

Received 7 January 2023, accepted 20 January 2023, date of publication 30 January 2023, date of current version 2 February 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3240648

## RESEARCH ARTICLE

# A Non-Reference Evaluation of Underwater Image Enhancement Methods Using a New Underwater Image Dataset

ASHRAF SALEEM<sup>ID1</sup>, (Member, IEEE), SIDIKE PAHEDING<sup>ID1</sup>, (Senior Member, IEEE), NATHIR RAWASHDEH<sup>ID1</sup>, (Senior Member, IEEE), ALI AWAD<sup>1</sup>, AND NAVJOT KAUR<sup>ID2</sup>

<sup>1</sup>Department of Applied Computing, College of Computing, Michigan Technological University, Houghton, MI 49931, USA

<sup>2</sup>Department of Computer Science, College of Computing, Michigan Technological University, Houghton, MI 49931, USA

Corresponding author: Ashraf Saleem (ashraf@mtu.edu)

**ABSTRACT** The rise of vision-based environmental, marine, and oceanic exploration research highlights the need for supporting underwater image enhancement techniques to help mitigate water effects on images such as blurriness, low color contrast, and poor quality. This paper presents an evaluation of common underwater image enhancement techniques using our new publicly-available Challenging Dataset for Underwater Image Enhancement (CDUIE). The collected dataset is comprised of 85 images of aquatic plants taken at a shallow depth of up to three meters from three different locations in the Great Lake Superior, USA, via a Remotely Operated Vehicle (ROV) equipped with a high-definition RGB camera. In particular, we use our dataset to benchmark nine state-of-the-art image enhancement models at three different depths using a set of common non-reference image quality evaluation metrics. Then we provide a comparative analysis of the performance of the selected models at different depths and highlight the most prevalent ones. The obtained results show that the selected image enhancement models are capable of producing considerably better-quality images with some models performing better than others at certain depths. The dataset is available at <http://www.github.com/ashrafrepo/underwater-image-enhancement>.

**INDEX TERMS** Underwater image dataset, deep learning, non-reference image quality metrics, ROV, underwater image enhancement.

## I. INTRODUCTION

Underwater image enhancement can be useful for researchers in several fields such as geology, ecology, and oceanography with many promising applications such as underwater classification of species, fish counting, coral reefs health monitoring, and infrastructure inspection. Images captured underwater suffer from haze, blurriness, and non-uniform lighting artifacts. This severe degradation of underwater images is due to selective attenuation and scattering effects caused by the travel of light through water [1], [2]. In addition, underwater scenes tend to be bluish or greenish since different light waves are distinctively absorbed by water [3], which in turn, affects the performance of image enhancement

The associate editor coordinating the review of this manuscript and approving it for publication was Felix Albu<sup>ID</sup>.

methods. The recent popularity of Remotely Operated Vehicles (ROVs) facilitated underwater data collection from various types of waterbodies at different depths [4], [5]. Although most ROVs can be equipped with a high-definition RGB camera, the captured frames are typically characterized by low contrast, blurry details, and distorted colors. This presses the need for automatic and robust image quality enhancement in underwater scenarios. While underwater image restoration focuses on reversing the physical transformations that cause the degradation in images, underwater image enhancement focuses only on the degraded images. Underwater image enhancement techniques are categorized into five categories [6] based on the frequency domain, spatial domain, color constancy, fusion, and deep learning, while underwater image restoration techniques are categorized into three categories [6] based on optics, polarization, and prior-knowledge.

Peng et al. [7] used a technique called General Common Channel Prior (GCCP) to determine ambient light to get the scene transmission. In other techniques [8], [9], a fusion strategy was introduced to enhance images based on spatial domain methods such as image inverse, histogram equalization, white balance, and luminance enhancement. In recent years, data-driven deep-learning models have shown promising results in improving underwater image quality. However, these models struggle with the deficit of real underwater data since they require large datasets to generate quality results [8], [10]. A group of researchers [11], [12] incorporated synthetic data in which noise, water, and haze effects are added to casual outdoor images to train and test underwater image enhancement models. Despite the promising results shown by image enhancement techniques that are based on synthetic data, such scenarios do not reflect the natural underwater environment due to the discrepancies between the synthetic data and real data. Hence, such image enhancement models become less effective when encountered by real underwater images. On the other hand, real underwater data can be acquired from publicly available content online to train image enhancement models. For example, Li et al. [10] generated underwater data using publicly available videos online. However, public underwater data are not always available, especially in remote areas that are difficult to reach. To this end, we collected our own real-world underwater image dataset from three different locations in the Great Lake Superior, Michigan's Upper Peninsula, USA. Unlike clear ocean water data, our collected lake-water data have reduced visibility and are significantly more challenging to enhance. One of the problems facing ecologists in those areas is widespread invasive plants that affect native species, reduce light and oxygen levels in the water, and passively harm other organisms [13], [14], [15]. Therefore, the collected data mainly focuses on plants to advocate ecology-related underwater research in the areas at which the data are collected. Then, we conduct a comparative study of the existing analytical and pre-trained deep learning-based image enhancement models by using our dataset as a benchmark. The quantitative evaluation is an integral part of the underwater image enhancement model development process. Therefore, we assess the enhanced images of the selected models using five different types of non-reference evaluation metrics, including 1) Blind/ Referenceless Image Spatial Quality Evaluator (BRISQUE) [16] which calculates the possible loss of naturalness in the image caused by distortion; 2) Naturalness Image Quality Evaluator (NIQE) [17] which assesses the quality of the image by quantifying deviations from statistical regularities in the image; 3) Perception-based Image Quality Evaluator (PIQE) [18]; 4) Entropy-based method [19], [20] which assesses the noise and blurring effects in the images concerning transmitted information on distorted images; 5) a specialized metric for underwater conditions called the CCF [21] which considers the colorfulness, contrast, and fog density indices to assess an underwater image. The primary contributions of this paper are as follows:

- Collect real-world underwater images from multiple areas suffering from ecological problems using an ROV equipped with a high-definition RGB camera. The collected images contain mainly plants and were taken at different water depths.
- Comparatively evaluate the selected state-of-the-art underwater image enhancement models on our real-world dataset using several non-references image quality evaluation metrics and provide insights on their performance.

The rest of the paper is organized as follows: Section II summarizes state-of-the-art image enhancement techniques. The data collection procedure is explained in Section III. Section IV details the utilized enhancement models and metrics while Section V discusses the obtained results. Finally, Section VI concludes the paper.

## II. RELATED WORK

Various image enhancement methods are developed and published by researchers to improve the quality of underwater images. Traditional image enhancement methods are categorized into physical and non-physical models, also known as restoration and enhancement models, respectively. In addition, newer deep learning-based methods can be categorized into convolutional neural networks (CNN)-based and Generative Adversarial Network (GAN)-based models [22]. On the other hand, various underwater datasets are also developed and published online to support the further development of image enhancement models.

### A. TRADITIONAL METHODS

Researchers developed numerous methodologies to enhance underwater images with a focus on either the cause of degradation (physical models) or the result of degradation (non-physical models). For instance, Hou et al. [23] proposed an underwater image synthesis algorithm (UISA) that utilizes hierarchical search and red channel prior algorithm to get underwater background light and transmission map from real-world underwater images. Then, a synthetic underwater image dataset (SUID) is generated using the proposed algorithm. However, the synthesized output images of the UISA don't reflect real-world underwater images that have a motion-blurring effect due to scattering. In [1], a natural-based underwater image enhancement (NUCE) model is proposed based on: 1) neutralizing underwater color cast where the gain factor is used to enhance the inferior color channels; 2) the dual-intensity image fusion to produce lower and upper stretched histograms; 3) a mean equalization technique to give a natural alike quality to output images based on a swarm intelligence algorithm; 4) a masking technique that sharpens the images. The approach presented in [1] is able to significantly reduce the underwater cast on the color based on its four components. Nonetheless, the effectiveness of each of its comprising components should be further evaluated by an ablation study.

Chang [24] developed two distinct transmission coefficient estimation approaches, namely 1) optical characteristics; 2) the essence of image processing knowledge. Weighted by saliency maps, the two transmission maps fused into one transmission map to get the outcome. Although the proposed model in [24] is capable of applying distinct enhancements to the background and foreground of an image, its performance compared to other baseline image restoration models is yet to be determined. Song et al. [25] developed a conventional model-based method using a manually annotated background lights database. The statistical models of background light estimation are provided using the relationship between the images in the dataset with histogram distribution. Then, the transmission map of the red channel is generated by the underwater dark channel prior (UDCP), and compensated by the adjusted reversed saturation map (ARSM) and underwater light attenuation prior (ULAP). Subsequently, the transmission maps of the green and blue channels are estimated depending on the attenuation ratio difference with the red channel. Finally, the output is post-processed by a white-balancing technique. While the method presented in [25] is novel, computationally inexpensive, and achieves superior performance, further optimization to the model is needed to incorporate the green and blue channels in the estimation process of the transmission map.

Paheding et al. [8] proposed a novel method for color image enhancement by designing an adaptive trigonometric transformation function that can help to improve visual quality. The utilized transformation function is based on a tangent with characteristics that vary depending on the luminance of the images. A Laplacian operator with a process for color restoration is also combined with the transformation function to obtain Images with well-balanced colors. Despite the fact that the novel model developed in [8] can adaptively adjust the intensity value of images, its performance in various noisy image conditions is yet to be evaluated. Muniraj and Dhandapani [4] combined a color constancy framework with a dehazing technique. Gamma correction, white patch retinex (WPR), and chromatic adaptation technique (CAT) are used for color constancy in the first step. Then, dehazing is performed using the estimation of artificial background light and transmission map depth, while the depth estimation is calculated using the difference of channel intensity prior (DCIP). Finally, the gamma-corrected HIS images are transformed into RGB images. The study presented in [4] is comprehensive in terms of providing an evaluation using reference and non-reference metrics, comparing the performance with non-traditional approaches, and conducting an ablation study on the model's components.

Bai et al. [26] introduced a novel underwater image enhancement method based on four stages of pixel intensity: 1) center regionalization; 2) global equalization of histograms; 3) local equalization of the histogram; 4) multi-scale fusion. Unlike [4], this method also utilized the gamma corrector to correct the problem of over-enhancement. Although

the methodology presented in [26] performs well on severely degraded images by noise, sand, fog, and low light, it is unable to produce consistent background colors when provided with images from different imaging sensors. Another group of researchers developed variational image enhancement models. For example, Xie et al. [27] proposed a variational framework that generates the transmission map depending on the hierarchical search and the red channel prior. Moreover, the sparse prior knowledge and the total variation item are incorporated, then blur kernel estimation is done by changing the resolution. Subsequently, the resultant optimization problem is solved depending on the alternating direction method of multipliers (ADMM). While the proposed model in [27] takes the forward scattering component into consideration because it represents a complete underwater image formation model (UIFM), it ignores texture details and focuses on other attributes. Li et al. [28] developed a novel framework based on pyramids technology and variational methods. Furthermore, the contrast is enhanced without impacting the textures by properly designing the total Laplacian model and the adaptive variational contrast enhancement (AVCE) model. The developed variation models are solved using the alternating direction method of multipliers (ADMM) and the gradient descent method (GDM). Even though the proposed method in [28] achieves good performance in naturally illuminated water, it poorly handles non-uniformly illuminated images using artificial lights. The authors in [29] presented a framework based on adaptive color and contrast enhancement (ACCE) and denoising. First, the low-frequency and high-frequency components are separated using two filters, then the low-frequency component is enhanced based on the ACCE while the high-frequency component is denoised. Afterward, the ACCE is solved with an accelerated pyramid-based method. Although the proposed method in [29] adds a considerable 5% improvement compared to other baseline methods, it doesn't yet handle all aspects of image degradation. Hou et al. [30] integrated the UIFM with the variational framework depending on a non-local differential operators approach. The UDCP and quad-tree subdivision are used to build the UIFM and estimate the background light and transmission map. Then, the formed optimization problem is solved based on the ADDM and the output is treated with a gamma correction procedure to improve the saturation. While the work done in [30] is effective in removing haze from real and synthetic underwater images, it cannot process all kinds of haze such as the haze resulting from fog. Zhuang et al. [31] designed a retinex-based variational model inspired by the priors of the hyper-Laplacian reflectance. Particularly, statistical methods are employed for color correction, then a conversion to the HSV color space is conducted to ensure less variant illumination and fewer artifacts. Subsequently, the V channel is decomposed into illumination and reflectance layers each of which is applied to the designed retinex-based variational model. The enhanced V channel is generated as the product

of the resultant enhanced illumination and reflectance layers. Finally, the images are converted from HSV color space back to RGB color space. Even though the quantitative and qualitative evaluation of the technique presented in [31] demonstrates its superiority compared to other methods, the runtime performance is still considerably higher than the runtime of state-of-the-art data-driven methods. Zhang et al. [32] introduced an underwater color and contrast image enhancement model named JOE-ACDC. In particular, the authors designed a special attenuation matrix to correct the color of underwater images based on the discrepancies between different color channels, then, both global and local contrasts in the resultant images were improved using histogram-based methods and fused using a multiscale fusion technique. The fused images are treated with an unsharp mask as a final stage of refinement. Although the method proposed in [32] can be extended to handle low-light and blurred images, it is computationally expensive and may degrade the color of low-quality images. The same researchers later developed a robust and efficient model dubbed the MMLE [33] to enhance underwater images. Specifically, the color of the images is corrected while considering two principles, namely map-guided maximum attenuation fusion and minimum color loss. Subsequently, the contrast is adaptively adjusted depending on the local statistical metrics of image blocks. Finally, the color of the ‘a’ and ‘b’ channels in the color space of CIELAB is balanced. Despite the fact that the method proposed in [33] is capable of enhancing the color of foggy and dusty images, it is still incapable of enhancing low-light underwater images.

## B. DEEP LEARNING-BASED METHODS

Some researchers developed underwater image enhancement methods based on CNNs while others based their work on GANs. For example, The CNN-based WaterNet model [10] adopted the fusion strategy based on image degradation characteristics to apply White Balance (WB), Histogram Equalization (HE), and Gamma Correction (GC) algorithms for an underwater image. Although the work proposed by [10] introduced and evaluated a relatively large underwater, the utilized strategy for the generation of reference images is hugely affected by backscatter. Li et al. [34] used an underwater scene prior and a synthesis algorithm to construct a UWCNN model based on a lightweight CNN model. The model took into consideration various types and degradations of underwater images to train the network. However, it poorly performs in handling low-contrast in-door synthetic training data.

Another group of researchers focused their work on GAN-based models. For instance, the model proposed in [35] is based on a fusion generative adversarial network, called the DewaterNet. The DewaterNet model eliminates the element-wise matrix product by adding the output of the simple network to the output of the network that is supplied with raw underwater images. The model was tested

on the Underwater Image Enhancement Benchmark Dataset (UIEBD) [10]. Although the authors in [35] claim to implement the first attempt of blending two inputs in underwater GANs, no ablation studies were conducted to verify the effectiveness of the proposed architecture. Liu et al. [36] proposed a new model called UResnet based on a very-deep super-resolution reconstruction model (VDSR). The study used cycle-consistent adversarial networks (CycleGAN) to generate the synthetic image data. UResnet is made up of ResBlocks and these ResBlocks learn to differentiate between label image and input image through the skip connection between the head and body section of the model. Two types of testing data were used: the first was 221 underwater images and the second was a synthetic dataset generated by CycleGAN using in-air images. Despite the fact that the model proposed in [36] achieves the best results compared to other baseline models, it is yet to be generalized to provide other image enhancement features such as dehazing. Chen et al. [37] developed a deep enhancement model using detection preceptors, named HybridDetectionGAN. The preceptors work as gradients to guide the model to generate good output images. Due to the lack of underwater data, a synthesis model is proposed based on fusing the data-driven cues and physical priors. While the developed data synthesis model can learn the translation between underwater and in-air images robustly, the proposed enhancement model can be improved by incorporating a GAN with Bayesian estimation allowing it to be generalized to handle diverse underwater datasets. In [38], a conditional generative adversarial network-based model, called FunIE-GAN, is presented. An objective function is constructed to evaluate the perceptual quality of an image using information such as color, style, and texture. The model presented in [38] offers good performance and considerably faster inference time compared to other baseline approaches. However, it poorly performs when trained with unpaired datasets. Hambarde et al. [39] developed UW-GAN to estimate the depth from a single underwater image and performed enhancement tasks. Various UW networks were used to estimate the depth such as UWC-Net for coarse-level depth and UWF-Net for fine-level depth. The proposed UWF-Net uses spatial and channel-wise squeeze in addition to excitation blocks to estimate the fine-level depth. The model uses real-world images from the internet combined with synthetic datasets. Although the model proposed in [39] achieved relatively good performance by virtue of its coarse and fine-level depth estimation, further optimization of the model is still needed to be able to run on embedded computing platforms. Han et al. [40] proposed a spiral generative adversarial framework, named Spiral-GAN using several deconv-conv blocks. The study presents a spiral learning strategy while considering the pixel-wise loss and angle error of the objective function. Despite the fact that the proposed model in [40] generates richer details and colors, it can be further developed to manage lower-level image enhancement tasks such as de-hazing and de-noising. The study in [41]

proposed a new solution to enhance underwater images using a multiscale dense GAN. The model used multiscale, dense concatenation, and residual learning to achieve good performance. While the proposed model in [41] performs well on real underwater images, its performance on synthetic underwater images still needs to be improved.

Li et al. [42] presented Ucolor, a robust model that uses medium transmission-guided multicolor space embedding with attention mechanisms to focus on discriminative features. Ucolor is inspired by underwater physical models and is comprised of a multi-color space encoder network with a medium transmission-guided decoder network. The Ucolor model achieves high performance across multiple benchmark datasets. However, it struggles, as well as state-of-the-art methods, when applied to low-light images. Fu and Cao [43] combined deep learning with a physical-based model. In particular, a two-network-based deep learning model was introduced to improve the distorted color and low contrast, as well as to deal with over-enhancement and compressed histogram equalization. The hybrid architecture in [43] is lightweight and fast since it efficiently utilizes the advantages of both traditional and non-traditional methods. However, other configurations of combined traditional and non-traditional methods may achieve even faster performance.

### C. UNDERWATER IMAGE DATASETS

Even though underwater data are relatively limited compared to in-air data, there is a growing number of publicly available underwater datasets that are used to benchmark image enhancement models [44] thanks to the recent contributions made by many researchers. Following is a brief description of common public underwater datasets. Named after the original SUIM dataset [45], the SUIM-E dataset [46] is a modified version in which the enhancement references are manually supplied to 1635 selected images which are divided into 1525 training samples and 110 testing samples. Furthermore, that dataset contains a plethora of underwater scenes such as fish, wrecks, and aquatic plants. The references are manually selected based on human judgment since using image enhancement metrics can result in biases towards over-enhanced images [47]. Researchers in [44] constructed the UID2021 underwater dataset by selecting 60 images from other publicly available datasets and some online websites. The selected images are cropped to a unified resolution of 512x384 and are categorized into six groups depending on the type of scene in each image such as bluish and greenish scenes. Then, 900 enhanced images are generated by 15 state-of-the-art image enhancement models using the selected images. Specialized software and a group of 52 volunteers were tasked with selecting the reference-enhanced images, then the mean of opinion score (MOS) method is used to generate the set of ground-truth images for the dataset. Li et al. [10] introduced the Underwater Image Enhancement Benchmark (UIEB) which is comprised of 950 images with 800 images used for training, 90 images used for testing, and

the rest are considered challenging images. The dataset is created from publicly-available videos online and privately collected videos containing a wide variety of underwater scenes. Researchers in [38] collected the Enhancement of Underwater Visual Perception (EUVP) dataset that contains a gigantic number of paired and unpaired underwater images. The images were collected using seven different cameras in variable visibility conditions and locations. Song et al. [25] composed the first underwater dataset of 500 images with manually annotated background lighting (MABLs). Various scenes, plants, animals, and organisms are contained within the dataset with multiple sources of distortions such as scatter and low visibility. Lui et al. [9] proposed the real-world underwater image enhancement (RUIE) dataset containing more than 4,000 images divided into three groups to accommodate different image enhancement aspects, namely quality (UIQs), color deviation (UCSS), and advanced mission drivers (UHTS) group. Berman et al. [48] proposed the SQUID, a comprehensive dataset collected under natural light at various depths, seasons, and water bodies. Images in the squid are associated with color charts and depth maps to assist in the evaluation of color correction techniques and other image enhancement tasks. Researchers in [49] selected images from the massive ImageNet dataset [50] and generated synthetic underwater images based on CycleGAN [36]. Hou et al. [23] established a benchmark for full-reference image enhancement evaluation called the synthetic underwater image dataset (SUID). The SUID is comprised of 900 synthesized underwater images and 30 ground-truth outdoor images.

### III. REAL-WORLD UNDERWATER DATASET

A common challenge in the research of image enhancement is the availability of real-world data. Many studies in the past either used synthetic underwater datasets generated by GAN models or underwater data collected from existing online streaming platforms. Such datasets do not necessarily and accurately reflect real-world data variations and artifacts. In addition, data generated from online video streaming platforms do not necessarily conform to a data collection standard, (e.g., collecting images of the same scene or object from various depths, angles, and at different times and conditions). In recent underwater datasets, Autonomous Underwater Vehicles (AUV) and Remotely Operated Vehicles (ROV) are usually used for data collection and image acquisition. In this work, real-world underwater RGB images are collected using Geneinno T1 Pro ROV [51] which is shown in Figure 1. The collected video data are taken from three different geographical locations, namely Lake Linden, Chassell Bay, and Portage Lake in Michigan, USA. The specifications of the CMOS camera that is equipped on the ROV are shown in Table 1. To construct our Challenging Dataset for Underwater Image Enhancement (CDUIE) dataset available [[here](#)], we manually extract convenient frames from the collected videos at the original 3840x2160 pixel resolution without any change or modification. The collected dataset



**FIGURE 1.** Geneinno T1 Pro ROV used for data collection.

**TABLE 1.** Camera hardware specifications.

Unit	Specification
CMOS	1/2.3" SONY CMOS
IRIS	F2.5
FOV	160°
Shutter Speed	1/30s
Still Image Size	4000 x 3000 - 12MP
Video Resolution	4K: 30 fps, FHD: 60 fps
Live Streaming Resolution	480 / 720 / 1080p
Max. Video Bitrate	2 Mbps (Live) / 12 Mbps (Video)
Video Format	MP4 (MPEG-4 H264/H265)
Lens anti-shake	EIS Anti Shake
Depth Sensor Precision	+/- 0.05m

contains 85 underwater plant images that are separated based on depth into three groups; 1) images taken at a depth of less than one meter; 2) images taken at a depth of one-two meters; 3) images taken at a depth of two-three meters. Sample images from the three proposed depths are shown in Figure 2 where the effect of even such a minor depth difference can be clearly and easily noticed on the three groups of images. Moreover, the collected images are very challenging compared to underwater images collected in other datasets due to the severe degradation added by the extremely turbid lake environment and low-light conditions. This causes some images to be blurry, poorly contrastive, and sometimes noisy due to the dust that is pulled by the ROV.

#### IV. IMAGE ENHANCEMENT

A detailed description of the nine employed state-of-the-art image enhancement models is provided in this section. In addition, the five utilized non-reference image enhancement metrics are to be described in detail later in this section.

#### A. METHODS UNDER STUDY

In this study, the following image enhancement methods are selected for a comparative evaluation: WaterNet [10], Ucolor [42], Global-Local Network and Compressed Histogram Equalization (GLN-HE) [43], UW-CNN [34], Statistical Model of Background Light and

Optimization of Transmission Map (SMBLOT) [25], Image Inverse [52], Adaptive Trigonometric Transformation Function (ATTF) [8], JOE-ACDC [32], and MMLE [33]. Our proposed dataset is used as a benchmark to evaluate the aforementioned image enhancement methods. In the performed experiments, we re-implemented the selected models with the default parameters using their publicly-available codes to generate the enhanced images. following is a detailed description of the enhancement models under study.

##### 1) WaterNet

WaterNet [10] is a gated fusion model. It first generates three inputs: White balance (WB), Histogram Equalization (HE), and Gamma Correction (GC), then the Feature Transformation Units (FTU) are added to the generated three inputs. This helps in reducing the color cast and artifacts in the input images. These generated inputs are then sent to the gated Network. The enhanced image can be expressed as:

$$I_{en} = R_{WB} \odot C_{WB} + R_{HE} \odot C_{HE} + R_{GC} \odot C_{GC}, \quad (1)$$

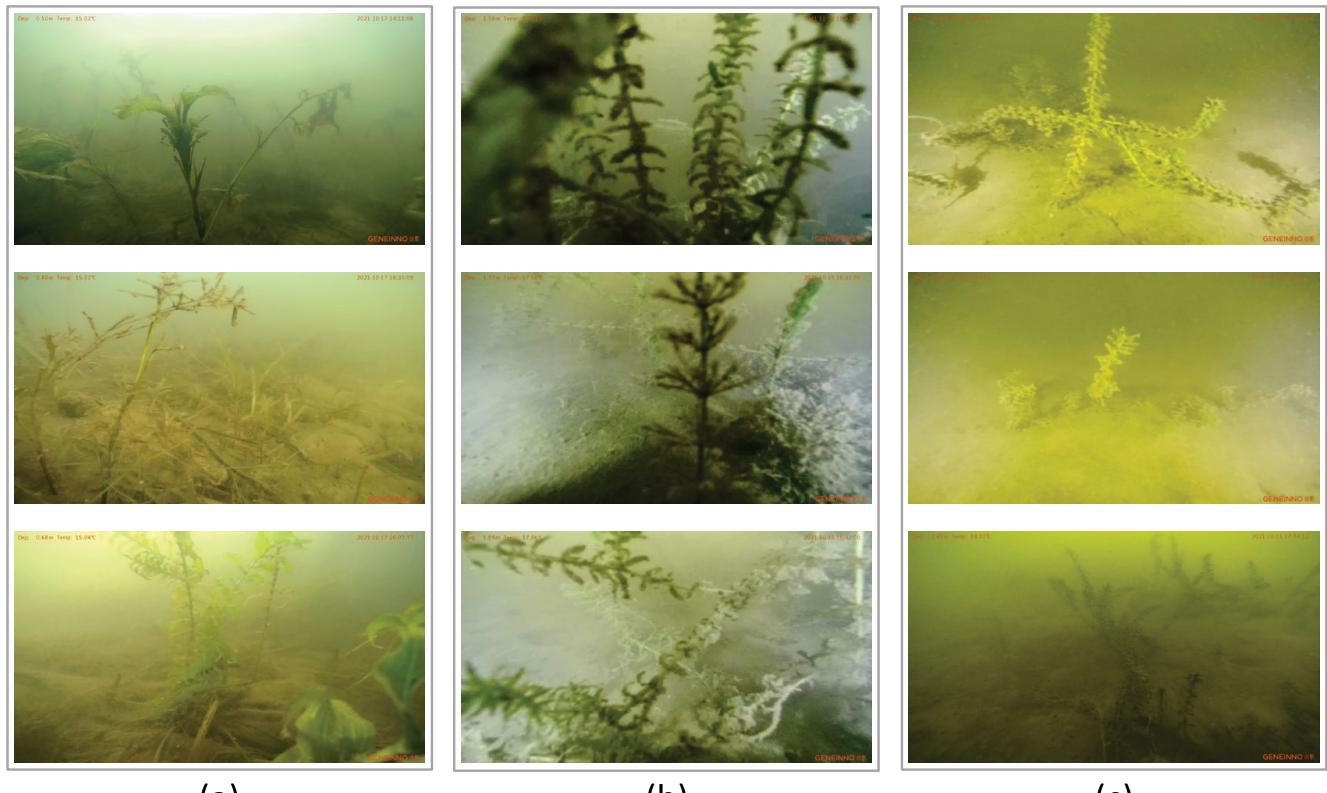
where  $\odot$  is the element-wise product of the refined input RWB, RHE, and RGC resulting from the WB, HE, and GC inputs, respectively. The CWB, CHE, and CGC are the learned confidence maps that are calculated to generate the output. The authors assume that the L1 and L2 pixel-wise loss functions can add artifacts, therefore, the perceptual loss function shown in Eq. (2) was minimized to learn the mapping function of underwater images to generate realistic results.

$$L_j^{\varphi} = \frac{1}{C_j H_j W_j} \sum_{i=1}^N \left\| \varphi_j \left( I_{en}^i \right) - \varphi_j \left( I_{gt}^i \right) \right\|^2, \quad (2)$$

where C, H, and W are the feature map's number, height, and width, respectively, j is the index of the VGG19 network convolutional layer, and N is the batch number. The model uses standard Gaussian distribution to initialize the filter weights. The initial learning rate is set to 1e-3, then decreased to 0.1 at the 10,000th iteration until the model converged.

##### 2) UColor

UColor model [42] uses medium transmission-guided multi-color space embedding in the underwater image enhancement network. The underwater images first go through color space transformation. The inputs are forwarded through three encoder paths, namely HSV, RGB, and Lab paths. Then the selected encoder features along with a reverse medium transmission (RMT) map of the same size are sent to the medium transmission guidance module. Different sizes of reverse medium transmission are achieved by a max pooling operation. Then, the output is forwarded into three residual-enhancement modules with a 2x up-sampling operation. The model considers image degradation components using a multicolor space to facilitate the measurements of color deviation. The objective of the residual-enhancement



**FIGURE 2.** Samples of the proposed CDUIE dataset. (a) depth < 1 meter, (b) depth 1-2 meters, (c) depth 2-3 meters. The full CDUIE is available on <http://www.github.com/ashrafrepo/underwater-image-enhancement>.

module is to preserve the data accuracy and solve the problem of the gradient vanishing. In the encoder, the filters are increased from 128 to 512, while in the decoder network the filters are decreased from 512 to 128 by a factor of 2, respectively. The kernel size is  $3 \times 3$  and the stride is one in the convolutional layers. The model extracts the channel features from different color spaces and exploits the Channel attention module to utilize the interconnection between them. The medium transmission guidance module implemented in the decoder network uses the reverse medium transmission (RMT) map as the pixel-wise attention map. In addition, the model uses the GDCP (general dark channel prior) to obtain the medium transmission due to the unavailability of the ground truth data. RMT weights serve to avoid gradient vanishing and tolerate errors caused by inaccurate medium transmission estimation.

### 3) GLOBAL-LOCAL NETWORK AND COMPRESSED HISTOGRAM EQUALIZATION (GLN-HE)

GLN-HE [43] is a two-branch network; Network-G (global network) and Network-L (local network). The Network-G takes the average of the mean  $\mu_I$  and standard deviation  $\sigma_I$  to provide first-order color measurement and second-order dispersion information. In this model, hidden layers are concatenated to anticipate the residual  $\Delta\hat{\mu}$  which helps to generate

the compensated average  $\hat{\mu}_j$  as shown in Eq. (3-7).

$$h_1 = \text{ReLU}(W_1 \cdot \text{concat}(\mu_1, \sigma_1) + b_1), \quad (3)$$

$$h_2 = \text{ReLU}(w_2 \cdot h_1 + b_2), \quad (4)$$

$$h_3 = \text{ReLU}(w_3 \cdot h_2 + b_3), \quad (5)$$

$$\Delta\hat{\mu} = w_4 \cdot \text{concat}(h_1, h_2, h_3) + b_4, \quad (6)$$

$$\hat{\mu}_j = \text{Sigmoid}(\mu_1 + \Delta\hat{\mu}), \quad (7)$$

where  $h$  indicates hidden features,  $w$  is the learnable weights,  $b$  is the bias,  $\text{ReLU}(\cdot)$  is the rectified linear units,  $\text{sigmoid}(\cdot)$  is a sigmoid activation function, and  $\text{concat}(\cdot)$  is concatenation. The architecture of Network-L is similar with the exception of conducting convolutional operations to process the input matrix that contains color distortions. The model establishes global residual-guide bias for accurate local contrast compensation. Moreover, Compressed-histogram equalization is implemented to obtain a uniform distribution by matching the cumulative input histogram. The compression of peaks in the input histogram aid to evade the over-enhancement problem. A logarithmic operation is used to alter the input histogram which can effectively compress large peak values while maintaining the order of the input histogram.

### 4) UWCNN

UWCNN model [34] uses a lightweight densely connected FCNN network and takes an RGB image  $U$  as an input.

This model handles the vanishing/exploding gradients at the training stage by enforcing the learning of the residual and adding the input to the output of the network before the loss function as shown in Eq. (8), where + is the element-wise addition operation.

$$I = U + \Delta(U, \theta). \quad (8)$$

The network also includes a chain of enhancement units connected to a final convolution layer. The layers in the network are categorized into three types: 1) Layer One which is made up of 16 convolutional layers of the size of  $3 \times 3 \times 3$  to generate 16 output feature maps. 2) Layer Two which consists of the ReLU activation function to establish the non-linearity. 3) Layer Three which is implemented to concatenate all other layers after each block. The input is fed to each block so that it can be applied to the chain of enhancement units. Each enhancement unit comprises three convolutional layers and a single output layer at the end. To avoid boundary artifacts and generate artifacts-free output, the pooling layers are not deployed and a value of zero is added before each convolutional layer.

##### 5) STATISTICAL MODEL OF BACKGROUND LIGHT AND OPTIMIZATION OF TRANSMISSION MAP (SMBLOT)

In SMBLOT [25], two approaches are used; the first is to estimate the overall background light of RGB channels and the other is to estimate the transmission map of RGB channels. This can be achieved based on the new UDCP integrated with the TM optimizer and an exponential decay function of the mean of the RGB channels. Afterward, these estimates are applied to dehaze the input images. Furthermore, a white balance color correction mechanism is used to restore the colors. To achieve this, a manually annotated Background Light (MABL) dataset is created, then a tight correlation between the MABL and its distribution on the histogram is discovered. Based on the correlation results of channels, the linear model of the average (Avg) and the standard deviation (Std) for the BL estimation of the G-B channel is defined as:

$$B^{c'} = \alpha \times Avg^{c'} + \beta \times Std^{c'} + \gamma, \quad (9)$$

where  $Avg^{c'}$  is the average and  $Std^{c'}$  is the standard deviation of the input image;  $\alpha$  and  $\beta$  are coefficients and  $\gamma$  is a constant. Finally, a non-linear model for the R channel is defined as follows where a, b, and c are coefficients.

$$B^r = \frac{a}{1 + b \times Exp(c \times Medr)} \Omega \quad (10)$$

##### 6) ATTF WITH CONTRAST IMPROVEMENT AND COLOR RESTORATION (ATTF-CCR)

ATTF-CCR model [8] consists of three major stages, namely, adaptive luminance enhancement by ATTF, contrast improvement through high-frequency (HF) boosting, and color restoration. To improve the colors, it first calculates the

luminance of the input image using the NTSC formula:

$$I(x, y) = 0.299 \times I_r(x, y) + 0.587 \times I_g(x, y) \\ + 0.114 \times I_b(x, y), \quad (11)$$

where  $(x, y)$  is the pixel location and  $I(x, y)$  indicates the intensity at a pixel location.  $I_r(x, y)$ ,  $I_g(x, y)$ , and  $I_b(x, y)$  represent R, G, and B values. Furthermore, to capture the minor details in the input images, high-frequency (HF) boosting for contrast improvement is used. The HF component is extracted after combining the original grayscale image with a high-pass filter, then both are added to the enhanced image as follows:

$$S_h(x, y) = 255 \cdot S(x, y) + I(x, y) * h, \quad (12)$$

where  $h$  is a smoothed Laplacian operator. The color content restoration is done using the linear color restoration approach, which results in the final enhanced image using the following equation:

$$E_j(x, y) = I_j(x, y) \cdot \frac{S_h(x, y)}{I(x, y)}, \quad (13)$$

where  $j$  represents the color channels and  $E$  is the enhanced image of the corresponding color channel.

##### 7) IMAGE INVERSE

Image Inverse [52] is widely used to convert dark or bright intensities in the input image to bright or dark intensities in the output image. Given that the intensity of a grayscale pixel falls in the range  $[0, 255]$ , the inverse is acquired by calculating the complementary value. Since a colored pixel has three values (red, green and blue) that fall in the range  $[0, 255]$ , the inverse is acquired by calculating the complementary of each of the three values as shown in Eq. (14).

$$P(i, j, k) = 255 - P(i, j, k), \quad (14)$$

where  $P$  is the colored image matrix,  $i$  and  $j$  represent pixel indices and  $k$  represents the color channel with  $1 \leq k \leq 3$

##### 8) JOE-ACDC

The JOE-ACDC model [32] corrects the color of attenuated channels and enhances the contrast without demolishing the details in the input images depending on four stages. First, the attenuation matrices of the most prevalent RGB channel are used to correct the color of other channels depending on the water type (e.g., the green channel is used for turbid water and the blue channel is used for deep water). The channel selection is done by calculating the total pixel intensity mean given by the following equation:

$$Mean_c = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N I_c(i, j), \quad c \in \{R, G, B\} \quad (15)$$

where  $M$  is the height and  $N$  is the width of a given image  $I$ . Then, the local contrast is improved using a method based on limited histogram and Rayleigh distribution while the global contrast is improved using a method based on dual histograms and iterative threshold. Subsequently, the weight

maps of saliency and brightness are generated to fuse the local and global contrast-enhanced versions. The weight map of brightness  $W_{B,k}$  is given by the equation below:

$$W_{B,k}(i,j) = \exp\left(-\frac{(L_k(i,j) - 0.5)^2}{2 \times (0.25)^2}\right), \quad (16)$$

where  $L_k(i,j)$  is the value of a pixel in the input image  $I_k$  at position  $(i,j)$ . The weight map of saliency  $W_{S,k}$  is given in the following equation:

$$\begin{aligned} W_{S,k}(i,j) = & (H_k(i,j) - MH_K)^2 \\ & + (S_k(i,j) - MS_K)^2 \\ & + (V_k(i,j) - MV_K)^2, \end{aligned} \quad (17)$$

where  $H_k$  is the hue,  $S_k$  is the saturation,  $V_k$  is the value,  $MH_K$  is the mean hue,  $MS_K$  is the mean saturation, and  $MV_K$  is the mean value. Finally, the output images are unsharpened based on an image-dehazing framework proposed in [53] with a reduced number of Gaussian Kernels using the following equation:

$$I\_Blur_n(i,j) = G_n(i,j) * I_{input}(i,j), \quad (18)$$

where  $n$  is the number of the Gaussian kernel function filtering scales,  $*$  is the operator of the convolution, and  $G_n$  is defined as:

$$G_n(x,y) = \frac{1}{2\pi\sigma_n^2} \exp(-(x^2 + y^2)/(2\sigma_n^2)). \quad (19)$$

### 9) MMLE

The MMLE model [33] produces color-corrected images based on the principle of minimum color loss and a map-guided fusion strategy that allows for an adaptive color and details adjustment. To generate the color transfer image, the mean of every color channel is calculated, then an iterative process is employed based on the minimum color loss principle defined by the equation below:

$$L_{color} = (\bar{I}_l - \bar{I}_m) + (\bar{I}_l - \bar{I}_s), \quad (20)$$

where  $\bar{I}_l$ ,  $\bar{I}_m$ , and  $\bar{I}_s$  are the mean values of the color channels with the largest, medium, and smallest mean values, respectively. Subsequently, the maximum attenuation map is selected and employed to mitigate some color distortions produced by the previous process. Finally, a fusion strategy that utilizes the color transfer image and modifies the colors of the input image while considering the loss of details is implemented. On the other hand, this model improves the contrast by converting the color-corrected images to CIELAB space. The contrast of the luminance channel is locally enhanced by first calculating the variance and mean of local image blocks using the integral map and squared integral map, then an enhancement transformation based on the equation defined below is applied:

$$L_{EB}(i,j) = \mu_B + \alpha(L_B(i,j) - \mu_B), \quad (21)$$

where  $\alpha$  is the control factor of enhancement and  $\mu_B$  is the mean of an image block B with a greyscale matrix  $L_B$ . Simultaneously, the mean values of both the a and b channels of the CIELAB space are computed and employed to compensate for the color imbalance between the two channels using the following equation:

$$\begin{cases} I_{bc} = I_b + \frac{(\bar{I}_a - \bar{I}_b)}{(\bar{I}_a + \bar{I}_b)} I_b, \bar{I}_a > \bar{I}_b \\ I_{ac} = I_a + \frac{(\bar{I}_b - \bar{I}_a)}{(\bar{I}_b + \bar{I}_a)} I_a, \bar{I}_a < \bar{I}_b, \end{cases} \quad (22)$$

where  $\bar{I}_a$  is the mean value of channel a,  $\bar{I}_b$  is the mean value of channel b, and  $I_{bc}$  and  $I_{ac}$  are the color balanced channels of a and b, respectively.

## B. IMAGE QUALITY METRICS

Our quantitative comparison is done by evaluating the image enhancement models under study using the five non-reference image enhancement metrics that are described in detail below.

### 1) BRISQUE

The Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) estimates the possible loss of naturalness in the image that can be the result of the distortion present in the image, including compression artifacts, gaussian pixel noise, and blurring [16]. The metric is based on human observer opinion of distortions, and a smaller BRISQUE score indicates better image quality. Using Eq. (23), the metric utilizes a pixel-wise preprocessing model referred to as mean subtracted contrast normalization (MSCN) to transform image pixel luminances  $I(i,j)$  to reduce the dependencies of Gaussian characteristics of neighboring pixels:

$$\hat{I}(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + C}, \quad (23)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of pixel intensities in a  $3 \times 3$  neighborhood of  $(i,j)$ , and the constant  $C = 1$  is used to prevent instabilities when the denominator tends to zero, for bland image regions such as the sky in a natural scene. The metric also relies on modeling the statistical relationships between each pixel, i. e. locations  $(i,j)$ , and its  $3 \times 3$  neighbors in the horizontal (H), vertical (V), and both diagonal directions (D1 and D2), using the pairwise products shown in Eq. (24-27):

$$H(i,j) = \hat{I}(i,j)\hat{I}(i,j+1), \quad (24)$$

$$V(i,j) = \hat{I}(i,j)\hat{I}(i+1,j), \quad (25)$$

$$D1(i,j) = \hat{I}(i,j)\hat{I}(i+1,j+1), \quad (26)$$

$$D2(i,j) = \hat{I}(i,j)\hat{I}(i+1,j-1). \quad (27)$$

### 2) NIQE

The Naturalness Image Quality Evaluator (NIQE) builds a model to assess the quality of the image by quantifying deviations from statistical regularities that are present in the natural image [17]. A lower NIQE score indicates better

image quality. Unlike the human-based scoring in BRISQUE, the NIQE metric is based on the space domain Natural Scene Statistic (NSS) model. NSS features are calculated from the image under processing and are compared to those obtained from an image database used to train the model. The features are expressed as a multi-variate Gaussian distribution. The metric's quality score is based on the deviation of image statistics from the Gaussian model, as shown in Eq. (28).

$$D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{(v_1 - v_2)^T \left( \frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (v_1 - v_2)} \quad (28)$$

where  $v_1$  and  $\Sigma_1$  are the mean vector and covariance matrix of the natural Multivariate Gaussian model.  $v_2$  and  $\Sigma_2$  are the mean vector and covariance matrix of the distorted Multivariate Gaussian model.

### 3) PIQE

The Perception-based Image Quality Evaluator (PIQE) mimics human behavior in judging image quality without training data. It is based on the perceptual significance of local structural features and the generation of a block-level distortion map [18]. The PIQE is calculated as follows:

$$PIQE = \frac{\left( \sum_{k=1}^{NSA} D_{sk} \right) + C_1}{NSA + C_1}, \quad (29)$$

where  $NSA$  is the number of Spatially Active blocks in a given image, and  $C = 1$  is used for numerical stability when  $NSA$  is small. In addition,  $D_{sk}$  is the spatial distortion assignment for a given block  $B_k$ . A smaller PIQE score indicates better perceptual image quality.

### 4) ENTROPY

The Entropy is a statistical measure of randomness [19] that can characterize the texture of images [20], [54]. We calculate it as the discrete entropy (DE):

$$DE(X) = - \sum_{k=1}^K p(x_k) \log p(x_k), \quad (30)$$

where  $p(x_k)$  is the probability of the pixel intensity  $x_k$  as obtained from the optical density histogram having  $K$  luminance levels. A higher ENTROPY score indicates richer details, i. e., higher image quality.

### 5) CCF

The CCF metric [21] is specially designed to evaluate underwater image enhancement and restoration techniques. It is developed as a linear regression model with three input features: 1) colorfulness index quantifying the color loss caused by absorption, 2) contrast index quantifying the blurring effect caused by forward scattering, 3) fog density index quantifying the fog effect caused by backward scattering. The multivariate linear regression model is trained to generate the proper weight for each of the above features using

a special underwater image dataset with the ground truth being supplied based on the mean opinion score (MOS) from twenty volunteers. The proposed CCF regression model can be expressed as follows:

$$CCF = \omega_1 \times Colorfulness + \omega_2 \times Contrast + \omega_3 \times Fogdensity, \quad (31)$$

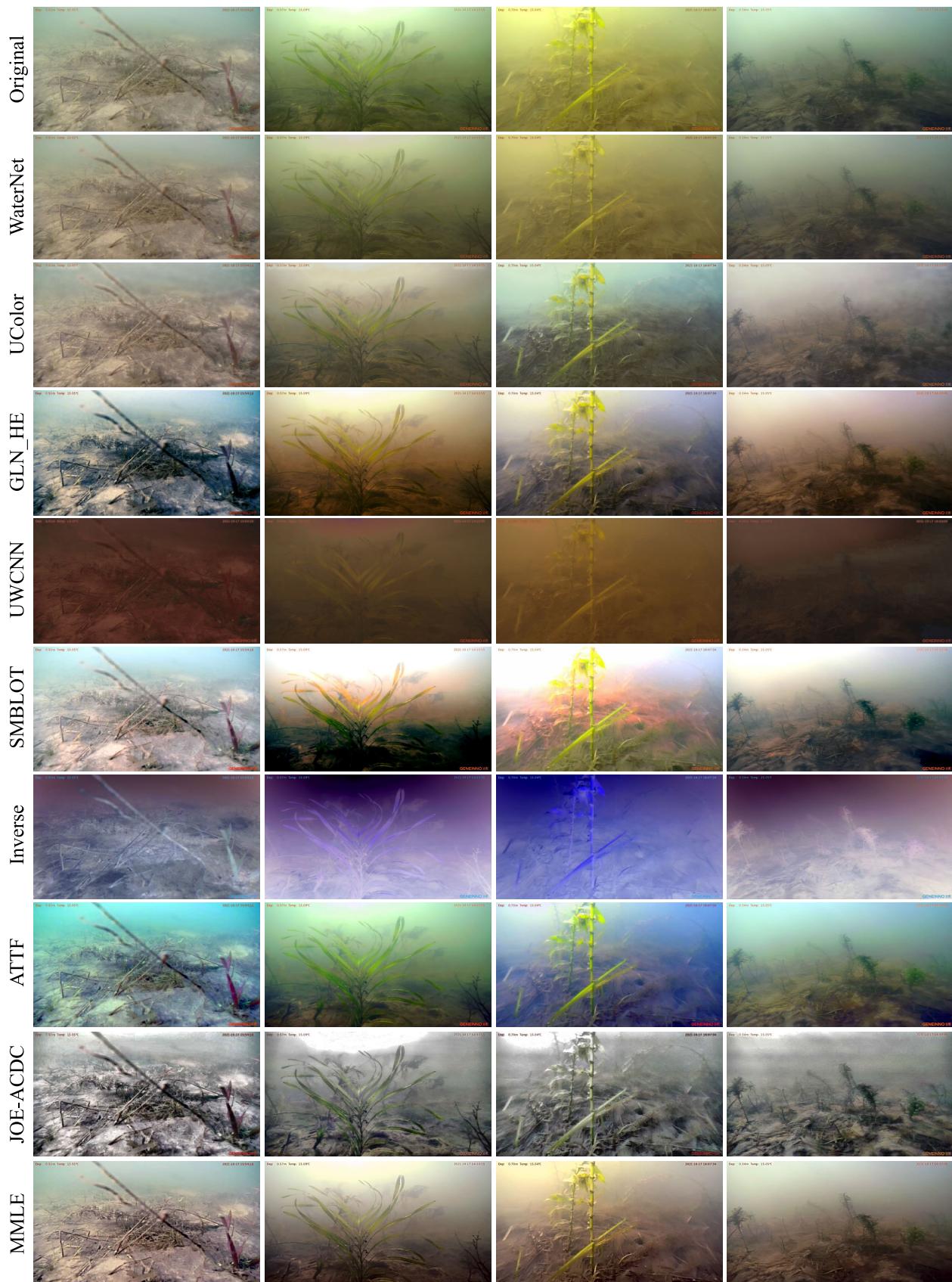
where  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  are the weights of the colorfulness, contrast, and fog density features, respectively. A higher CCF score indicates better image quality.

## V. RESULTS AND DISCUSSION

To ensure a fair evaluation of all the models, a set of images are randomly selected from various depths which are later used as inputs to all the methods under study. The objective of this section is to assess the performance of the selected seven methods on the proposed dataset. In the absence of reference images, we opt-in for the aforementioned popular no-reference image quality assessment (IQA) metrics to evaluate the effectiveness of the selected models.

### A. IMAGE ENHANCEMENT EVALUATION AT DEPTH LESS THAN ONE METER

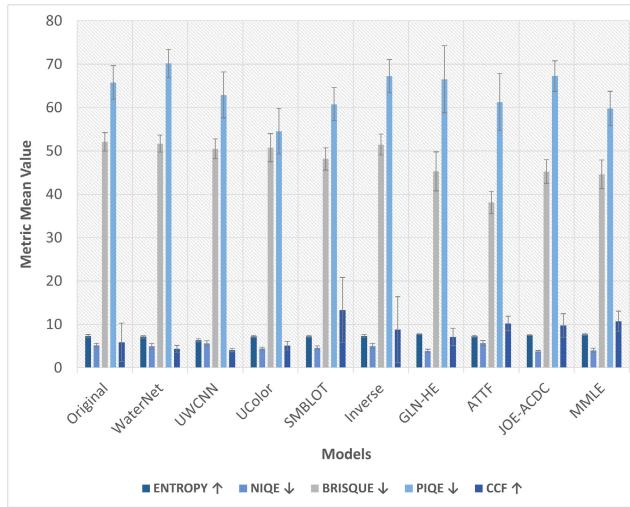
Figure 3 shows samples of the image enhancement results after applying the image enhancement models, considered in this study, to the images captured at a depth of less than 1 meter. Qualitative visual inspection shows that all aforementioned image enhancement methods perform differently for a given input. From Figure 3, it can be seen that WaterNet produces similar output as the input image, which can be witnessed by scores from IAQ metrics presented in Table 2. In contrast, Ucolor which utilizes an encoder and decoder deep network framework that incorporates channel attention mechanism and medium transmission-guiding to extract richer features from color spaces produces a visually pleasing output compared to WaterNet. MMLE also produces a visually pleasing output with balanced color and contrast. The other competing methods introduce undesirable color artifacts where either color distortion or loss of structure details of the underwater scene. For instance, GLN-HE produces an unbalanced color content, although the contrast is improved to a certain degree. SMBLOT is overexposed in the light regions. On the other hand, while it is hard to spot any improvement in the output of UWCNN and Image Inverse, the output images seem to have deteriorated the brightness and contrast quality of the images. Table 2 and Figure 4 show the mean and standard deviation of each metric after applying each model to 37 images from the proposed BRUD dataset. To give better insights into which metrics can be used to separate the enhancement performance of each of the models. It is evident from the obtained results that the GLN-HE model stands out from the rest in terms of image distortion which is measured by NIQE and has improved the contrast of the original images (as seen in Fig. 3), with the highest entropy score of 7.76. On the other hand, the ATT model scores



**FIGURE 3.** Four image samples from the CDUIE dataset at a depth of less than one meter where each row represents the output images of the enhancement model whose name is appended to the left of that row.

**TABLE 2.** Evaluation of original and enhanced images at a depth of less than one meter with bold values indicating the best performance. The symbols  $\uparrow$  and  $\downarrow$  indicate a higher score is better and a lower score is better, respectively.

Metrics	ENTROPY $\uparrow$		CCF $\uparrow$		NIQE $\downarrow$		BRISQUE $\downarrow$		PIQE $\downarrow$	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Original	7.34	0.37	5.87	4.42	5.17	0.46	52.08	2.14	65.81	3.87
WaterNet [10]	7.15	0.26	4.37	0.78	5.01	0.63	51.7	1.98	70.2	3.24
UWCNN [34]	6.34	0.36	4.08	0.39	5.64	0.6	50.49	2.27	62.91	5.28
UColor [42]	7.24	0.24	5.11	0.98	4.44	0.3	50.74	3.24	<b>54.55</b>	5.22
SMBLOT [25]	7.21	0.3	<b>13.34</b>	7.48	4.63	0.41	48.15	2.59	60.76	3.8
Inverse [52]	7.35	0.37	8.8	7.61	5.03	0.59	51.47	2.41	67.27	3.82
GLN-HE [43]	<b>7.76</b>	0.07	7.13	2	3.9	0.37	45.28	4.51	66.55	7.73
ATTF [8]	7.32	0.21	10.21	1.66	5.7	0.57	<b>38.1</b>	2.55	61.31	6.53
JOE-ACDC [32]	7.52	0.1	9.77	2.74	<b>3.8</b>	0.21	45.27	2.74	67.27	3.49
MMLE [33]	7.67	0.2	10.74	2.31	4.04	0.46	44.63	3.3	59.79	3.96



**FIGURE 4.** Graphical representation of the evaluation at depth less than one meter.

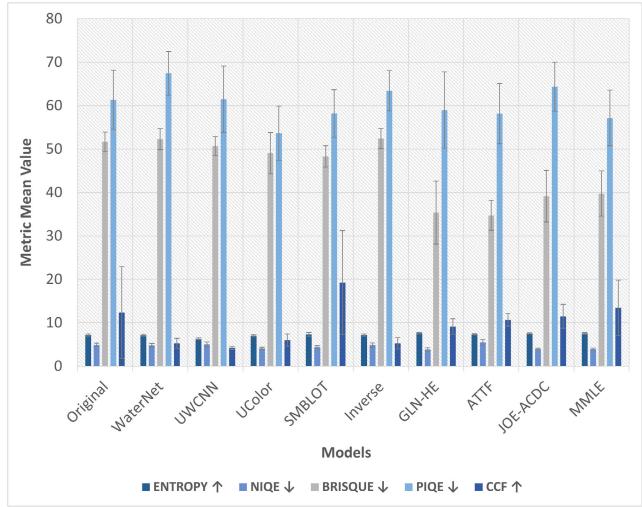
the minimum BRISQUE value, which measures the loss of image naturalness based on local luminance changes due to distortions. Similarly, UColor model has the best PIQE score, which is a perceptual quality, at 54.5 with a pleasing color balance. SMBLOT scores the highest CCF which evaluates the colorfulness, contrast, and fog density features in underwater images.

### B. IMAGE ENHANCEMENT EVALUATION AT DEPTH BETWEEN ONE AND TWO METERS

Going deeper into the water results in low light, more noise, and less contrast. Therefore, image enhancement should improve image quality so that information within the image can be extracted and used. Figure 6 shows the original and enhanced output images of depths from 1 to 2 meters from various models. It is noticed that all enhancement algorithms produced color artifacts to some degree. Table 3 and Figure 5 provide the mean and standard deviation of the calculated IQA metrics for 31 images in this depth range. As shown

**TABLE 3.** Evaluation of original and enhanced images at a depth of one-two meters with bold values indicating the best performance. The symbols  $\uparrow$  and  $\downarrow$  indicate a higher score is better and a lower score is better, respectively.

Metrics	ENTROPY $\uparrow$		CCF $\uparrow$		NIQE $\downarrow$		BRISQUE $\downarrow$		PIQE $\downarrow$	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Original	7.16	0.25	12.36	10.56	4.91	0.4	51.62	2.25	61.33	6.77
WaterNet [10]	7.11	0.25	5.25	1.17	4.8	0.46	52.28	2.41	67.42	5.02
UWCNN [34]	6.33	0.2	4.19	0.35	5.01	0.57	50.69	2.18	61.49	7.64
UColor [42]	7.03	0.28	6	1.38	4.09	0.29	49.06	4.74	<b>53.62</b>	6.23
SMBLOT [25]	7.32	0.43	<b>19.23</b>	11.96	4.36	0.36	48.28	2.43	58.2	5.45
Inverse [52]	7.17	0.25	5.24	1.31	4.88	0.5	52.39	2.31	63.4	4.64
GLN-HE [43]	<b>7.59</b>	0.18	9.13	1.79	<b>3.84</b>	0.41	35.36	7.26	58.97	8.76
ATTF [8]	7.32	0.14	10.67	1.48	5.5	0.66	<b>34.7</b>	3.45	58.15	6.92
JOE-ACDC [32]	7.56	0.12	11.47	2.75	3.91	0.22	39.12	5.94	64.36	5.64
MMLE [33]	7.53	0.22	13.42	6.4	3.92	0.28	39.74	5.2	57.15	6.41

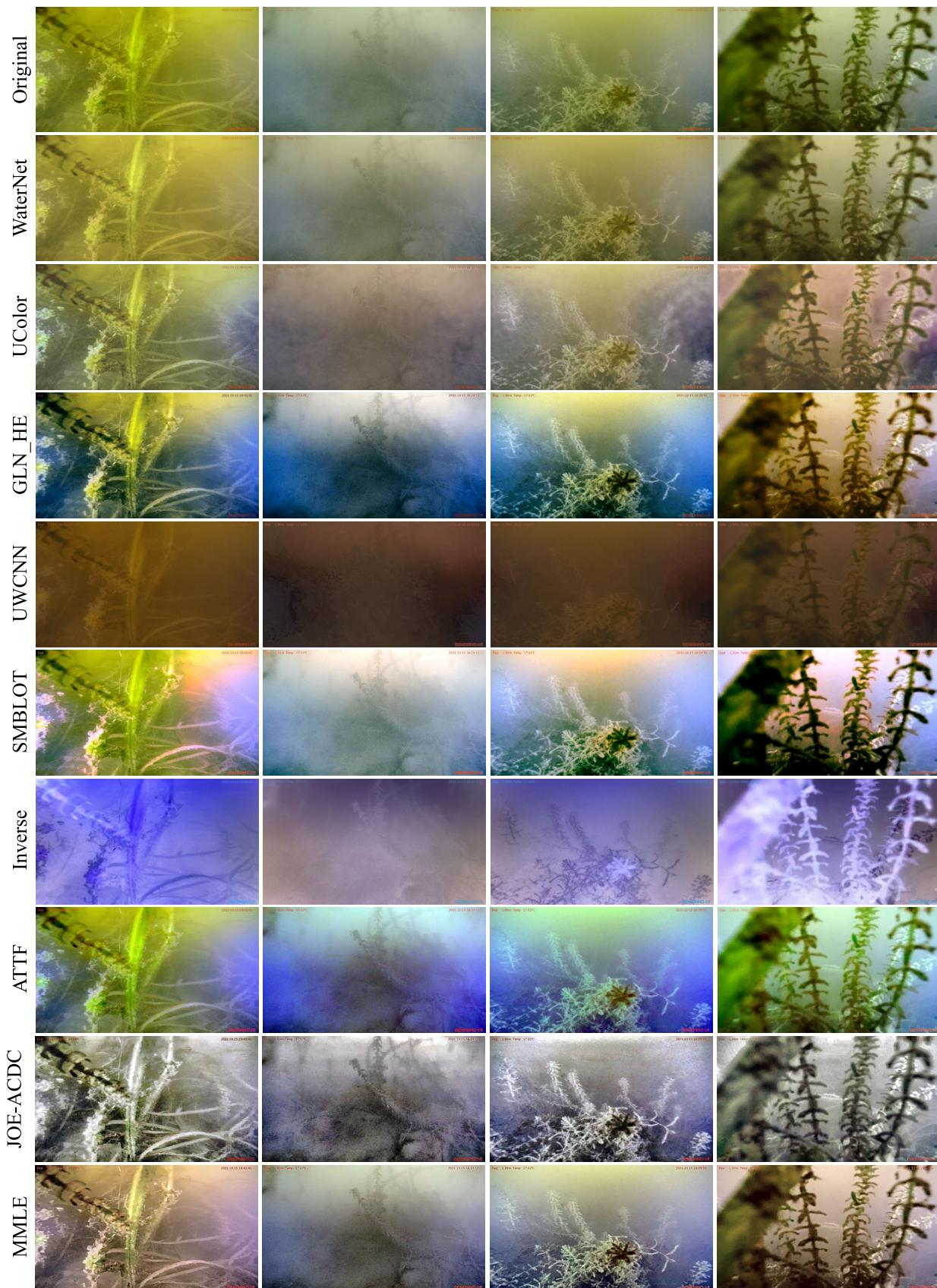


**FIGURE 5.** Graphical representation of the evaluation at depth from one to two meters.

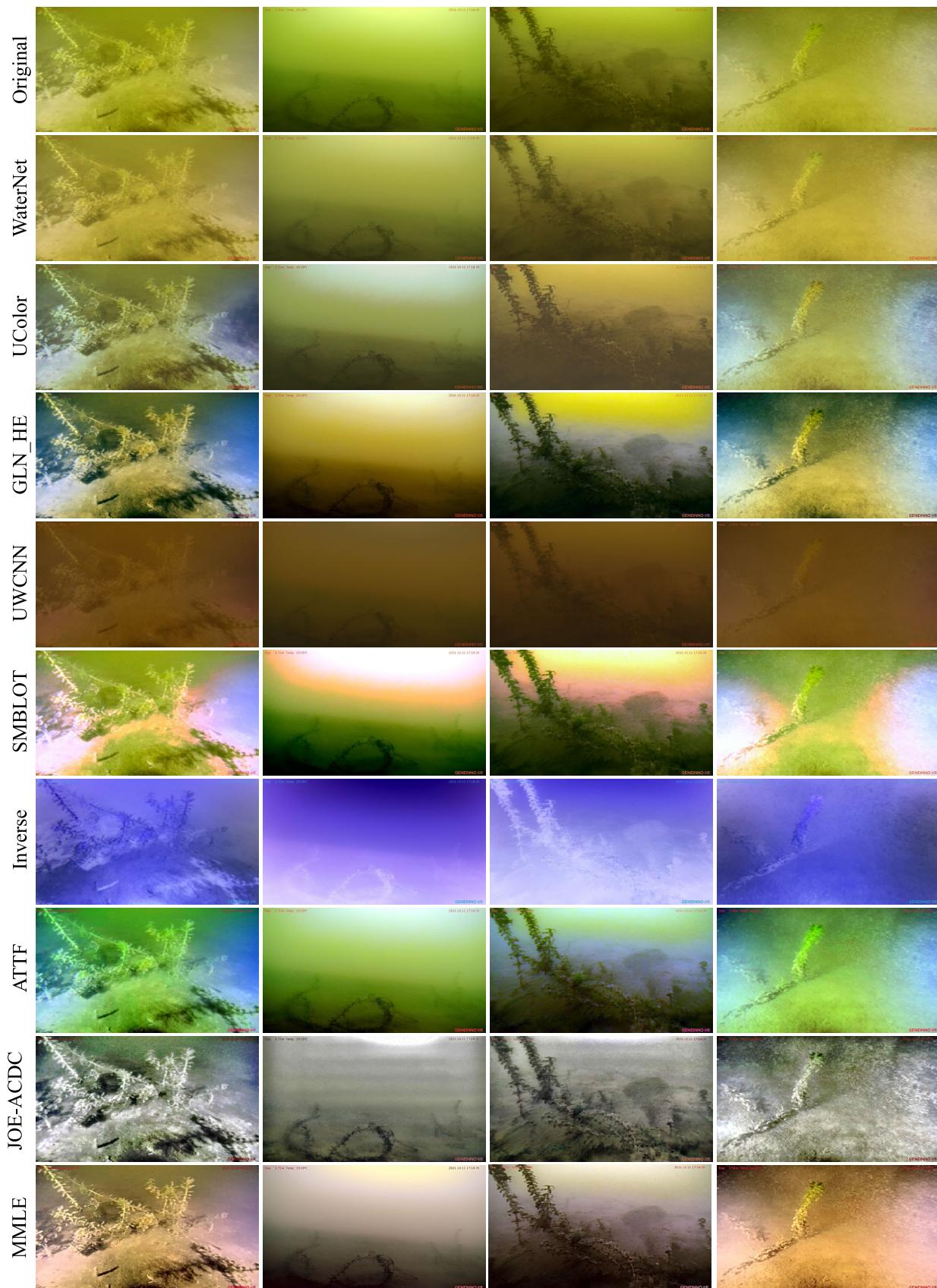
in Table 3, NIQE and ENTROPY scores show that GLN-HE achieves higher image correction than the other competing methods, confirming its good image enhancement performance. SMBLOT has also provided the best CCF score at this depth. On the other hand, it can be observed that the ATTF method has the lowest BRISQUE score which indicates better perceptual quality, while UColor yields the lowest mean of PIQE at 53.62.

### C. IMAGE ENHANCEMENT EVALUATION AT DEPTH BETWEEN TWO AND THREE METERS

Figure 7 shows the input and enhanced images for a depth of 2 to 3 meters underwater. Qualitatively, there are several models with color distortions; however, looking only at the structural contrast and enhancement, it is possible to compare the models' performance. The MMLE method outperforms the GLN-HE method in terms of Entropy score at this depth while the latter has better quality scores, in terms of NIQE and BRISQUE as shown in Table (4) and Figure (8). The results



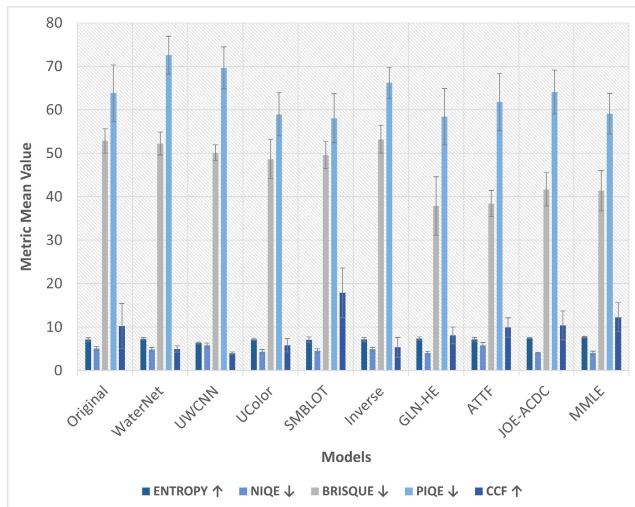
**FIGURE 6.** Four image samples from the CDUIE dataset at a depth of one-two meters where each row represents the output images of the enhancement model whose name is appended to the left of that row.



**FIGURE 7.** Four image samples from the CDUIE dataset at a depth of two-three meters where each row represents the output images of the enhancement model whose name is appended to the left of that row.

**TABLE 4.** Evaluation of original and enhanced images at a depth of two-three meters with bold values indicating the best performance. The symbols ↑ and ↓ indicate a higher score is better and a lower score is better, respectively.

Metrics	ENTROPY ↑		CCF↑		NIQE ↓		BRISQUE ↓		PIQE ↓	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Original	7.13	0.44	10.22	5.22	5.1	0.43	52.85	2.79	63.83	6.47
WaterNet [10]	7.32	0.23	4.97	0.73	4.84	0.48	52.23	2.63	72.57	4.34
UWCNN [34]	6.32	0.2	3.98	0.3	5.85	0.49	50.13	1.8	69.65	4.84
UColor [42]	7.2	0.25	5.84	1.52	4.34	0.47	48.68	4.49	58.96	4.98
SMBLOT [25]	7.08	0.67	<b>17.86</b>	5.71	4.54	0.45	49.58	3.1	<b>58.05</b>	5.65
Inverse [52]	7.15	0.41	5.35	2.26	4.91	0.43	53.24	3.21	66.24	3.57
GLN-HE [43]	7.35	0.39	8.09	1.92	<b>4</b>	0.41	<b>37.86</b>	6.73	58.4	6.46
ATTF [8]	7.16	0.44	9.91	2.23	5.85	0.62	38.44	3.01	61.84	6.57
JOE-ACDC [32]	7.53	0.15	10.39	3.34	4.01	0.18	41.71	3.83	64.1	5.08
MMLE [33]	<b>7.72</b>	0.16	12.27	3.36	4.03	0.42	41.38	4.62	59.14	4.64



**FIGURE 8.** Graphical representation of the evaluation at depth from one-two meters.

also revealed that SMBLOT outperforms all other methods in CCF and PIQE score. These results show that at higher depths, the original images are of poor illumination, contrast, sharpness and color balance. This seems to have affected the tested models, as they were not developed for such scenarios.

## VI. CONCLUSION

In this work, we provided a comprehensive survey and evaluation of the state-of-the-art image enhancement methods using real-world underwater images. We carefully selected data collection sites and the depth of the water to examine the effect of various turbid lake environments and low-light conditions on the performance of the image enhancement methods. Although the proposed CDUIE dataset is relatively small, the included images are captured at a very high resolution and are very challenging to enhance. As expected, by visual inspection, all models suffered from degraded enhancement at higher depths of the water due to poor illumination, contrast, sharpness, and color balance, compared

to lower depths. Throughout systematic evaluation using five popular non-reference evaluation metrics, it is found that: 1) None of the evaluated methods consistently performed best when tested on the CDUIE dataset; 2) The SMBLOT algorithm yielded the best CCF score at all depths; 3) For water depths of less than 1 meter and 1-2 meters, the ATTF method captured better perceptual quality according to scores from BRISQUE. In contrast, for the same depths, the UColor approach obtained the best score of PIQE while GLN-HE achieved the highest entropy score indicating richer textures/details. The conclusion obtained from this study could potentially facilitate future research on underwater image processing. And it is also beneficial for the other scientific community, such as biological sciences, to select proper image enhancement methods for their own study.

In future work, we plan to include more images in our datasets by collecting samples from various water depths and spatial locations. In addition, different resolutions of the images will be also considered and evaluated. Last but not the least, a further evaluation of image enhancement methods will be conducted in terms of their contributions to underwater object (e.g., different types of plants) detection and segmentation.

## REFERENCES

- [1] K. Z. M. Azmi, A. S. A. Ghani, Z. M. Yusof, and Z. Ibrahim, “Natural-based underwater image color enhancement through fusion of swarm-intelligence algorithm,” *Appl. Soft Comput.*, vol. 85, Dec. 2019, Art. no. 105810.
- [2] S. Anwar and C. Li, “Diving deeper into underwater image enhancement: A survey,” *Signal Process., Image Commun.*, vol. 89, Nov. 2020, Art. no. 115978.
- [3] W. Zhang, L. Dong, and W. Xu, “Retinex-inspired color correction and detail preserved fusion for underwater image enhancement,” *Comput. Electron. Agricult.*, vol. 192, Jan. 2022, Art. no. 106585.
- [4] M. Muniraj and V. Dhandapani, “Underwater image enhancement by combining color constancy and dehazing based on depth estimation,” *Neurocomputing*, vol. 460, pp. 211–230, Oct. 2021.
- [5] L. Zhishen, D. Tianfu, and W. Gang, “ROV based underwater blurred image restoration,” *J. Ocean Univ. Qingdao*, vol. 2, no. 1, pp. 85–88, Apr. 2003.
- [6] W. Zhang, L. Dong, X. Pan, P. Zhou, L. Qin, and W. Xu, “A survey of restoration and enhancement for underwater images,” *IEEE Access*, vol. 7, pp. 182259–182279, 2019.
- [7] Y.-T. Peng, K. Cao, and P. C. Cosman, “Generalization of the dark channel prior for single image restoration,” *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2856–2868, Jun. 2018.
- [8] P. Sidike, V. Sagan, M. Qumsiyeh, M. Maimaitijiang, A. Essa, and V. Asari, “Adaptive trigonometric transformation function with image contrast and color enhancement: Application to unmanned aerial system imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 404–408, Mar. 2018.
- [9] R. Liu, X. Fan, M. Zhu, M. Hou, and Z. Luo, “Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 12, pp. 4861–4875, Dec. 2020.
- [10] 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, 2020.
- [11] S. Anwar, C. Li, and F. Porikli, “Deep underwater image enhancement,” 2018, *arXiv:1807.03528*.
- [12] Z. Wang, L. Shen, M. Yu, Y. Lin, and Q. Zhu, “Single underwater image enhancement using an analysis-synthesis network,” 2021, *arXiv:2108.09023*.

- [13] C. N. Brooks, A. G. Grimm, A. M. Marcarelli, and R. J. Dobson, "Multi-scale collection and analysis of submerged aquatic vegetation spectral profiles for Eurasian watermilfoil detection," *J. Appl. Remote Sens.*, vol. 13, no. 3, Aug. 2019, Art. no. 037501.
- [14] C. Brooks, A. Grimm, A. M. Marcarelli, N. P. Marion, R. Shuchman, and M. Sayers, "Classification of Eurasian watermilfoil (*Myriophyllum spicatum*) using drone-enabled multispectral imagery analysis," *Remote Sens.*, vol. 14, no. 10, p. 2336, May 2022.
- [15] R. R. Van Goethem, C. J. Huckins, and A. M. Marcarelli, "Effects of invasive watermilfoil on primary production in littoral zones of north-temperate lakes," *Diversity*, vol. 12, no. 2, p. 82, Feb. 2020.
- [16] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.
- [17] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Mar. 2013.
- [18] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani, "Blind image quality evaluation using perception based features," in *Proc. 21st Nat. Conf. Commun. (NCC)*, Feb. 2015, pp. 1–6.
- [19] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, Oct. 1948.
- [20] X. Fu, J. Wang, D. Zeng, Y. Huang, and X. Ding, "Remote sensing image enhancement using regularized-histogram equalization and DCT," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2301–2305, Nov. 2015.
- [21] Y. Wang, N. Li, Z. Li, Z. Gu, H. Zheng, B. Zheng, and M. Sun, "An imaging-inspired no-reference underwater color image quality assessment metric," *Comput. Electr. Eng.*, vol. 70, pp. 904–913, Aug. 2017.
- [22] J. Zhou, T. Yang, and W. Zhang, "Underwater vision enhancement technologies: A comprehensive review, challenges, and recent trends," *Appl. Intell.*, vol. 53, pp. 3594–3621, Jun. 2022.
- [23] 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.
- [24] H.-H. Chang, "Single underwater image restoration based on adaptive transmission fusion," *IEEE Access*, vol. 8, pp. 38650–38662, 2020.
- [25] W. Song, Y. Wang, D. Huang, A. Liotta, and C. Perra, "Enhancement of underwater images with statistical model of background light and optimization of transmission map," *IEEE Trans. Broadcast.*, vol. 66, no. 1, pp. 153–169, Mar. 2020.
- [26] L. Bai, W. Zhang, X. Pan, and C. Zhao, "Underwater image enhancement based on global and local equalization of histogram and dual-image multi-scale fusion," *IEEE Access*, vol. 8, pp. 128973–128990, 2020.
- [27] J. Xie, G. Hou, G. Wang, and Z. Pan, "A variational framework for underwater image dehazing and deblurring," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 6, pp. 3514–3526, Jun. 2022.
- [28] N. Li, G. Hou, Y. Liu, Z. Pan, and L. Tan, "Single underwater image enhancement using integrated variational model," *Digit. Signal Process.*, vol. 129, Sep. 2022, Art. no. 103660.
- [29] X. Li, G. Hou, K. Li, and Z. Pan, "Enhancing underwater image via adaptive color and contrast enhancement, and denoising," *Eng. Appl. Artif. Intell.*, vol. 111, May 2022, Art. no. 104759.
- [30] G. Hou, Z. Pan, G. Wang, H. Yang, and J. Duan, "An efficient nonlocal variational method with application to underwater image restoration," *Neurocomputing*, vol. 369, pp. 106–121, Dec. 2019.
- [31] P. Zhuang, J. Wu, F. Porikli, and C. Li, "Underwater image enhancement with hyper-Laplacian reflectance priors," *IEEE Trans. Image Process.*, vol. 31, pp. 5442–5455, 2022.
- [32] W. Zhang, Y. Wang, and C. Li, "Underwater image enhancement by attenuated color channel correction and detail preserved contrast enhancement," *IEEE J. Ocean. Eng.*, vol. 47, no. 3, pp. 718–735, Jul. 2022.
- [33] 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.
- [34] C. Li, S. Anwar, and F. Porikli, "Underwater scene prior inspired deep underwater image and video enhancement," *Pattern Recognit.*, vol. 98, Feb. 2020, Art. no. 107038.
- [35] H. Li and P. Zhuang, "DewaterNet: A fusion adversarial real underwater image enhancement network," *Signal Process., Image Commun.*, vol. 95, Jul. 2021, Art. no. 116248.
- [36] P. Liu, G. Y. Wang, H. Qi, C. F. Zhang, H. Y. Zheng, and Z. B. Yu, "Underwater image enhancement with a deep residual framework," *IEEE Access*, vol. 7, pp. 94614–94629, 2019.
- [37] L. Chen, Z. Jiang, L. Tong, Z. Liu, A. Zhao, Q. Zhang, J. Dong, and H. Zhou, "Perceptual underwater image enhancement with deep learning and physical priors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 8, pp. 3078–3092, Aug. 2021.
- [38] 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, Apr. 2020.
- [39] P. Hambarde, S. Murala, and A. Dhall, "UW-GAN: Single-image depth estimation and image enhancement for underwater images," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021.
- [40] R. Han, Y. Guan, Z. Yu, P. Liu, and H. Zheng, "Underwater image enhancement based on a spiral generative adversarial framework," *IEEE Access*, vol. 8, pp. 218838–218852, 2020.
- [41] Y. Guo, H. Li, and P. Zhuang, "Underwater image enhancement using a multiscale dense generative adversarial network," *IEEE J. Ocean. Eng.*, vol. 45, no. 3, pp. 862–870, Jul. 2020.
- [42] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren, "Underwater image enhancement via medium transmission-guided multi-color space embedding," *IEEE Trans. Image Process.*, vol. 30, pp. 4985–5000, 2021.
- [43] X. Fu and X. Cao, "Underwater image enhancement with global-local networks and compressed-histogram equalization," *Signal Process., Image Commun.*, vol. 86, Aug. 2020, Art. no. 115892.
- [44] G. Hou, Y. Li, H. Yang, K. Li, and Z. Pan, "UID2021: An underwater image dataset for evaluation of no-reference quality assessment metrics," *ACM Trans. Multimedia Comput., Commun., Appl.*, Jan. 2023, doi: 10.1145/3578584.
- [45] M. J. Islam, C. Edge, Y. Xiao, P. Luo, M. Mehtaz, C. Morse, S. S. Enan, and J. Sattar, "Semantic segmentation of underwater imagery: Dataset and benchmark," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 1769–1776.
- [46] Q. Qi, K. Li, H. Zheng, X. Gao, G. Hou, and K. Sun, "SGUIE-Net: Semantic attention guided underwater image enhancement with multi-scale perception," 2022, arXiv:2201.02832.
- [47] Y. Wang, W. Song, G. Fortino, L.-Z. Qi, W. Zhang, and A. Liotta, "An experimental-based review of image enhancement and image restoration methods for underwater imaging," *IEEE Access*, vol. 7, pp. 140233–140251, 2019.
- [48] D. Berman, D. Levy, S. Avidan, and T. Treibitz, "Underwater single image color restoration using haze-lines and a new quantitative dataset," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2822–2837, Aug. 2021.
- [49] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 7159–7165.
- [50] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [51] Geneinno T1 Pro Underwater Drone—Geneinno.us. Accessed: Nov. 22, 2022. [Online]. Available: <https://geneinno.us/products/geneinno-t1-pro-underwater-drone>
- [52] Image Enhancement in Python—Code. Tutsplus.com. Accessed: Nov. 22, 2022. [Online]. Available: <https://code.tutsplus.com/articles/image-enhancement-in-python--cms-29289>
- [53] W. Zhang, L. Dong, X. Pan, J. Zhou, L. Qin, and W. Xu, "Single image defogging based on multi-channel convolutional MSRCR," *IEEE Access*, vol. 7, pp. 72492–72504, 2019.
- [54] X. Chen, Q. Zhang, M. Lin, G. Yang, and C. He, "No-reference color image quality assessment: From entropy to perceptual quality," *EURASIP J. Image Video Process.*, vol. 2019, no. 1, pp. 1–14, Dec. 2019.



**ASHRAF SALEEM** (Member, IEEE) received the B.Sc. degree in electrical and computer engineering from Philadelphia University, Jordan, in 2000, and the M.Sc. and Ph.D. degrees in mechatronics engineering from DeMontfort University, U.K., in 2003 and 2006, respectively. His research interests are unified under the theme "Deployment of robotics systems and artificial intelligence in the field of remote sensing." His research interests include solving real-life problems, such as monitoring environmental pollution and developing real-time smart controllers for different engineering systems, which include electromechanical, electro-pneumatic, and piezoelectric-based systems.



**SIDIQE PAHEING** (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Dayton, OH, USA. Currently, he is an Assistant Professor with the Department of Applied Computing, Michigan Technological University. His research interests include image/video processing, machine learning, deep learning, computer vision, and remote sensing. He is an Associate Editor of the *Signal, Image, and Video Processing* (Springer) and *Journal Photogrammetric Engineering and Remote Sensing* (ASPRS). He serves as a guest editor/reviewer for several reputed journals.



**ALI AWAD** received the B.S. degree in computer engineering from Philadelphia University, Jordan, in 2017, and the M.S. degree in computer engineering from the German-Jordanian University, Jordan, in 2019. He is currently pursuing the Ph.D. degree in computational science and engineering with Michigan Technological University, USA. He has excellent practical experience in hardware and low-level control systems. His research interests include autonomous robotics, digital signal processing, and data science.



**NATHER RAWASHDEH** (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Kentucky, in 2005. He was an Associate Professor at the Mechatronics Engineering Department, German Jordanian University, where he spent ten years. He has five years of industry experience. He was a Senior Software Engineer at Laser Color Science and Imaging Department, Lexmark International, Inc., Lexington-Kentucky, and at MathWorks, Inc., Natick-Massachusetts, working on software quality engineering for embedded DSP programming using MATLAB and Simulink. He joined as an Assistant Professor with the Department of Applied Computing, in 2019. His research interests include unmanned vehicles and image analysis.



**NAVJOT KAUR** received the B.C.A. degree from Maharaja Ganga Singh University, India, in 2012, the M.B.A. degree from Punjab Technical University, India, in 2014, and the M.Sc. degree in data science from Michigan Technological University, in 2022. From 2014 to 2018, she worked as a Senior Software Developer at Ariel Software Solutions Pvt. Ltd. She is currently pursuing industrial exposure in data science. Her interests include artificial intelligence, applied statistics, and applied machine learning.

• • •