. Additionally, the limited size of underwater image datasets causes GANs and transformers to over-fit in model generalization ability. These constraints highlight the importance of designing lightweight models, which are prioritized over other visual quality metrics [1]. Therefore, efficient computational design and real-time processing capability, along with the lowest energy consumption, have emerged as a key concern and remain a research gap in UIE. In this paper, we propose a solution to reduce high computational costs and balance the model for UIE in resource-constrained devices.

The main contributions of this paper are as follows: (1) A lightweight U-shaped encoder-decoder architecture is proposed for UIE. (2) The incorporation of Depthwise Separable Convolution (DWSConv) with the Simple Attention Module (SimAM) effectively reduces computational overhead while maintaining enhancement quality. (3) The model with low FLOPs and parameter count is well-suited for resource constrained devices. (4) Experimental evaluations with high Structural Similarity Index (SSIM) and Underwater Image Quality Measure (UIQM) scores, along with competitive Peak Signal-to-Noise Ratio (PSNR) values, confirm that the proposed Deep2.6GNet achieves an optimal trade-off between computational efficiency and image quality.

This paper is structured as follows. Section II reviews related works on UIE and outlines the limitations of existing approaches. Section III describes the proposed methodology. Section IV presents the experimental setup, including dataset information, evaluation metrics and visual comparisons. Finally, Section V presents the conclusion of the study and future research directions.

II. Literature Review

Early research in UIE primarily relied on traditional techniques, such as histogram equalization, filtering, and retinex-based methods. These approaches are simple and easy to implement, but they often fail in complex underwater images, particularly in deep-sea environments [2]. In 2022, MLLE [3] enhanced color in underwater images by combining a minimum color loss with attenuation maps. However, it struggles to restore accurate colors and produces overly vivid results with noise or artifacts [3]. Physical model-based approaches, like UDCP [4], relied on the dark channel prior that may not accurately restore color under underwater conditions. As a result, UDCP often leads to over-enhancement that appears unnaturally vivid and inconsistent color [4].

With the innovation of DL, data-driven approaches have become mainstream, providing more reliable features for object and target detection in computer vision tasks. Convolutional Neural Networks (CNNs), known for their strong feature extraction, are commonly applied to UIE. In 2019, Li *et al.* [5] introduced WaterNet, a CNN-based gated fusion network trained on the Underwater Image Enhancement Benchmark (UIEB), which includes 890 pairs with real-world images along with their reference counterparts. However, WaterNet requires extensive computation and is less suitable for AUVs and ROVs. Then, it struggles to generalize in color cast, low contrast, and haze.

UWCNN [6], a residual learning-based CNN, provided more natural colors than traditional methods. Nevertheless, its dense connections and large number of parameters result in high memory and battery consumption, causing challenges for resource-constrained underwater robots. Shallow-UWNet [7], a lightweight neural network, has lower computational requirements and is well-suited for real-time processing. Despite of its efficiency, the model struggles to handle complex underwater conditions, particularly in cases of turbidity and restoring reddish color [7]. UICE^2-Net [8] is an end-to-end CNN that utilizes both Red, Green, and Blue (RGB) and Hue, Saturation, and Value (HSV) pixel-level blocks to improve overall image quality. However, it increases computational requirements and struggles with severe color degradation.

PRWNet [9] enhanced underwater images by combining spatial and frequency domain refinement using a wavelet-based boosting strategy. PUIE-Net [10] employed a Conditional Vibrational Auto Encoder (CVAE) integrated with adaptive instance normalization for UIE. However, it needs longer inference time, and its performance is dependent on datasets quality. In addition, the refinement process and wavelet decomposition cause more computational resources, and their performance declines under unseen or highly complex underwater conditions.

Ling *et al.* [11] proposed UHD-SFNet, a U-shaped transformer network integrated channel-wise multi-scale feature fusion and spatial-wise global feature modeling. DEHRFormer [12] was built on a transformer with shared encoder and decoder. It handles haze removal with better generalization. NU2Net [13] employed a convolutional attentional image transformer, effectively capturing both global and local degradations. While achieving high-quality enhancement, UHD-SFNet [11], DEHRFormer [12] and NU2Net [13] suffer from high computational cost due to their attention modules of the transformer. Therefore, the transformer's computational demand is difficult to use on underwater robots.

AUVs and ROVs are constrained by limited onboard computing capacity because their design balances the hardware, restricted power supply, and real-time processing [14]. Consequently, recent approaches have focused on designing lightweight models [14, 15]. Jiang *et al.* [15] proposed Five A$^+$, a lightweight network in which a robust prior stage is followed by a detailed refinement stage. Five A$^+$ limits to handle complex or unseen underwater image degradations. Yang *et al.* [14] proposed LU2Net, a U-shaped model with a channel Attention layer (CA layer) to improve image quality. It is fast and lightweight, but the CA layer focuses only on channel information, which can cause problems with spatial details and color accuracy. Even though the CA layer helps the model learn better features, it can be computationally expensive. As a result, LU2Net cannot restore colors well in dark or low-light images. To solve these issues observed in LU2Net [14], we incorporate the SimAM parameter free attention module into the U-Shaped encoder-decoder architecture. It enhances feature learning without increasing model complexity.

There exist DL-models for image segmentation and enhancement [16–23] and some recent review papers [24–26]. In Ref. [24], deep-sea environments are discussed. UIE methods for deep-sea environments differ from shallow-water, as artificial lighting has a strong impact on deep-sea images. Existing UIE models are not effective for deep-sea conditions. Therefore, it is emphasized that these models are needed to handle low brightness, light absorption and backscattering effects in deep-sea images [24]. On the other hand, it is described that future studies in UIE need to prioritize real-time processing and produce more natural and realistic underwater images in [25, 26]. In Ref. [27], the need of advanced DL techniques is discussed. Many existing learning-based methods depend on large datasets, which require a large number of paired underwater images. Therefore, a lightweight GAN based model is explored for UIE in Ref. [27].

III. METHODLOGIES

This section presents the proposed Deep2.6GNet architecture and its related components. It begins with the lightweight U-shaped encoder–decoder architecture, followed by DWSConv, and then the SimAM.

A. Lightweight U-Shaped Encoder-Decoder Architecture

In recent years, the U-Shaped encoder-decoder architecture, known as UNet, has been extensively applied in computer vision tasks to improve efficiency by reducing computational overhead. It is composed of an encoder and a decoder.

The encoder captures hierarchical low and high features through Convolution (Conv) and pooling operations and then the decoder reconstructs these features from the encoder using skip connections to preserve image details. It is widely used due to its lightweight architecture and significantly reducing parameters while preserving spatial details and feature learning in medical image segmentation and restoration [16].

UNeXt [16], the U-Shaped model with a Multilayer Perceptron (MLP) for image segmentation, can reduce parameters and computational cost. It improves network accuracy in medical image segmentation [16]. Inspired by this, the U-shaped architecture for UIE can help to reduce computation demands. Earlier UIE methods suffer from limited underwater image dataset sizes, inefficient color correction, and high computational complexity. The U-shaped encoder-decoder model works well with small datasets while achieving outstanding results [14, 16], as its architecture allows encoder features to flow directly to the decoder, enhancing learning efficiency. Motivated by this, we propose Deep2.6GNet, a U-shaped encoder-decoder architecture to minimize model parameters and computational costs while effectively enhancing underwater images.

B. Depthwise Separable Convolution (DWSConv)

Nowadays, DWSConv has emerged as an innovative design of CNNs that is effectively used to achieve computational efficiency and reduced model size. It is employed in well-known deep convolutional networks, including XceptionNet, EfficientNet and MobileNet.

In a conventional Conv, a kernel operates across all input channels, which results in a high computational result. DWSConv decomposes into two sequential operations. First, Depth Wise Convolution (DW Conv) applies separate filters to each input channel independently. Then, Point Wise Convolution (PW Conv) merges the outputs across channels [17].

This separation of filtering and integration not only decreases the number of learnable parameters but also reduces FLOPs, thereby improving computational efficiency and resulting in faster inference. It can perform 100 times fewer multiplications as compared to a standard Conv, which is suitable for mobile and resource-constrained devices. The two-step processes seem inefficient, however, in DW Conv, each individual filter operates to each input channel independently. Then, the PW Conv linearly combines the output results across all channels. This factorization makes the model lighter and less prone to over-fitting [17]. Algorithm 1 provides a detailed explanation of DWSConv operations.

Algorithm 1. Depth Wise Separable Convolution (DWSConv)

Input: A 3 dimensional tensor $\boldsymbol{I}$ with $\boldsymbol{n}$ channels
Output: A 3 dimensional tensor $\boldsymbol{O}$ with $\boldsymbol{n}$ channels
for $\boldsymbol{i = 0}$ **to** $\boldsymbol{n - 1}$ **do**
  $\boldsymbol{DW = DW\ Conv(I[i])}$
end for
  $\boldsymbol{O = PW\ Conv(DW)}$

C. SimAM

SimAM is adopted over other attention modules in the proposed Deep2.6GNet due to its parameter-free design, lightweight structure, and high computational efficiency, making it suitable for underwater robots. The SimAM is a simple yet effective mechanism that uses 3D attention weights. Other conventional attention modules refine features along spatial or channel dimensions and incorporate pooling layers in their structure. They can involve complex factors of the model and increasing parameter counts.

SimAM uses the energy function inspired by neuroscience principles. Consequently, SimAM does not require extra parameters to the network while achieving performance comparable to other attention modules [18]. The energy weight (w_i) can be calculated as follows:

$$w_i = \frac{(i-\mu)^2 + 2(\varphi+\vartheta)}{4(\varphi+\theta)} \tag{1}$$

where μ and φ are the mean and variance values, respectively, w_i represents the energy weight of the neuron i, and θ is the regularization coefficient. In this work, the symbol θ is assigned to the value 0.0001.

D. The Proposed Deep2.6GNet Architecture

The architecture of Deep2.6GNet proposed in this paper is shown in Fig. 1. Deep2.6GNet is a U-shaped based encoder-decoder architecture with DWSConv and SimAM attention module. It aims to enhance the model generalization efficiency by effectively capturing local and global image features for UIE. The first layer of the network is a Conv layer and it converts RGB channel input into 16 feature maps to expand the image's features. An encoder is used to increase channel depth and get informative features of each input image. Depthwise Separable Convolution Block (DW Separable Block) in the Deep2.6GNet helps richer feature learning and focuses on important image details. Each DW Separable Block in Deep2.6GNet consists of DW Conv and PW Conv, which is followed by Batch Normalization (BatchNorm), Gaussian Error Linear Unit (GELU) activation and SimAM attention module. Finally, a MaxPooling Layer is applied in the encoder to reduce the dimensionality by half, thereby accomplishing the feature extraction process.

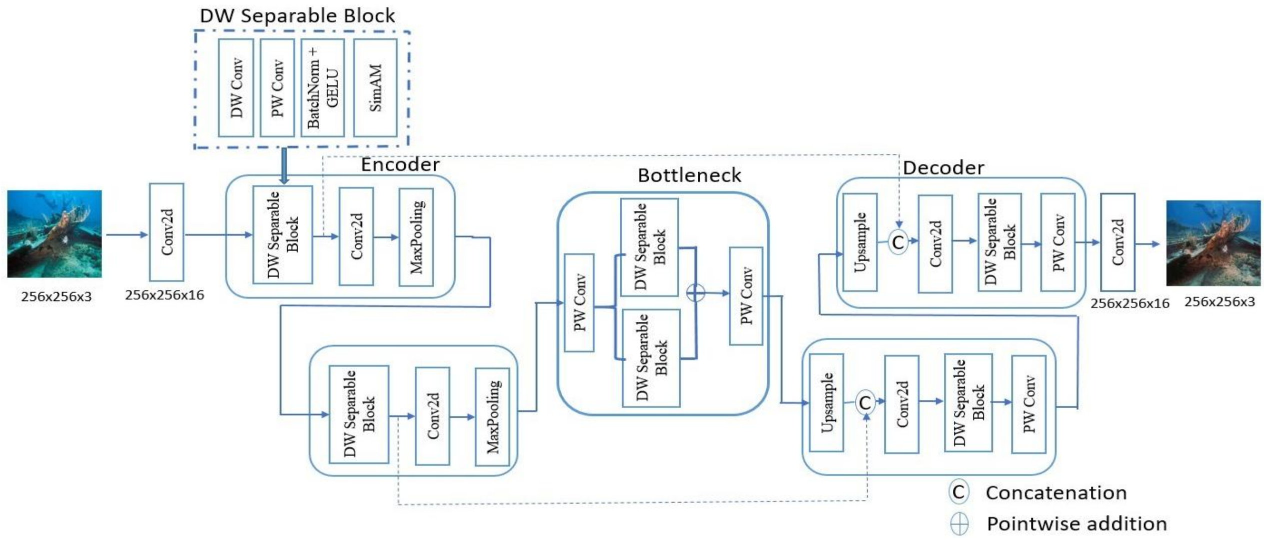


Fig. 1. Deep2.6GNet architecture.

The bottleneck block is employed to efficiently extract both fine local and global features. It consists of a dual-branch DW Separable block and is followed by a PW Conv layer to refine the extracted features. The dual-branch DW Separable block enables parallel learning of spatial patterns and contextual patterns simultaneously. Then, the PW Conv layer merges these two branches to get rich features with minimal parameter overhead. The decoder block reconstructs fine image details while integrating features from the encoder. It starts from an upsampling, after that skip-connected features by the corresponding encoder are concatenated to recover spatial information. Subsequently, a Depthwise Separable Block is applied to extract refined spatial features, followed by PW Conv to fuse channel-wise information and enhance feature representation. Finally, a Conv layer is applied to map the refined feature to a three-channel RGB image. The total loss function (L_{TL}) consists of five losses: pixel-wise loss, structural similarity loss, perceptual loss, Lab color space (Lab) loss and Lightness, Chroma, and Hue (LCH) color space loss as:

$$L_{TL} = 0.1L_{pixel} + 100\ L_{SSIM} + 100L_{VGG} \\ +0.001L_{lab} + 1L_{LCH} \quad (2)$$

The pixel-wise loss (L_{pixel}) calculates the mean squared error between images, focusing on low-level accuracy and detail preservation. The structural similarity loss (L_{SSIM}) evaluates perceptual similarity by comparing luminance, contrast, and structural features. The perceptual loss (L_{VGG}) compares high-level features extracted from a pre-trained VGG network. Additionally, the Lab loss (L_{Lab}) and LCH loss (L_{LCH}) measure color differences in perceptually color spaces.

IV. Experiments

The experimental evaluation of the Deep2.6GNet is organized into five subsections. Firstly, the experimental details and the datasets employed are described. Next, the image quality assessment metrics and the evaluation of performance are presented. Finally, the section concludes with visual comparisons.

A. Experimental Details

The Deep2.6GNet model was implemented in PyTorch and trained for 150 epochs. The training utilized paired input-output underwater images from a Large-Scale Underwater scene dataset (LSUI), and the model was optimized using the Adam optimizer with an initial learning rate of 0.0005 and momentum parameters of 0.5 and 0.999. The learning rate scheduler (StepLR) was applied, reducing the learning rate by a factor of 0.8 every 40 epochs. The loss function compiled pixel-wise loss, SSIM loss, perceptual (VGG) loss and color space losses (LAB and LCH). The input images were resized to 256×256 and normalized to the range of [−1, 1]. The experiments were performed on a computer with an NIVDA GeForce MX230, and an Intel ® core TM i7 model and 24 GB of RAM for training.

B. Datasets

The LSUI dataset comprises 4279 real underwater images along with their corresponding high-quality reference images, capturing diverse underwater conditions, including varying lighting, water types, and object categories [11, 14]. The dataset is partitioned into training and testing subsets using an 8:2 ratio, resulting in 3995 pairs for training and 284 pairs for testing. The UIEB dataset provides 890 high-quality underwater images along with their ground truth counterparts, while the UFO-120 dataset offers 120 paired images. Additionally, 60 particularly challenging images from UIEB are designated as the C60 dataset. The U45 dataset comprises 45 authentic images, divided into three subsets: blue, green, and haze according to the differences in color distortion, contrast, and blur. Both C60 and U45, which are non-reference datasets, are evaluated for the quality of the models [18].

C. Image Quality Assessment Metrics

Underwater image quality can be assessed using both subjective and objective methods. However, human-based subjective methods are impractical for underwater images, so objective metrics such as PSNR, SSIM, and non- reference measures like UIQM and Underwater Color Image Quality Evaluation (UCIQE) are designed to align. PSNR calculates the error between corresponding pixels of the reference and enhanced underwater images, and SSIM evaluates the images' brightness, contrast, and structural similarity. On the other hand, UCIQE and UIQM are no-reference metrics: UCIQE focuses on evaluating the detailed information of color in distorted images, while UIQM assesses perceptual quality based on color, sharpness, and contrast. PSNR, SSIM, UIQM and UCIQE are calculated as follows:

$$PSNR = 10log_{10}(\frac{{Max_p}^2}{mse}) \tag{3}$$

where Max_p refers the maximum pixel value and mse denotes the mean square error.

$$SSIM(x,y) = \frac{(2\beta_x\beta_y+c_1)(2\gamma_{xy}+c_2)}{(\beta_x^2+\beta_y^2+c_1)(\gamma_x^2+\gamma_y^2+c_2)} \tag{4}$$

where β and γ denote the average brightness and the contrast of images x and y, respectively, and c_1 and c_2 are small constants.

$$UIQM = \omega_1.uicm + \omega_2.uism + \omega_3.uiconm \tag{5}$$

where $uicm$, $uism$ and $uiconm$ correspond to the color, sharpness and contrast measure values of the underwater image, and $\omega_1, \omega_2, \omega_3$ are weighting factors.

$$UCIQE = W_1.\delta + W_2.l_{con} + W_3.\emptyset \tag{6}$$

where δ, l_{con} and $\emptyset$ denote to the standard deviation, luminance contrast and mean value saturation of the underwater image, respectively, and W_1, W_2, W_3 are fixed weights [7].

D. Evaluation Performance

To evaluate both UIE quality and computational efficiency of Deep2.6GNet, we compared our Deep2.6GNet with recent State-of-the-Art (SOTA) methods. An upward arrow represents that higher values indicate better performance, whereas a downward arrow represents that lower values indicate better performance in tables.

Table I presents the PSNR and SSIM results of various UIE techniques on the LSUI dataset. The conventional method, MLLE [3], gets the lowest performance and WaterNet [5], UWCNN [6] and U-Trans [19] show moderate results. Among the compared methods, LU2Net [14] attains the highest performance with 25.55 dB PSNR and 0.87 SSIM. The proposed Deep2.6GNet also performs competitively with a highest SSIM of 0.87, although its PSNR is slightly lower than LU2Net.

TABLE I. EVALUATION OF PSNR/SSIM ON THE LSUI DATASET (BEST RESULTS IN BOLD, SECOND-BEST UNDERLINED)

Method	PSNR	SSIM
MLLE [3]	18.13	0.73
WaterNet [5]	19.52	0.81
UWCNN [6]	22.62	0.83
U-Trans [19]	<u>25.03</u>	<u>0.84</u>
LU2Net [14]	**25.55**	**0.87**
Deep2.6GNet	24.42	**0.87**

Tables II and III show the performance comparison of PSNR, SSIM, and UIQM on the UIEB and UFO-120 datasets. Based on experiments conducted with the UIEB dataset, PUGAN [23] achieves the highest overall image quality in terms of PSNR and UIQM, while the proposed Deep2.6GNet has the highest SSIM, indicating strong structural preservation. Other methods, such as WaterNet [5], FUnIE-GAN [21], DeepSESR [22], and Shallow-UWNet [7], have moderate performance in both pixel-level fidelity and perceptual quality. On the UFO-120 dataset, DeepSESR [22] leads in PSNR, while PUGAN [23] performs best in SSIM. The proposed Deep2.6GNet has a balanced performance across PSNR and SSIM, achieving high UIQM scores effectively. Overall, Deep2.6GNet has competitive results on both the UIEB and UFO-120 datasets.

TABLE II. EVALUATION OF PSNR/SSIM/UIQM ON THE UIEB DATASET (BEST RESULTS IN BOLD, SECOND-BEST UNDERLINED)

Method	PSNR	SSIM	UIQM
WaterNet [5]	19.11	<u>0.79</u>	3.02
FUnIE-GAN [21]	19.13	0.73	2.99
DeepSESR [22]	19.26	0.73	2.95
Shallow-UWNet [7]	18.99	0.67	2.77
PUGAN [23]	**21.67**	0.78	**3.28**
Deep2.6GNet	<u>19.7</u>	**0.81**	<u>3.17</u>

TABLE III. EVALUATION OF PSNR/SSIM/UIQM ON THE UFO-120 DATASET (BEST RESULTS IN BOLD, SECOND-BEST UNDERLINED)

Method	PSNR	SSIM	UIQM
WaterNet [5]	23.12	0.73	<u>2.94</u>
FUnIE-GAN [21]	24.72	0.74	2.88
DeepSESR [22]	**26.46**	0.78	**2.98**
Shallow-UWNet [7]	<u>25.20</u>	0.73	2.85
PUGAN [23]	23.7	**0.82**	2.85
Deep2.6GNet	23.6	<u>0.81</u>	**2.98**

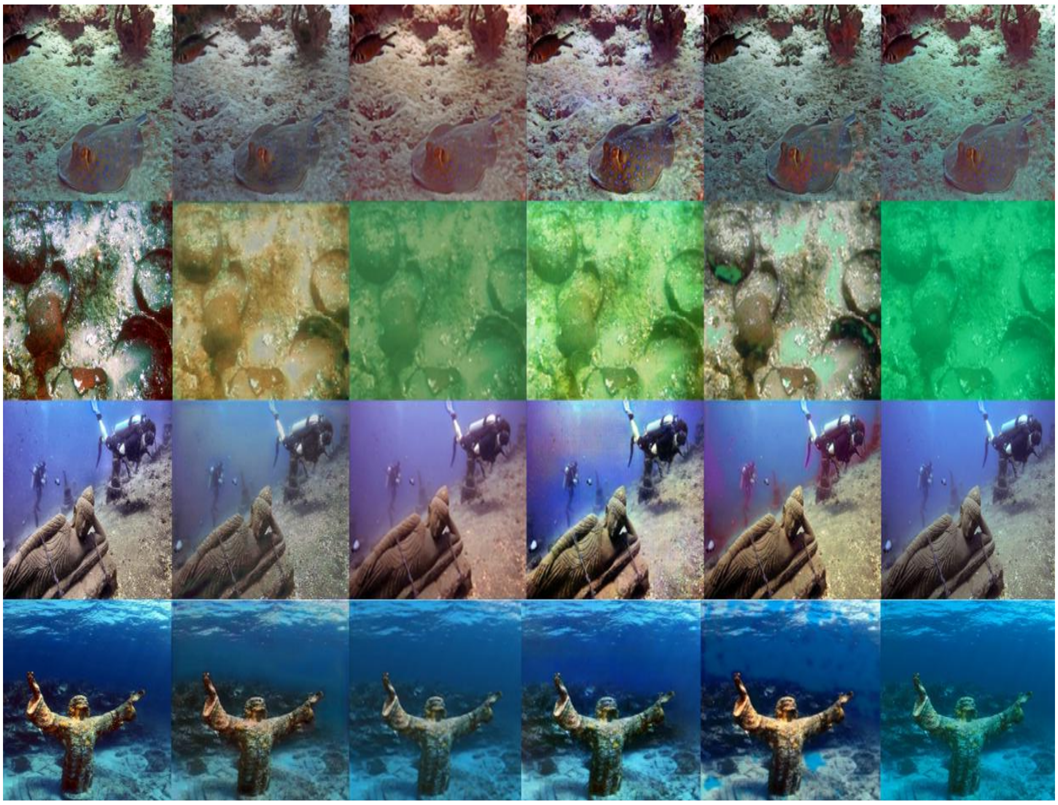
TABLE IV. EVALUATION OF UIQM/UCIQE ON THE C60 AND U45 DATASETS (BEST RESULTS IN BOLD, SECOND-BEST UNDERLINED)

Method	C60		U45	
	UIQM	UCIQE↑	UIQM↑	UCIQE↑
WaterNet [5]	2.113	0.550	2.957	0.576
UWCNN [6]	2.222	0.492	3.063	0.527
PRW-Net [9]	<u>2.717</u>	0.572	3.026	**0.625**
Shallow-UWNet [7]	2.212	0.521	3.109	0.545
Ucolor [20]	2.167	0.530	3.148	0.554
UIEc2Net [8]	2.228	<u>0.580</u>	3.125	0.604
MLLE [3]	1.977	0.581	2.454	0.597
UHD-SFNet [11]	1.741	0.528	2.826	0.585
PUIE-Net [10]	2.155	0.543	**3.192**	0.563
NU2Net [13]	2.222	0.555	<u>3.185</u>	0.593
Five A$^+$ [15]	2.088	**0.593**	3.174	<u>0.609</u>
Deep2.6GNet	**2.840**	0.544	3.176	0.56

In Table IV, the proposed Deep2.6GNet is evaluated with other SOTA models on the C60 and U45 datasets. The comparison includes CNN-based methods (WaterNet [5], UWCNN [6], PRW-Net [9], Shallow-UWNet [7], UIEC2Net [8], UHD-SFNet [11], PUIE-Net [10], NU2Net [14], and Five A$^+$ [15]), GAN-based approaches (FUnIE-GAN [21], DeepSESR [22], and PUGAN [23]), and transformer-based methods (U-Trans [19] and UColor [20]). Moreover, NU2Net [13], which is based on a Conv attentional image transformer framework, is also considered.

On the C60 dataset, Deep2.6GNet achieves the highest UIQM of 2.84, while its UCIQE of 0.544 is slightly lower than Five A$^+$. On the U45 dataset, Deep2.6GNet gets a UIQM score of 3.176 and a UCIQE of 0.56, showing competitive performance compared with recent SOTA methods. Overall, Deep2.6GNet has strong visual quality, achieving a balance between contrast enhancement, color restoration, and structural preservation compared with both conventional CNN- and GAN-based approaches.

Table V provides a comparative analysis of the computational efficiency of different UIE methods. The evaluation includes FLOPs, the number of parameters, and runtime performance [24]. Among the CNN- and GAN-based models, WaterNet [5] and PRW-Net [9] shows relatively high computational costs with FLOPs, respectively, while Ucolor [20] and PUIE-Net [10] are the most resource-intensive, demanding 443.8G and 423.1G FLOPs. Lightweight networks such as Five A$^+$ [15], LU2Net [14], and the proposed Deep2.6GNet achieve a significant reduction in computational complexity, with Deep2.6GNet requiring only 2.6G FLOPs, 0.067M parameters and 0.01 s for runtime performance. Using an NVIDIA GeForce MX230 GPU, Intel® Core™ i7 CPU, and 24 GB RAM, Deep2.6GNet operates at 20 frame per second. Therefore, Deep2.6GNet is extremely lightweight and highly computational efficient, making it suitable for resource-constrained underwater robots.



(a) (b) (c) (d) (e) (f)

Fig. 2. Comparison of visual results obtained by different approaches on the UIEB dataset (a) Raw Image, (b) WaterNet, (c) FUnIE_GAN, (d) Shallow_UWNet, (e) Deep2.6GNet, and (f) Ground Truth.

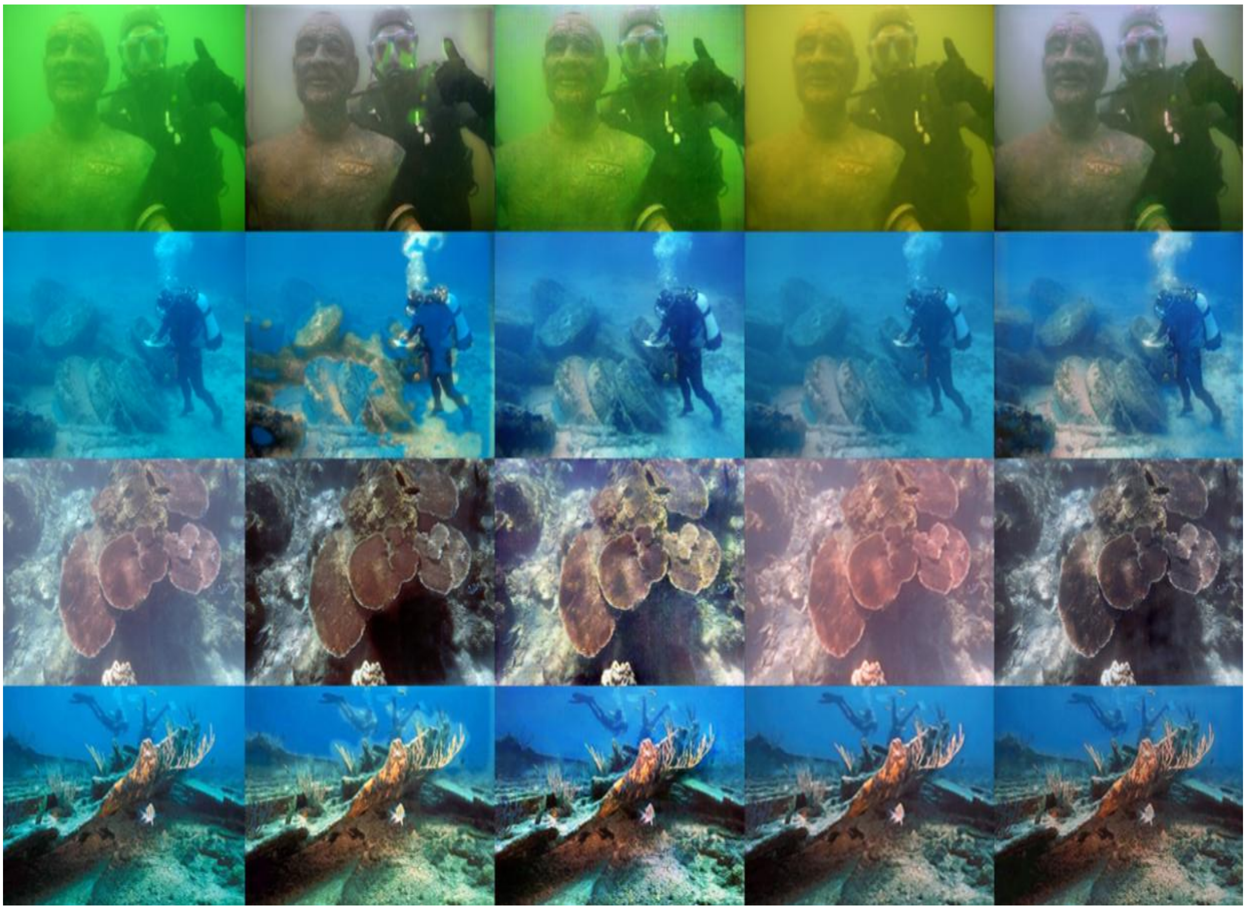
TABLE V. EFFICIENCY COMPARISON ON DIFFERENT MODELS (BOLD FOR THE BEST RESULT AND UNDERLINE FOR THE SECOND BEST RESULT)

Methods	FLOPs (G)	#Parameters	Runtime
WaterNet [5]	193.7 G	24.81 M	0.68
PRW-Net [9]	223.4 G	6.3 M	0.22
Shallow-UWNet [7]	30.75 G	0.22 M	0.03
FUnIE-GAN [21]	10.23 G	7.01 M	-
DeepSESR [22]	146.1 G	2.46 M	-
PUGAN [23]	72.05 G	95.66 M	-
Ucolor [20]	443.8 G	157.4 M	2.58
UIEC^{2}Net [8]	367.5 G	0.53 M	0.17
UHD-SFNet [11]	15.24 G	37.3 M	0.06
PUIE-Net [10]	423.1 G	1.4 M	0.07
NU2Net [13]	16.64 G	3.15 M	<u>0.02</u>
Five A$^+$ [15]	8.33 G	**0.009 M**	**0.01**
U-Trans [19]	60.2 G	22.8 M	-
LU2Net [14]	<u>2.8 G</u>	0.176 M	-
Deep2.6GNet	**2.6 G**	<u>0.067 M</u>	**0.01**

E. Visual Comparisons

Fig. 2 compares a visual comparison of UIE results on the UIEB dataset. The raw images have prominent color casts, low contrast, and poor visibility. WaterNet shows obvious noise and color distortions in the third image. FUnIE-GAN fails to effectively eliminate the green color cast in the second image, whereas Shallow-UWNet produces hazy outputs in the first and third images, as shown in Fig. 2. In contrast, the proposed Deep2.6GNet has more natural color correction, enhanced contrast, and structural details.

Fig. 3 illustrates the enhancement results on the U45 dataset. The raw underwater images suffer from color casts, low contrast, and reduced clarity. WaterNet improves contrast but shows visible noise and color artifacts in the first and second images in the complex textures. FUnIE-GAN struggles to correct the green color cast in the first image. Shallow-UWNet enhances visibility but fails to restore accurate green tones in the first image and over-boosts red color in the third image. On the other hand, Deep2.6GNet achieves more natural color restoration, clearer structural details, and contrast in comparison with other methods.



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

Fig. 3. Comparison of visual results obtained by different approaches on the U45 dataset (a) Raw Image, (b) WaterNet, (c) FUNIE_GAN, (d) Shallow_UWNet, and (e) Deep2.6GNet.

V. CONCLUSION

In this study, we proposed Deep2.6GNet, a lightweight U-shaped encoder-decoder architecture with DWSConv and SimAM attention module for UIE. Comprehensive experiments performed using multiple benchmark datasets, such as LSUI, UIEB, UFO-120, C60, and U45, and showed that Deep2.6GNet has the highest results in SSIM and UIQM, along with competitive results across PSNR and UCIQE metrics on SOTA models. Thus, Deep2.6GNet has the ability to enhance underwater images by improving contrast, restoring colors and preserving fine details across diverse underwater scenes. Moreover, Deep2.6GNet outperforms existing SOTA methods while significantly reducing FLOPs and the number of parameters to only about 2.6G and 0.067M. Overall, it achieves an excellent balance between computational efficiency and underwater image enhancement quality. Due to its lightweight design and effective performance results, Deep2.6GNet is well-suited for deployment in underwater devices with limited computational resources.

Here, we discuss potential directions for future research in UIE. Future work will focus on improving performance while maintaining model simplicity, such as using lightweight architectures, and enhancing the ability to handle realistic images. Current methods still face challenges like over-enhancement, which can result in extreme contrast or unnatural colors, so research will aim to improve images while preserving a realistic appearance. In addition, advanced DL models, including lightweight GANs, will be explored for UIE. Furthermore, many UIE methods emphasize visual enhancement but pay little attention to their impact on high-level tasks. Therefore, the relationship between low-level enhancement and high-level vision tasks requires further investigation.

Conflict of Interest

The authors declare that they have no conflict of interest.

Author Contributions

Tetsuya Shimamura provided conceptual guidance, supervised the research work, and contributed to and performed a critical review of the manuscript. May Thet Tun conducted the research, and prepared the initial draft. All authors have read and approved the final version of the manuscript.

References

[1] Y. Guan, X. Liu, Z. Yu, Y. Wang, X. Zheng, S. Zhang, and B. Zheng, "Fast underwater image enhancement based on a generative adversarial framework," *Front. Mar. Sci.*, vol. 9, 964600, 2023.

[2] S. Zhang, T. Wang, J. Dong, and H. Yu, "Underwater image enhancement via extended multi-scale Retinex," *Neurocomputing*, vol. 245, pp. 1–9, 2017.

[3] W. Zhang, P. Zhuang, H. H. Sun, G. Li, S. Kwong, and C. Li, "Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement," *IEEE Trans. Image Process.*, vol. 31, pp. 3997–4010, 2022.

[4] H. Y. Yang, P. Y. Chen, C. C. Huang, Y. Z. Zhuang, and Y. H. Shiau, "Low complexity underwater image enhancement based on dark channel prior," in *Proc. 2nd Int. Conf. Innovations in Bio-Inspired Comput. Appl.*, 2011, pp. 17–20.

[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 Process.*, vol. 29, pp. 4376–4389, 2019.

[6] S. Wu, T. Luo, G. Jiang, M. Yu, H. Xu, Z. Zhu, and Y. Song, "A two-stage underwater enhancement network based on structure decomposition and characteristics of underwater imaging," *IEEE J. Ocean. Eng.*, vol. 46, no. 4, pp. 1213–1227, 2021.

[7] A. Naik, A. Swarnakar, and K. Mittal, "Shallow-UWNet: Compressed model for underwater image enhancement (student abstract)," in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, no. 18, 2021, pp. 15853–15854.

[8] Y. Wang, J. Guo, H. Gao, and H. Yue, "UIEC²-Net: CNN-based underwater image enhancement using two color space," *Signal Process. Image Commun.*, vol. 96, 116250, 2021.

[9] F. Huo, B. Li, and X. Zhu, "Efficient wavelet boost learning-based multi-stage progressive refinement network for underwater image enhancement," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 1944–1952.

[10] Z. Fu, W. Wang, Y. Huang, X. Ding, and K. K. Ma, "Uncertainty inspired underwater image enhancement," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2022, pp. 465–482.

[11] Y. Wei, Z. Zheng, and X. Jia, "UHD underwater image enhancement via frequency-spatial domain aware network," in *Proc. Asian Conf. Comput. Vis. (ACCV)*, 2022, pp. 299–314.

[12] S. Chen, T. Ye, J. Shi, Y. Liu, J. Jiang, E. Chen, and P. Chen, "DEHRFormer: Real-time transformer for depth estimation and haze removal from varicolored haze scenes," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2023, pp. 1–5.

[13] C. Guo, R. Wu, X. Jin, L. Han, W. Zhang, Z. Chai, and C. Li, "Underwater Ranker: Learn which is better and how to be better," in *Proc. AAAI Conf. Artif. Intell.*, vol. 37, no. 1, 2023, pp. 702–709.

[14] H. Yang, J. Xu, Z. Lin, and J. He, "LU2Net: A lightweight network for real-time underwater image enhancement," arXiv Preprint, arXiv:2406.14973, 2024.

[15] J. Jiang, T. Ye, J. Bai, S. Chen, W. Chai, S. Jun, and E. Chen, "Five A$^+$ Network: You only need 9K parameters for underwater image enhancement," arXiv Preprint, arXiv:2305.08824, 2023.

[16] J. M. J. Valanarasu and V. M. Patel, "UNeXt: MLP-based rapid medical image segmentation network," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2022, pp. 23–33.

[17] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 1251–1258.

[18] M. T. Tun, Y. Sugiura, and T. Shimamura, "Lightweight underwater image enhancement via impulse response of low-pass filter based attention network," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2024, pp. 1697–1703.

[19] L. Peng, C. Zhu, and L. Bian, "U-shape transformer for underwater image enhancement," *IEEE Trans. Image Process.*, vol. 32, pp. 3066–3079, 2023.

[20] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren, "Underwater image enhancement via medium transmission-guided multi-color imaging applications," *IEEE Access*, vol. 11, pp. 111715–111737, 2023.

[21] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 3227–3234, 2020.

[22] M. J. Islam, P. Luo, and J. Sattar, "Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception," arXiv Preprint, arXiv:2002.01155, 2020.

[23] R. Cong, J. Xu, X. Liu, C. Zhao, and Y. Zhao, "PUGAN: Physical model-guided underwater image enhancement using GAN with dual discriminators," *IEEE Trans. Image Process.*, vol. 32, pp. 4472–4485, 2023.

[24] 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," *Vis. Comput.*, vol. 41, no. 6, pp. 3735–3783, 2025.

[25] N. Deluxni, P. Sudhakaran, and M. F. Ndiaye, "A review on image enhancement and restoration techniques for underwater optical imaging applications," *IEEE Access*, vol. 11, pp. 111715–111737, 2023.

[26] M. Wang, K. Zhang, H. Wei, W. Chen, and T. Zhao, "Underwater image quality optimization: Researches, challenges, and future trends," *Image Vis. Comput.*, vol. 146, 104995, 2024.

[27] A. Saleh, M. Sheaves, D. Jerry, and M. R. Azghadi, "Adaptive deep learning framework for robust unsupervised underwater image enhancement," *Expert Syst. Appl.*, vol. 268, 126314, 2025.