

Real-world Underwater Enhancement: Challenges, Benchmarks, and Solutions under Natural Light

Risheng Liu, *Member, IEEE*, Xin Fan, *Senior Member, IEEE*, Ming Zhu, Minjun Hou, and Zhongxuan Luo

Abstract—Underwater image enhancement is such an important low-level vision task with many applications that numerous algorithms have been proposed in recent years. These algorithms developed upon various assumptions demonstrate successes from various aspects using *different* data sets and *different* metrics. In this work, we setup an undersea image capturing system, and construct a large-scale *Real-world Underwater Image Enhancement* (RUIE) data set **divided into three subsets**. The three subsets target at three challenging aspects for enhancement, i.e., **image visibility quality, color casts, and higher-level detection/classification**, respectively. We conduct extensive and systematic experiments on RUIE to evaluate the effectiveness and limitations of various algorithms to enhance visibility and correct color casts on images with hierarchical categories of degradation. Moreover, underwater image enhancement in practice usually serves as a preprocessing step for mid-level and high-level vision tasks. We thus exploit the object detection performance on enhanced images as a brand new *task-specific* evaluation criterion. The findings from these evaluations not only confirm what is commonly believed, but also suggest promising solutions and new directions for visibility enhancement, color correction, and object detection on real-world underwater images. The benchmark is available at: <https://github.com/dlut-dimt/Realworld-Underwater-Image-Enhancement-RUIE-Benchmark>

Index Terms—Underwater image enhancement, Benchmark, Visibility, Color cast, Object detection.

I. INTRODUCTION

The development and utilization of ocean resources is of great significance to human beings, demanding remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) equipped with imaging systems for effective investigation. Unfortunately, poor-quality underwater images bring failures to intelligent computer vision systems for visual inspections, environmental sensing, and object detection and recognition. Therefore, it is crucial to develop underwater image enhancement technology for the benefit of more underwater computer vision tasks.

As shown in Fig. 1, there are two major factors leading to the degradation of underwater images. Firstly, the reflected light from underwater scene is absorbed and scattered by

This work is partially supported by the National Natural Science Foundation of China (Nos. 61733002, 61922019, 61672125 and 61572096), LiaoNing Revitalization Talents Program (XLYC1807088), the Fundamental Research Funds for the Central Universities (DUT19TD19) and the Fundamental Research Funds for the Central Universities.

R. Liu, X. Fan, M. Zhu, H. Hou, and Z. Luo are with DUT-RU International School of Information Science, Dalian University of Technology, Dalian, China. R. Liu, X. Fan, M. Zhu, and Z. Luo are also with the Peng Cheng Laboratory, Shenzhen, China, 518052. (E-mail: xin.fan@ieee.org)

Copyright ©2019 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

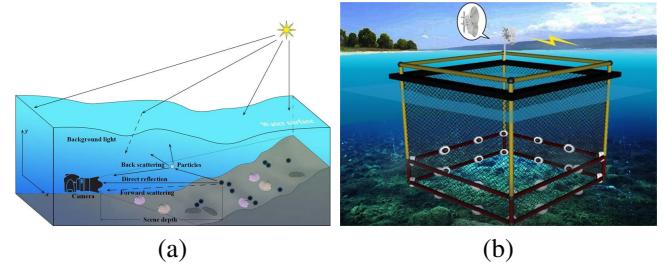


Fig. 1: Schematic diagram of underwater optical imaging and the underwater image capturing system of the proposed RUIE benchmark.

suspending particles in the medium before reaching the camera, resulting in **low contrast and haze-like effects**. Secondly, the attenuation of light, depending on optical wavelength, dissolved organic compounds, and water salinity, causing various degrees of color casts. For examples, underwater images always look bluish or greenish as the red light having a longer wavelength is absorbed more than the green and blue. Accordingly, two essential objectives of underwater image enhancement (UIE) algorithms are to combat the effect of light scattering (similar to dehazing) and to correct color casts. Improving the accuracy for the subsequent higher-level detection/classification tasks is one additional objective of enhancement when the UIE algorithms serve as a preprocessing step. Numerous UIE algorithms have been proposed upon various assumptions to address these degradation issues. Generally, UIE algorithms can be categorized into three types including **model-free, prior-based, and data-driven convolutional neural networks (CNNs)**. Traditional white balance adjustment, histogram equalization and fusion techniques [1], [2], [3], [4] fall into the model-free category. Researchers also develop underwater imaging models to characterize the physical formation process of underwater images. The prior-driven methods employ various types of domain knowledge to estimate depth-dependent transmission map, and further recover images of higher visibility [5], [6], [7], [8], [9], [10], [11], [12]. Recent data-driven frameworks either design deep CNNs [13], [14], [15], [16], [17], or integrate CNNs with physical priors [18], [19] to learn these essential parameters or transmission maps from the degraded inputs. Moreover, UIE has been increasingly applied as a pre-processing step to improve the accuracy of detection/classification tasks [14], with significant advances in underwater high-level tasks in recent years.

Existing UIE algorithms in literature are generally evaluated using *different* data sets, among which many are synthetic, and using *different* metrics upon a certain quality index such as contrast, saturation and luminance. In view of the rapid progress of underwater image enhancement algorithms, it is necessary to enrich a large-scale real-world benchmark for the algorithm evaluation as well as the generation of synthetic images for training data-driven networks¹. Moreover, the evaluation merely on the visibility is not enough for the enhancement with multiple objectives/aspects. Specifically, three major limitations exist in current underwater datasets [6], [3], [10], [20], [13], especially for the purpose of performance evaluation: 1) many existing data sets are unsuitable for evaluating the performance of visibility enhancement, especially those prior model driven methods, as the scene depth is shallow and the scattering effect is subtle in these sets; 2) the scenes and tones of these data sets are relatively monotonous, making it difficult to evaluate how the algorithms work under different illuminations and color casts; 3) there are few marine organisms in images, which limits the application of these databases for evaluating the effectiveness of enhancement to higher-level tasks.

In this study, we build a benchmark data set with *real world sea* images to overcome the three limitations, and also compare the state-of-the-art to suggest the effective and efficient solutions to the problem of underwater image enhancement. Our contributions are summarized as follows:

- We setup a **multi-view imaging system under sea water**, and construct a large-scale underwater benchmark under natural light, the *Real-world Underwater Image Enhancement* (RUIE) data set, **with over 4,000 images**. Table II lists the profile of RUIE, and Fig. 2 shows several image examples. Compared with exiting realistic image sets from underwater scenes, the RUIE includes a large diversity of images, which are divided into three subsets targeting at the evaluation for three objectives of image enhancement algorithms.
- We conduct **substantial and systematic experiments on RUIE** to evaluate the performance of various algorithms **in processing images with multiple degrees of degradation and different types of color casts**. Both quantitative and qualitative analysis demonstrate the advantages and limitations of every algorithm for evaluation. Not only do the findings from these experiments confirm what is commonly believed, but also bring insights for new research directions in underwater image enhancement.
- We also apply a **task-specific evaluation protocol for enhancement algorithms that exploits the object detection accuracy on the enhanced images**. Experimental results reveal that there is not always a strict positive correlation between classification accuracy and visual quality when the underwater images are preprocessed by the existing enhancement methods. This discovery may suggest a new research perspective that considers low-level enhance-

ment and higher-level detection/classification as a whole instead of two cascaded independent processes.

This paper is organized as follows. Section II surveys algorithms addressing the challenges for the enhancement of underwater images. Section III details the RUIE benchmark set, followed by experimental evaluations on RUIE and discussions on the results that suggest solutions to enhancement in Section IV. Section V concludes the paper.

II. UNDERWATER IMAGE ENHANCEMENT ALGORITHMS

Typical UIE algorithms aim to produce a high-quality image that human favor from a single degraded input. These algorithms either increase the visibility or alleviate color casts by combating the light scattering and other ambient circumstance factors during capturing underwater scenes. According to the means of modeling imaging process, we roughly categorize existing UIE algorithms into the following three types. Table I summarizes some algorithm features and experimental schemes of existing representative UIE methods, including the enhancement techniques of model-free methods, the domain knowledge of prior-based methods, the network architecture of data-driven methods, and the experimental datasets and evaluation criteria of these UIE methods.

A. Model-free Methods

Model-free method can be divided into spatial domain method and frequency domain method. Among them, spatial domain methods adjust pixel values of a given image without explicitly modeling the image formation process. The adjustments may be performed in the spatial or transform domain. The spatial domain methods include histogram equalization [29], the Gray World [30], contrast limited adaptive histogram equalization (CLAHE) [2], multi-scale retinex with color restore (MSRCR) [1], automatic white balance [31], and color constancy [32], [33]. Ancuti *et al.* proposed a fusion method [3], and improved it by applying a multiscale fusion strategy to avoid the artifacts of the reconstructed image [34]. Moreover, in more challenging scenes with artificial illumination, Ancuti *et al.* [35] made use of the polarization property to separate the background from the target to achieve the purpose of underwater descattering. The frequency domain methods map image pixels into a specific domain where we exploit the physical properties to perform adjustments [36], [37]. The commonly used transforms include Fourier and wavelets.

The model-free methods can improve the visual quality to some extent, but may accentuate noise, introduce artifacts, and cause color distortions. The transform domain methods perform well in smearing noise, but suffer from low contrast, detail loss, and color deviations. Due to the complexity of underwater environment and illumination conditions, these enhancement techniques, merely relying on the observed information, can hardly recover high quality images from underwater degradation.

¹Considering the emerging and popularity of generative adversarial networks (GANs), real-world images are required for discriminators in GANs to generate ‘synthetic’ training examples.

TABLE I: Overview of representative UIE algorithms. The labels “R”, “S”, and “CC” in the “test data” column represent real world, synthetic, and ColorChecker images. The “Criterion” column lists the metrics of Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), Contrast-to-noise ratio (CNR), Patch-based Contrast Quality Index (PCQI) [21], UCIQE, UIQM, and Blind/ Referenceless Image Spatial Quality Evaluator (BRISQUE) [22].

Model-free methods					
Methods	Enhancement technique	Test data	Criterion	Code	
UCM [23]	Unsupervised color balance and histogram stretching		R	Histogram distribution	✓
MSRCR [1]	Multiscale retinex with color restoration		R	X	✓
CLAHE [2]	Contrast-limited adaptive histogram equalization		R	X	✓
CLAHE-mix[24]	Mixture RGB and HSV CLAHE		R	MSE, PSNR	X
Fusion [3]	White Balance, bilateral filtering, image fusion		R, CC	SIFT	✓
Ghani <i>et al.</i> [25]	minimizes under-enhanced and over-enhanced areas		R	Entropy, MSE, PSNR	X
Prior-based methods					
Methods	Physical prior	Post process	Test data	Criterion	Code
BP [5]	Difference of radiance attenuation	X	R, CC	X	✓
Drews <i>et al.</i> [9]	Underwater DCP on g, b	X	R, CC	RMSE	✓
NOM [8]	Underwater DCP	X	R	X	✓
RCP [10]	Red channel prior	X	R	Edge number, Gradient ratio	X
WCID [6]	DCP	Channel compensation	R, CC	X	✓
Yang <i>et al.</i> [26]	DCP with median filter	Unsupervised correction [23]	R	X	X
Peng <i>et al.</i> [12]	Blurriness and Light Absorption	Histogram equalization	R, S	PSNR, SSIM, BRISQUE, UIQM	X
UHP [7]	Color distribution	Global white balance	R	RGB Median angle	✓
Lu <i>et al.</i> [27]	Underwater DCP on r, b	Channel compensation	R	PSNR, CNR, SSIM	X
LDP [28]	Dark channel and light prior	Histogram distribution prior	R, S, CC	MSE, PSNR, PCQI, UCIQE	X
Data-driven methods					
Methods	Network		Test data	Training data	Code
Ding <i>et al.</i> [15]	Transmission estimation CNN and background light estimation CNN		R	Synthetic images	X
Cao <i>et al.</i> [16]	Transmission estimation multi-scale CNN		R, S	Synthetic images	X
GAN-RS [14]	End-to-end supervised GAN		R	Filter-enhanced images	✓
waterGAN [13]	End-to-end encoder-decoder CNN		R, S	GAN-synthesized images	✓
DPATN [18]	Transmission estimation lightweight learning-based framework		R	Synthetic images	✓

B. Prior-based Methods

The model based methods explicitly characterize the physical imaging process, and estimate the parameters of the imaging model from the observation and various prior assumptions. The clear underwater scene can be restored by inverting this degradation process. One common underwater imaging model is derived from the Jaffe-McGlamery model as [38], [39]:

$$I_c(x) = J_c(x)t_c(x) + A_c(1 - t_c(x)), c \in \{r, g, b\}, \quad (1)$$

where $I_c(x)$ is the observed degraded image, $J_c(x)$ is the clear scene radiance to be recovered, c represents a color channel. There are two critical parameters for the restoration, i.e., the global atmospheric light A_c , and transmission matrix $t_c(x)$. The transmission denotes the portion of the scene radiance that reaches the camera, defined as:

$$t_c(x) = e^{-\beta_c d(x)}, \quad (2)$$

where $d(x)$ represents the scene depth, and β_c is the scattering coefficient of the transmission medium depending on water quality, depth and salinity, for underwater images. Most recent

underwater enhancement methods estimate these two key parameters A_c and $t_c(x)$ in order to improve visibility, and correct color casts using traditional model-free techniques, e.g., color balance or histogram equalization.

Many UIE algorithms attempt to extend the prior model based dehazing algorithms to underwater scenes by noting that the underwater imaging model shares commonalities with the one for hazy images. The original priors for dehazing have to adapt the serious attenuation of red light through water so that those dehazing algorithms are applicable to underwater scenarios. For instance, several prior-based UIE methods are derived from the dark channel prior (DCP) [40], one of the most effective means to estimate the transmission (depth) map of a hazy image. Chiang *et al.* modified DCP by compensating the attenuation to restore the color balance [6]. Drews-Jr *et al.* applied the modified underwater DCP only to the blue and green channels [9]. Wen *et al.* [8] estimate transmission by underwater dark channel with a new optical model (NOM). Galdran *et al.* proposed the Red Channel Prior (RCP) upon DCP by characterizing the attenuation on the red channel [10].



(a) Underwater Image Quality Set (UIQS).

From left to right, images from the five subsets A~E are ranked according to a non-reference image quality metric, and the corresponding image quality successively decreases.



(b) Underwater Color Cast Set (UCCS). The set is divided into “Green”, “Green-blue”, and “Blue” according to the degree of color cast.



(c) Underwater Higher-level Task-driven Set (UHTS). Various sea life appear in the images of this set. The five subsets A~E are also ranked according to the image quality.

Fig. 2: Example images from the three sets of RUIE.

Researchers have also proposed various physical priors designated to underwater images, other than those derived from DCP. Carlevaris *et al.* proposed Bianco Prior (BP) [5] which took advantage of the characteristics of channel discrepancies to estimate the underwater transmission. Wang *et al.* developed a maximum attenuation identification method (MAI), which only used the red channel information to generate the depth map and atmospheric light [41]. Peng *et al.* presented a depth estimation method using image blurriness and light absorption [12]. Wang *et al.* proposed an adaptive attenuation-curve prior applicable to both UIE and dehazing [42]. Zhou *et al.* [43] handled the scattering and absorption problems of light with different wavelengths based on the color-line model. Berman *et al.* [7] proposed underwater Haze-line Prior (UHP) and estimated the attenuation ratios of the blue-red and blue-green color channels.

Nevertheless, one of the common drawbacks of the aforementioned prior-based UIE algorithms lie in that these priors are invalid to some specific environmental/scenery configurations and/or severe color casts. For example, it is well known that DCP is inapplicable to white objects or regions.

C. Data-driven Enhancement Neural Networks

Very recently, deep CNNs on large-scale datasets have delivered excellent performance in many high-level computer

vision and recognition tasks [44], [45], which promoted the development of data-driven UIE methods. Meanwhile, various complex factors in underwater imaging, including dynamic water flow, color deviations, and low illuminations, require a more complex network structure and/or a well-designed loss function.

Li *et al.* designed the WaterGAN to synthesize training examples and an end-to-end network consisting of a depth estimation module followed by a color correction module [13]. Chen *et al.* combined a filtering-based with a GAN-based restoration scheme that adopted a multi-stage loss strategy for the training [14]. Ding *et al.* used a new white balance algorithm to correct color cast, then adopted traditional CNN to estimate background light and transmission, and finally restored the underwater images based on image formation model. Cao *et al.* proposed a 5-layer ConvNet and a multi-scale deep network to estimate background light and scene depth respectively. Recently, Hou *et al.* proposed an underwater residual network to jointly optimize transmission and correct color casts [17]. Different from the above end-to-end UIE networks, Liu *et al.* established the basic propagation scheme based on the fundamental image modeling cues and then introduced CNNs [19] or a lightweight residual learning framework Data and Prior Aggregated Transmission Network (DPATN) [18], to integrate both physical priors and data driven

cues for solving various image enhancement tasks including the underwater aspects.

The three categories of UIE algorithms were evaluated on different data sets in terms of various metrics for visual quality. A comprehensive study is still highly demanded on the extent to which these UIE algorithms achieve the objectives of improving visibility, correcting color casts, and increasing accuracy for the following higher-level vision tasks.

III. THE PROPOSED DATA SET

A successful UIE algorithm has to address one or all of the following issues in underwater imaging including visibility degradation, color casts, and accuracy decrease of higher-level detection tasks. These multiple objectives of UIE algorithms require a diverse portfolio of testing examples in a benchmark for UIE. The images to evaluate the capability of visibility improvements typically need a larger scene depth² so that the degradation effects caused by water scattering are evident. On the other hand, the evaluation of color correction performance demands a data set containing a wide range of color tones. Moreover, the calculation of detection/classification accuracy demands object/target labels as the groundtruth in the benchmark. However, most existing underwater image data sets generally target at evaluating either one or two of the three objectives for UIE algorithms. Therefore, the establishment of a large-scale, diverse, and task-specific database is important for fair and comprehensive comparisons of UIE algorithms. Additionally, this type of benchmark may lay a ‘data’ foundation for training intelligent underwater vehicles equipped with automatic computer vision systems.

We setup a multi-view underwater image capturing system with twenty-two water-proof video cameras in order to collect image examples for our RUIE benchmark, this system is simplified and modeled as shown in Figure 1 (b). These cameras were mounted along a 10 meters by 10 meters square frame, and placed 0.5 meter above the sea bed close to the Zhangzi island in the Yellow Sea, of which the geographic coordinate locates at ($N39.186, E44.625$). We carefully adjusted the view angles of the cameras so that the maximum scene depths may vary from 0.5 to 8 meters. We installed no external lighting system and captured all videos under natural light during two time slots from 8 AM to 11 AM and 1 PM to 4 PM each day between September 21st and 22nd, 2017. The water depth varied from 5 to 9 meters owing to the periodic tide. The changing of lighting and water depth produce varying color tones. More importantly, this area maintains a natural marine ecosystem containing abundant sea life including fish, sea urchin, sea cucumber, scallops, etc. This ecosystem makes it possible for us to provide labels for underwater object detection tasks.

The captured videos over 250 hours cover a wide range of diversities on illuminations, depths of fields, blurring degrees, and color casts. We manually picked about four thousand images, and divided them into three subsets according to specific tasks of UIE algorithms. We list their respective profiles and objectives for evaluation as follows.

²The depth is defined as the distance from the scene to the imaging plane.

Underwater Image Quality Set (UIQS): This subset is used to test UIE algorithms for the improvement of image visibility. Specifically, we assessed the quality of images according to the underwater color image quality evaluation (UCIQE) metric [46], and ranked these images by their corresponding UCIQE scores. The UCIQE metric is a linear combination of chroma, saturation, and contrast of underwater images. Then we equally divided them into five subsets, denoted as [A, B, C, D, E], in descending order of the UCIQE values, in order to facilitate testing the performance of different algorithms in various underwater conditions. Figure 2 shows image examples with different levels of image quality.

Underwater Color Cast Set (UCCS): This set aims to evaluate the ability of correcting color casts for UIE algorithms. According to the average value of the blue (b) channel (red-green bias) in the CIElab color space, we produced the 300-image UCCS set. It contains three 100-image subsets of the bluish, greenish and blue-green tones. The corresponding examples are shown in the second row of Fig. 2.

Underwater Higher-level Task-driven Set (UHTS): One of the tasks of UIE is to improve the accuracy for subsequent higher-level object detection/classification when UIE has been increasingly serving as a pre-processing step. It is necessary to investigate the effectiveness of UIE algorithms from the perspective of high-level tasks. Therefore, we specially build up the task-driven dataset UHTS, considering the increasing demands of UIE for numerous computer vision applications.

The UHTS set contains 300 images embodying several types of sea life for the purpose of evaluating the effects of UIE algorithms to higher-level computer vision tasks, e.g., classification and detection. Currently, we label the bounding boxes and types of three classes of sea life, i.e., scallop, sea cucumbers and sea urchins, in these underwater images. The detection and classification of these three types greatly challenge recent computer vision algorithms because their appearance is quite similar to the ambient environment thus difficult to distinguish. The accuracy of higher-level algorithms is sensitive to image quality. Additionally, these sea lives are of great interest for marine ecology. Therefore, we provide these labels for our higher-level task driven set. Furthermore, similar to UIQS, in order to explore the impact of image quality on the detection accuracy, the images of UHTS are sorted into five subsets, each containing sixty images.

Since enhanced images are often fed to higher-level computer version tasks, it is noted that the objective of UIE is not only pixel-level or perceptual-level quality improvements, but also the utility of enhanced images in given semantic analysis tasks. We thus propose the higher-level task-driven evaluation for UIE algorithms, and study the problem of object detection in the presence of visibility degradation as an example. We trained an underwater object detection CNN using the network structure of YOLO-v3 as the baseline [47]. The training set consists of 1,800 labeled pictures captured from shallow waters with the depth of less than three meters. We apply the trained CNN to detect three types of objects from the enhanced results given by various UIE algorithms. The detection accuracy is evaluated in terms of the mean Average Precision (mAP).

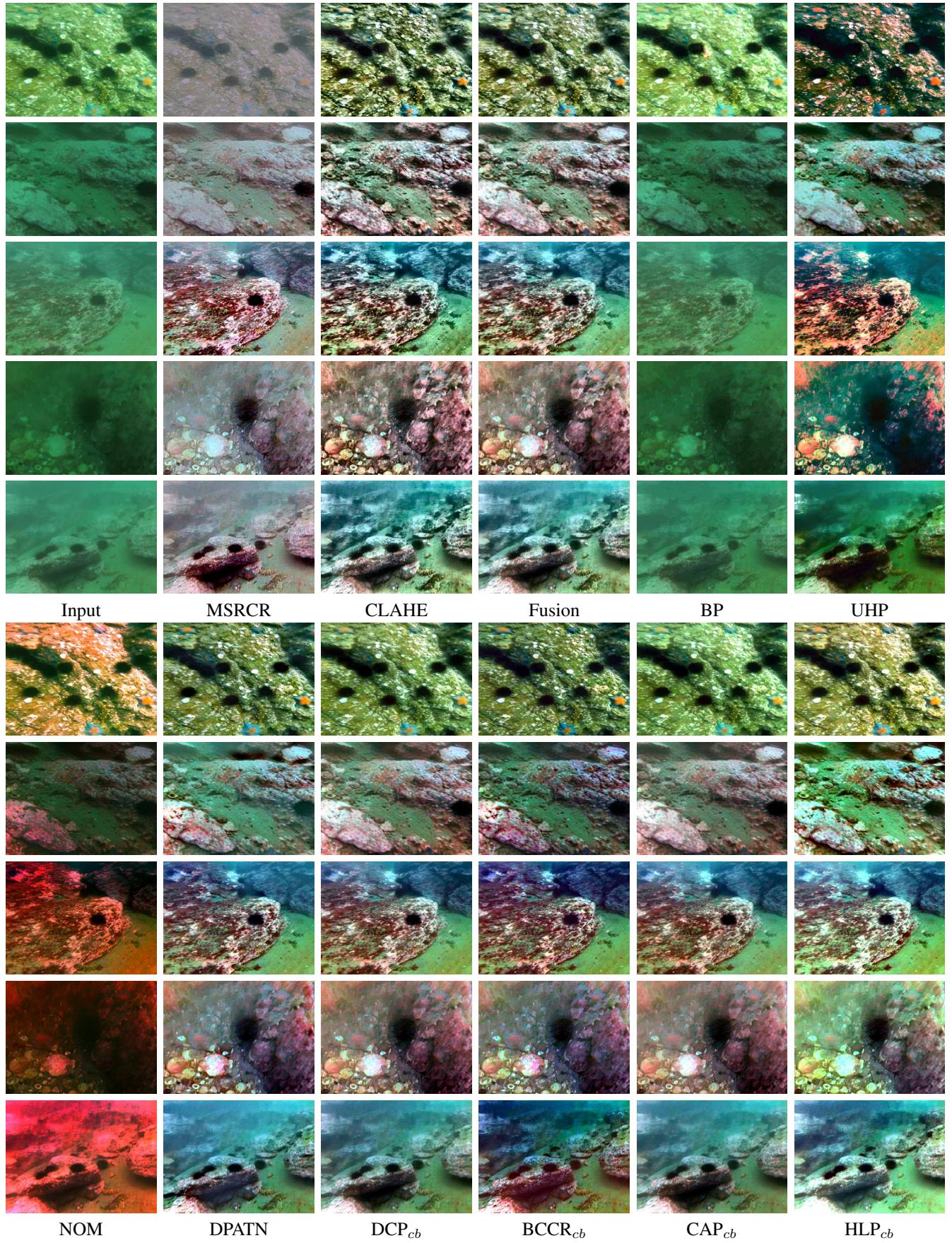


Fig. 3: Comparison on UIQS sub-dataset. In sequence the quality levels of the five inputs are A , B, C, D and E respectively.

TABLE II: Subsets of RUIE

Subset	Image Number
Underwater image quality set (UIQS)	3630 (726×5)
Underwater color cast set (UCCS)	300 (100×3)
Underwater task-driven testset (UHTS)	300 (60×5)

IV. EVALUATION RESULTS AND DISCUSSIONS

We applied our RUIE benchmark to quantitatively and qualitatively evaluate *eleven* representative underwater image enhancement algorithms³ including original underwater enhancement methods MSRCR [1], CLAHE [2], Fusion [3], BP [5], UHP [7], NOM [8] and DPATN [18], and underwater extended application of dehaze methods DCP [40], Boundary Constrained Context Regularization [48], Color Attenuation Prior [49] and Haze Line Prior [50]. They are denoted as DCP_{cb}, BCCR_{cb}, CAP_{cb} and HLP_{cb}, respectively, and the subscript *cb* indicates that the algorithm was cascaded with color balance as a post-processing step. Table I gives more information about these algorithms. In this section, we provide experimental results on the three subsets of our benchmark, and also discuss the results especially on the higher-level task-driven subset. For the sake of evaluation in reasonable time, we resized all images to 300×400, that produces stable outputs for both enhancement and detection, in the following experiments.

A. Comprehensive Image Quality Comparisons on UIQS

Qualitative comparisons: We compared the capabilities of the eleven methods to improve the image visibility on the subset UIQS. The qualitative comparison in Fig. 3 demonstrate that most methods are able to achieve better enhancement for images with the quality levels A,B and C where the underwater scattering effect is subtle. The results of MSRCR appear appropriate color tones but not enough saturation and contrast. CLAHE and Fusion notably improve the image brightness, saturation and contrast, but the lack of an imaging model leads to evident hazy effects. Additionally, these three methods are unable to adaptively work in various scenarios due to their fixed parameter settings. BP is effective for removing the effects caused by light scattering, but cannot deal with color cast well, especially when the water is greenish. UHP generates over-saturation and excessive contrast, smearing image details. DPATN as well as those algorithms with dehazing priors extended to underwater imaging can effectively remove haze-like effects and produce more natural scene.

For the underwater images with the quality grades of D and E, the algorithm of MSRCR with a fixed set of parameters even aggravates the scattering effect. CLAHE and Fusion improve the contrast of these categories of images, but introduce considerable artifacts and keep the severe haze-like effects. BP works little on improving these severely degraded inputs. Another prior-based method UHP can yield relatively clearer results, especially for scenes farther from the camera. NOM tends to make the results severely reddish. It is worth

³The implementations of compared algorithm are available at <https://github.com/dlut-dmft/Underwater-image-enhancement-algorithms>

noting that these reddish results still exhibit high UICM scores because the UICM measure favors the reddish hue. In contrast, DPATN, HLP_{cb} and BCCR_{cb} are able to remove the haze-like effects well in these challenging images. Among the three, HLP_{cb} improves the image brightness recovering more image details, and BCCR_{cb} performs the best on improving visibility.

Quantitative comparisons: We employ two *non-reference* metrics for the quantitative assessment of underwater image quality as no ground truth scene is available as the reference for real world sea images. One is the underwater image quality measure (UIQM) [51], consisting of three underwater image attribute measures, i.e., the underwater image colourfulness measure (UICM), underwater image sharpness measure (UIS-M), and underwater image contrast measure (UIConM). UIQM is expressed as a linear combination as:

$$\text{UIQM} = c_1 \times \text{UICM} + c_2 \times \text{UIS-M} + c_3 \times \text{UIConM}, \quad (3)$$

where c_1 , c_2 and c_3 are the scale factors. We set $c_1 = 0.0282$, $c_2 = 0.2953$, and $c_3 = 3.5753$ as the original paper [51]. The other one is the underwater color image quality evaluation (UCIQE) [46], which uses a linear combination of the chroma, saturation, and contrast of underwater images in the CIElab color space. The UCIQE score can be obtained as:

$$\text{UCIQE} = c_1 \times \omega_c + c_2 \times \text{con}_l + c_3 \times \mu_s, \quad (4)$$

where ω_c is the standard deviation of chroma; con_l is the contrast of brightness; μ_s is the average of saturation; c_1 and c_2 are the scale factors. Again, we set $c_1 = 0.4680$, $c_2 = 0.2745$, and $c_3 = 0.2576$ as the original paper [46].

Table III gives the quantitative scores of the eleven UIE algorithms averaged on all images of the UIQS data set. In addition to the two comprehensive quality metrics, UICQM and UCIQE, we also list the values of the components of UICQM including UICM, UIS-M and UIConM, reflecting colourfulness, sharpness, and contrast of an image, respectively. The highest two values are marked bold. It should be noted that, the excessive redness in results of NOM boosts extremely high UICM and UICQM scores. Unfortunately, these results are so visually poor that they are considered as outliers and marked in green. We can see from Tab. III that the following algorithms perform over the others in terms of one single metric: MSRCR and BCCR_{cb} have advantages in balancing color casts; CLAHE and prior-CNN-aggregated DPATN can generate sharper images; Fusion, DPATN and BCCR_{cb} output images with higher contrasts. In terms of the two comprehensive metrics, DPATN, DCP_{cb}, BCCR_{cb}, and the three model-free methods (MSRCR, CLAHE, and Fusion) stably improve image quality on all the five categories, and the gaps over the others are more evident on the categories of C, D and E with lower image quality. Among them, CLAHE and BCCR_{cb} obtain higher UIQM scores, while UHP and BCCR_{cb} yield higher UCIQE scores.

As a summary, the prior-based BP algorithm is more suitable to process images with less degradation, while those model-free based methods including Fusion and CLAHE, and prior-aggregated DPATN are the better choices for severely degraded images.

TABLE III: Non-reference Underwater Image Quality Evaluation of algorithms on UIQS.

Metric	Input	MSRRCR	CLAHE	Fusion	BP	UHP	NOM	DPATN	DCP _{cb}	BCCR _{cb}	CAP _{cb}	HLP _{cb}	
E	UICM	-74.29	-2.07	-20.80	-21.75	-69.04	-22.82	42.45 ⁴	-16.18	-10.11	0.057	-14.12	-31.96
	UISM	1.657	4.925	5.532	4.784	0.162	3.251	2.317	4.947	3.042	3.805	3.098	3.650
	UIConM	0.459	0.712	0.807	0.818	0.563	0.743	0.684	0.811	0.778	0.833	0.759	0.782
	UIQM	0.035	3.942	3.933	3.725	0.114	2.974	4.327	3.902	3.396	4.103	3.232	2.971
	UCIQE	0.240	0.493	0.451	0.469	0.264	0.500	0.486	0.479	0.486	0.501	0.475	0.462
D	UICM	-77.51	-0.718	-16.32	-16.72	-69.59	-19.40	35.48	-15.24	-2.001	5.208	-5.765	-28.98
	UISM	1.825	4.852	5.598	4.845	0.257	3.369	2.298	4.857	3.136	3.805	3.193	3.798
	UIConM	0.512	0.717	0.805	0.825	0.619	0.745	0.665	0.806	0.781	0.822	0.759	0.784
	UIQM	0.184	3.975	4.071	3.909	0.328	3.112	4.058	3.886	3.661	4.210	3.493	3.108
	UCIQE	0.266	0.483	0.452	0.470	0.293	0.500	0.477	0.475	0.491	0.505	0.476	0.472
C	UICM	-82.01	-1.497	-19.50	-19.19	-70.57	-18.59	37.84	-14.83	-1.85	6.256	-7.920	-31.69
	UISM	1.836	4.663	5.525	4.719	0.262	3.376	2.251	4.793	3.094	3.793	3.180	3.731
	UIConM	0.520	0.703	0.791	0.811	0.635	0.742	0.652	0.802	0.766	0.809	0.740	0.758
	UIQM	0.089	3.848	3.909	3.753	0.358	3.124	4.061	3.863	3.600	4.190	3.363	2.918
	UCIQE	0.283	0.475	0.453	0.470	0.311	0.501	0.478	0.478	0.493	0.502	0.472	0.478
B	UICM	-84.82	-2.079	-16.27	-17.00	-71.70	-18.49	29.53	-16.97	1.695	8.351	-10.65	-28.55
	UISM	1.889	4.431	5.422	4.548	0.351	3.235	2.289	4.814	3.076	3.671	3.131	3.608
	UIConM	0.539	0.688	0.778	0.799	0.637	0.727	0.639	0.799	0.754	0.787	0.723	0.725
	UIQM	0.092	3.709	3.925	3.719	0.361	3.034	3.794	3.651	4.134	3.800	3.210	2.852
	UCIQE	0.301	0.463	0.455	0.471	0.338	0.508	0.466	0.477	0.496	0.500	0.472	0.486
A	UICM	-77.19	0.971	-23.68	-26.65	-64.64	-22.23	11.58	-17.79	-16.30	-7.198	-32.91	-36.94
	UISM	3.291	3.650	6.085	5.057	1.131	4.153	3.513	3.870	4.486	4.805	4.252	4.729
	UIConM	0.696	0.630	0.785	0.841	0.715	0.755	0.721	0.790	0.804	0.815	0.787	0.738
	UIQM	1.285	3.358	3.937	3.749	1.068	3.300	3.940	3.467	3.738	4.129	3.142	2.991
	UCIQE	0.362	0.375	0.430	0.449	0.413	0.504	0.457	0.489	0.484	0.486	0.462	0.485

Discussion: Underwater image quality assessment

There exist discrepancies between the qualitative images in Fig. 3 and the quantitative scores in Table III. The algorithms producing results with higher scores do not always exhibit favorable appearance for human visual perception. Additionally, UCIQE and UIQM may yield inconsistent assessments on images. For example, NOM always tends to produce severe reddish color shift due to excessive enhancement, resulting in higher UIQM scores, especially on images of lower visual quality. The metric UCIQE favors the results with high contrasts, even for those of UHP showing unnatural excessive contrast. One possible explanation lies in that both UCIQE and UIQM metrics focus on the intensities of low-level features, e.g., contrast and saturation, but ignores higher semantic or prior knowledge from human perception. Also, the calculation of these metrics fails to test whether the intensity values fall within a reasonable range over the whole image. Therefore, the development of an appropriate and objective metric for underwater image assessment is still an open issue in this field.

Recently, data-driven CNNs comprising the information from human labels for non-reference quality assessment of natural images have rapidly developed and achieved remarkable performance [52], [53], [54]. It is a promising direction to investigate how to immigrate the deep architectures for natural images to underwater scenarios. Training examples also play

an important role in any deep learning approaches. From this respect, the real-world images showing different quality levels in our RUIE data set may contribute to this type of studies.

B. Color Correction Comparisons on UCCS

Different lighting conditions, water depths, and levels of salinity produce significant changes on color tones of underwater images. Correcting to a natural tone is one of the important objectives for UIE. Therefore, we construct UCCS having great diversities of color tones, and use images of UCCS to evaluate the capabilities of UIE algorithms for color correction.

Qualitative comparisons: Figure 4 demonstrates representative resultant images of the eleven methods performing on UCCS. At this point, we focus on the ability to correct color distortions. MSRRCR can correct both greenish and bluish tones well. As for CLAHE and Fusion, the ability to handle greenish tones is superior than that to blue tones. BP enhances the contrast in bluish scenes, but tends to failure on greenish pictures. UHP may produce partial darkness, and NOM always generates over-corrected reddish results. Among the four prior-based algorithms with direct applications of dehazing priors,

⁴Excessive redness in results always boosts extremely high UICM and UIQM scores, these results are so visually poor that they are considered as outliers and marked in green.

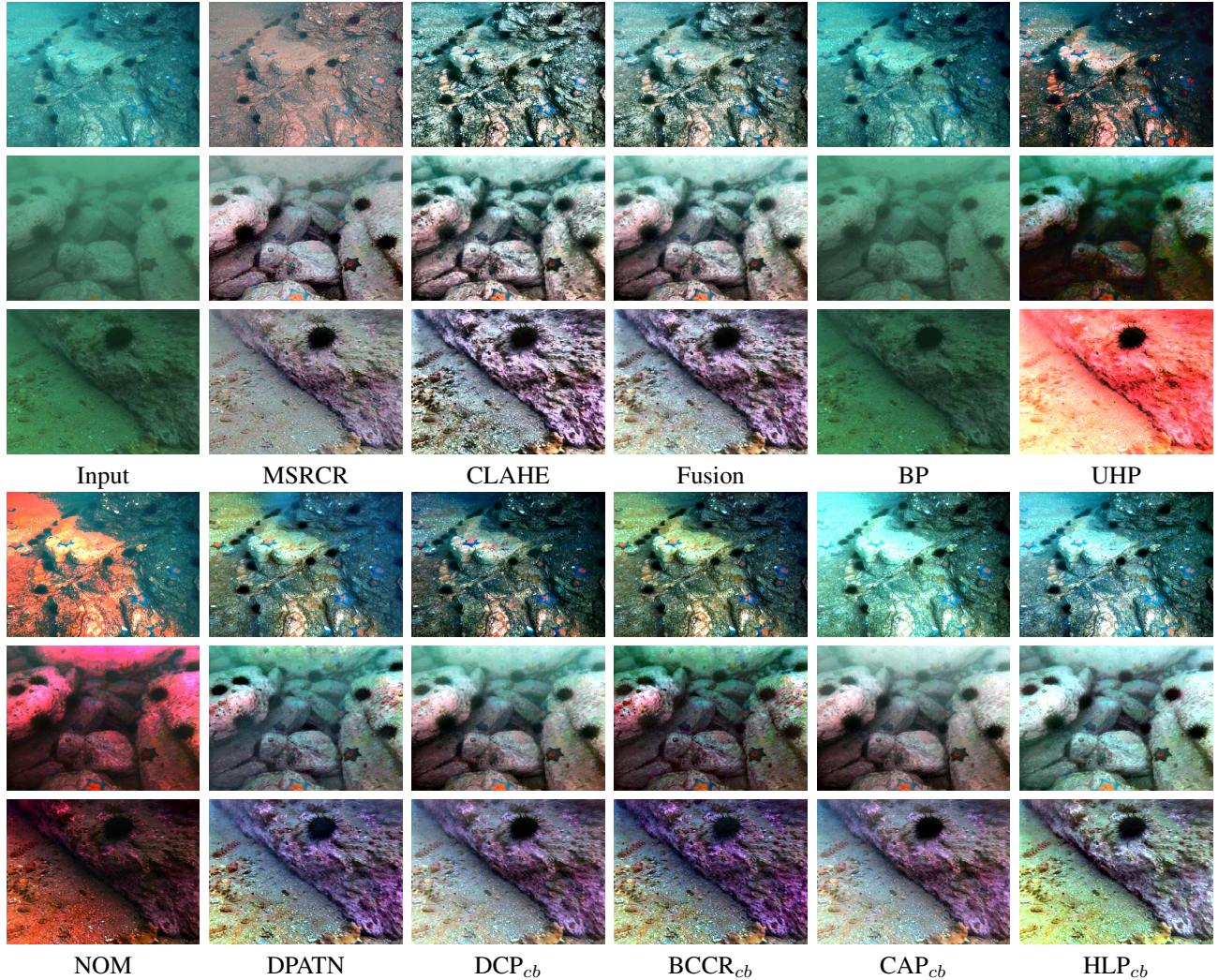


Fig. 4: The comparison on the dataset UCCS. The three input images from top to bottom are from subsets “Blue”, “Green-blue”, and “Green” of UCCS, respectively.

DCP_{cb} and $BCCR_{cb}$ can correct blue tone well giving more natural results, while HLP_{cb} performs the best when dealing with low illumination and greenish tone.

Quantitative comparisons: We also quantitatively evaluate the color correction ability of these methods by two metrics Avg_a and Avg_b , i.e., the average values of the channel a and b in the CIElab space, respectively. The metric Avg_a characterizes the component of green to red, with the green in the negative direction while the red in the positive one. Similarly, Avg_b represents the component from the blue to yellow. Larger absolute values of average chromaticity indicate more severity of color casts. Table IV gives the quantitative scores of the inputs images and outputs of the eleven UIE algorithms averaged in the three sub-datasets of UCCS. The most serious green bias appears in the “Green” set, and the green bias of the “Blue” set is not much different from that of the “Green-blue”. In comparison, the blue bias of the “Blue” set is much more serious, and the three subsets are gradually positive biased in the blue-yellow component.

MSRCR performs the best on the “Blue” set, and exhibits

excellent ability of correcting green bias in all three subsets, but tends to push Avg_a and Avg_b to positive values so that the results appear to be visually reddish. DCP_{cb} and $BCCR_{cb}$ work better on the blue-green than the other tones. On the “Green” set, DPATN, DCP_{cb} , and the three model-free methods have good ability to correct green bias. Unfortunately, BP can handle none of the three subsets well, and sometimes produces resultant images showing extremely subtle difference with the input of low quality. NOM over-corrects color casts, resulting in poor visual effects.

Considering both visual effects and quantitative analysis on UCCS, we can see that DCP_{cb} , $BCCR_{cb}$ and DPATN show more satisfactory performance on blueish images, while the model-free methods are more suitable for processing images with more green components.

Discussion: Color correction and visibility improvement

Color cast is one of challenging issues in underwater image enhancement as its effects vary with complex environmental factors such as water depth and salinity. The attenuation of red light through water causes the color cast as well as image

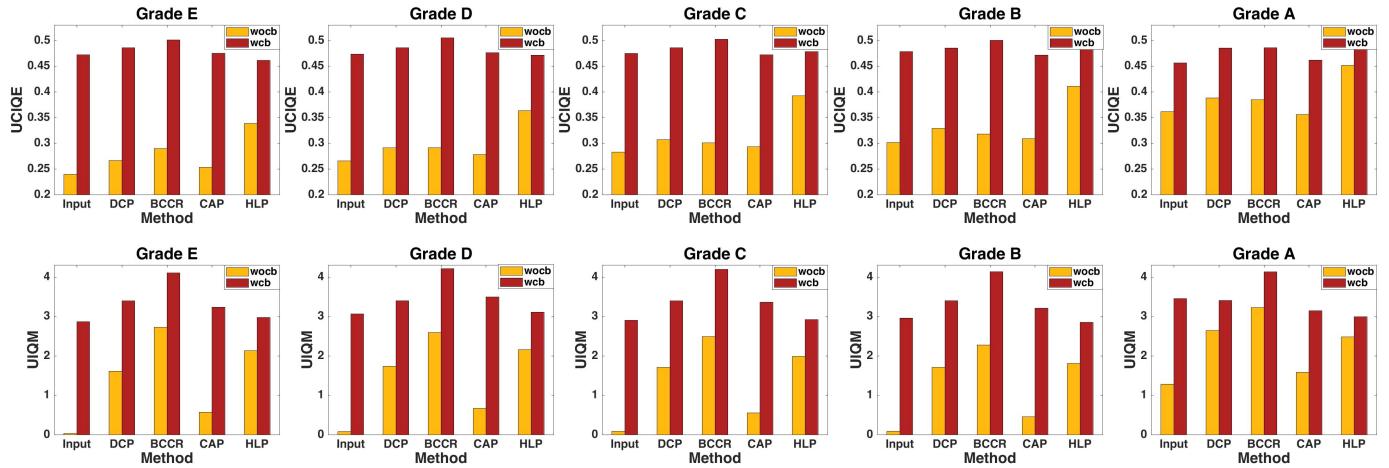


Fig. 5: Comparison of “underwater dehazing without color balance (wocb)” strategy and the “dehazing with color balance (wcb)” strategy on the five sub-datasets of UIQS.

TABLE IV: Average Avg_a / Avg_b scores on UCCS. Avg_a means the green-red component, with green in the negative direction. Avg_b represents the blue-yellow component, with blue in the negative direction. The best two are shown in bold.

Method	Blue	Green-blue	Green
Input	-25.84 / -6.56	-24.26 / 4.24	-30.97 / 12.10
MSRCR	1.17 / 0.47	2.42 / 1.05	2.58 / 0.21
CLAHE	-10.95 / -2.71	-6.73 / 1.67	-1.68 / 1.46
Fusion	-10.27 / -2.67	-6.07 / 1.66	-1.21 / 1.42
BP	-24.23 / -5.90	-23.15 / 4.85	-29.77 / 12.25
UHP	-11.70 / -1.08	-7.87 / 6.59	-9.84 / 6.42
NOM	0.88 / 6.90	35.25 / 9.05	36.01 / 22.25
DPATN	-10.15 / -3.14	-4.16 / 2.03	-1.15 / 1.49
DCP _{cb}	-12.21 / -2.27	1.77 / 0.84	0.76 / 1.58
BCCR _{cb}	-8.21 / 1.00	1.59 / 4.76	3.76 / 2.43
CAP _{cb}	-15.83 / -2.15	-4.69 / 5.46	1.70 / 2.79
HLP _{cb}	-16.14 / -6.16	-7.68 / 0.94	-8.20 / 5.17

details loss. This difficulty stumbles many UIE algorithms based on the transmission-related priors that work well in the context of image dehazing. For instance, the prior-based BP algorithm performs poor on greenish images. On the other hand, the four prior-based algorithms cascaded with a color balance module can achieve more appealing results.

We further compare the performance of the strategies with and without color balance (CB) [3] as postprocessing on the UIQS data set in order to peer into the impact from this simple technique tackling color cast. Figure 5 demonstrates that every prior-based dehazing-like UIE method somewhat improves the image quality. Among them, HLP increases the values of UCIQE the most, and BCCR is favored by UIQM the most. However, as shown in the first dark red bar of each chart of Figure 5, even a single CB module improves the input of low quality much more than anyone of the four prior-based algorithms without CB, i.e., DCP, BCCR, CAP and HLP, in terms of either UIQM or UCIQE.

Figure 5 also evidently illustrates that the strategy combining both prior-based visibility enhancement and color correction works superior to either individual module. Different priors may exhibit different affinities with the color balance module. DCP and BCCR gain more evident increases when combining CB, while CAP_{cb} and HLP_{cb} produce a tiny gap over the single CB module especially on the C, D and E subsets showing severe degradation. Therefore, in the future, it is worthy developing a more elaborate scheme to collaborate these two modules instead of simple cascade. In a recent preliminary study, we jointly learn both prior-based transmission and color correction in a residual learning framework, showing promising results [17].

C. Higher-level Task-driven Comparison on UHTS



Fig. 6: Example high-quality images from the training set.

We apply a common marine object classification module to the enhanced images given by the eleven UIE algorithm on UHTS, and evaluate the detection accuracy in terms of the mean Average Precision (mAP) and detection number (Num). Several dehazing studies [55], [56], [57] introduced similar task-driven evaluation on the performance of dehazing algorithms. Nevertheless, to our best knowledge, the task-specific evaluation on UIE algorithms still remains untouched, mainly attributing to no label of marine objects of interest available in any existing underwater benchmarks.

We labeled 1,800 clear captured underwater images (some examples are shown in Figure 6) and re-trained the YOLO-

V3 [47] network to detect and classify three types of marine objects, i.e., sea urchins, sea cucumbers, and scallops. Figure 8 compares the underwater object detection results on an UHTS degraded image and after applying different UIE algorithms. Figure 7 lists the values of mAP and Num. Most UIE algorithms can improve the detection numbers, but hardly bring significant benefits to mAP. The prior-based algorithms, BCCR_{cb} and HLP_{cb}, perform the best by significantly improving Num while stably increasing the values of mAP on all the subsets of UHTS. Model-free CLAHE and Fusion play a positive role for Num, and CLAHE keeps or slightly increases mAP while Fusion has a negative effect on mAP. Unfortunately, NOM performs little or even negative on both mAP and Num, especially on the subsets of B and C.

Furthermore, we employed WaterGan [13] to *synthesize* sixty underwater images containing 21 classes of outdoor objects (e.g., bicycle, person, and bus) by using 3,000 randomly selected real world underwater images from UIQS and 3,000 outdoor RGB-D images [56]. Subsequently, we applied the original trained YOLO-V3 model in [47] to detect the outdoor objects from the resultant images of the eleven UIE algorithms. Table V and Figure 10 provide the example detection results on this synthetic data set. The mAP results on the synthetic data in Table V are consistent with those on the real world set UHTS. The prior-based UIE algorithms followed by a simple color balance module, CAP_{cb} and DCP_{cb}, and the aggregated prior-data network, DPATN, obtain higher mAP values, while model-free based MSRCR, UHP and NOM increase little or even decrease mAP.

Discussion: The role of underwater image enhancement on higher-level object detection

It is natural for human perception that one can achieve more accurate detection on sharper images. Interestingly, for current automatic object detection algorithms using deep networks, our experimental results show that classification accuracy does not always have a positive correlation with visual quality. For instance, on the subset B, both CLAHE and Fusion can greatly improve the UCIQE and UIQM scores, but their mAP results sometimes are even lower than those of degraded inputs. On the subsets of A, D and E, UHP significantly improves mAP, but its quality evaluation is the worst among all methods. Badly blurred images, e.g., those in the subset E of UHTS, definitely downgrade the performance of object detection but high contrast or saturation, favoring human inspection, may yield undesired features for detection. Additionally, recent deep detectors highly depend on the consistency between training and testing sets. This data dependency may bring unexpected correlation between image quality and automatic object detection.

Actually, studies on haze removal [57], [56] also found a similar phenomenon that dehazing methods cannot always improve the performance of image classification performance. Pei *et al.* [57] reported that the objective of low-level enhancement typically differs from that of classification so that enhancement algorithms can hardly recover features favoring higher-level tasks. Therefore, higher-level classification may prefer to train end-to-end deep models directly from labeled degraded examples other than two separate steps. There are

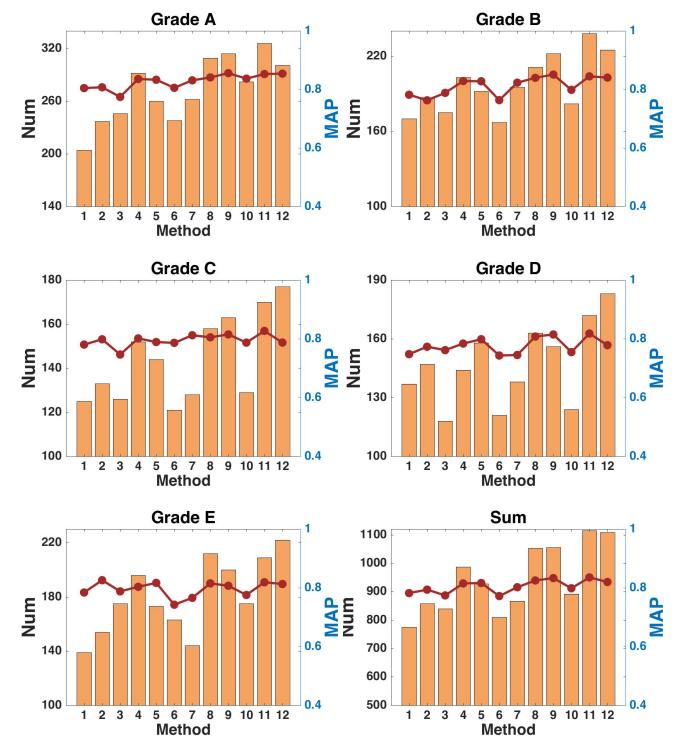


Fig. 7: Object detection number and mAP on UHTS. The histogram represents the detection number and the polyline represents mAP. Number 1 to 12 are respectively Input, MSRCR, CLAHE, Fusion, BP, UHP, NOM, DPATN, DCP_{cb}, BCCR_{cb}, CAP_{cb} HLP_{cb}.

two possible approaches to obtaining training examples that are critical to this end-to-end model: one is to directly label real world degraded images; the other is to use labeled images of natural scene together with unlabeled real world underwater images to synthesize degraded training examples with abundant labels, which requires designated GAN. For this respect, real-world images with different levels of image quality in our RUIE are not only helpful for evaluating low-level UIE but also boosting high-level tasks.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we setup an undersea imaging system and construct the RUIE benchmark. The benchmark consists of three subsets UIQS, UCCS and UHTS, targeting at the three challenging aspects or tasks for enhancement, i.e., visibility degradation, color cast, and higher-level detection/classification, respectively. Moreover, we evaluate eleven representative UIE algorithms on these three subsets in terms of various metrics on respective tasks. Table VII comprehensively shows the performance of the eleven methods compared on the three sub-datasets, with ticks representing the top four best. In terms of improving image quality, model-free CLAHE and prior-based BCCR_{cb} show relatively ideal performance on all images. But overall, model-free methods are more advantageous, especially in the case of poor-quality images. In terms of correcting color deviation, most methods cannot effectively solve all

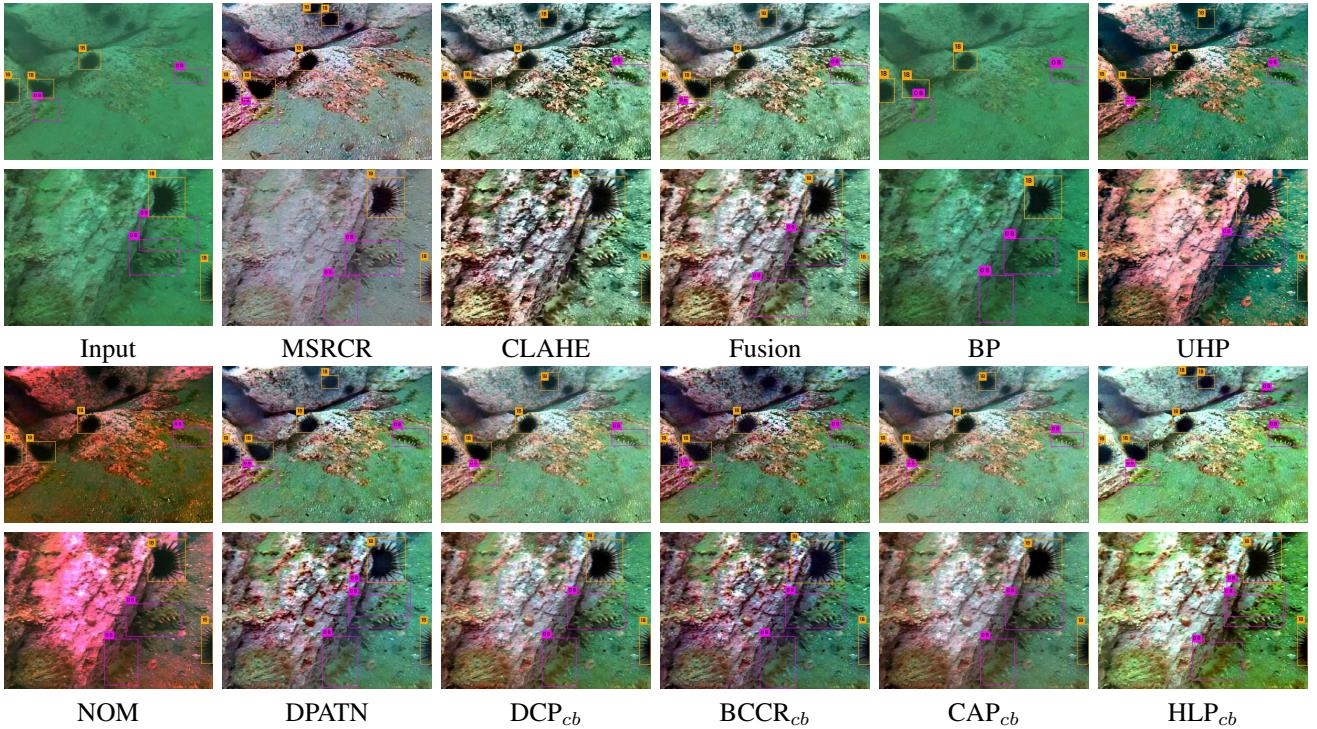


Fig. 8: Comparisons on UHTS.

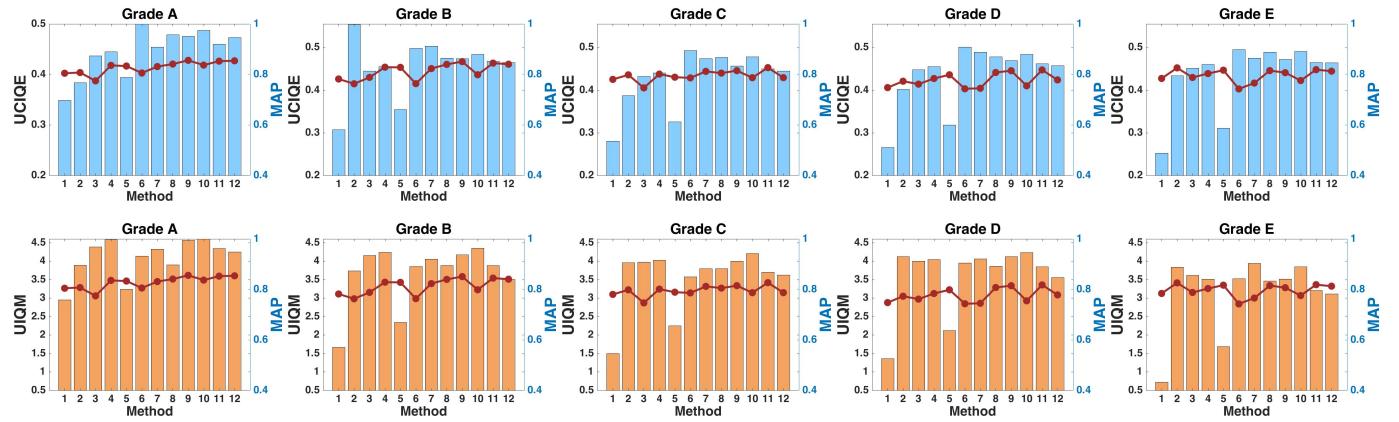


Fig. 9: The mAP and UCIQE/ UIQM scores on UHTS. The histogram represents UCIQW/UIQM scores and the polyline represents mAP, and methods 1 to 12 are respectively Input, MSRCR, CLAHE, Fusion, BP, UHP, NOM, DPATN, DCP_{cb}, BCCR_{cb}, CAP_{cb} HLP_{cb}.

TABLE V: MAP and total detection number on the synthetic dataset.

Category	Input	GT	MSRCR	CLAHE	Fusion	BP	UHP	NOM	DPATN	DCP _{cb}	BCCR _{cb}	CAP _{cb}	HLP _{cb}
mAP	0.703	0.725	0.707	0.720	0.717	0.714	0.702	0.695	0.723	0.726	0.721	0.725	0.721
Num	650	1038	893	882	888	767	733	668	804	823	837	819	835

TABLE VI: Average running time on 300×400 images.

Method	MSRCR	CLAHE	Fusion	BP	UHP	NOM	DPATN	DCP _{cb}	BCCR _{cb}	CAP _{cb}	HLP _{cb}
Time(s)	0.0119	0.0126	0.0232	4.5847	0.7985	0.6012	0.8758	0.8565	0.6568	0.5112	0.4256

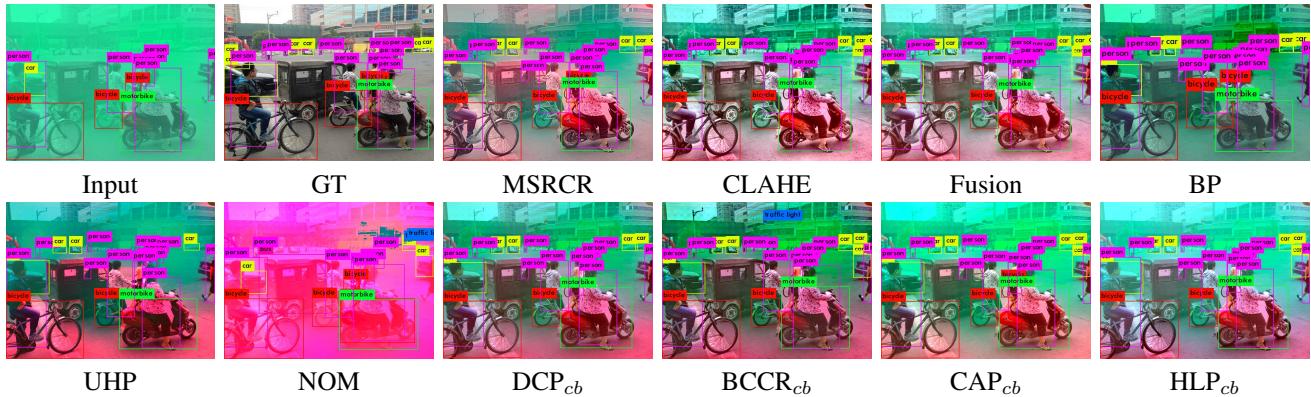


Fig. 10: Visualization of object detection results (on synthetic dataset) after applying different dehazing algorithms.

TABLE VII: The comprehensive comparison of eleven methods on three datasets, with ticks representing the best four. The three types of methods that are separated from top to bottom are “model-free”, “prior-based” and “data-driven”.

Dataset Methods	UIQS					UCCS		UHTS					
	E	D	C	B	A	blue	gb	green	E	D	C	B	A
MSRCR	✓	✓	✓			✓							
CLAHE	✓	✓	✓	✓	✓			✓					
Fusion		✓		✓	✓			✓					
BP									✓	✓			
UHP									✓	✓			
NOM											✓		
DCP _{cb}				✓	✓			✓			✓	✓	✓
BCCR _{cb}	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓
CAP _{cb}								✓	✓	✓	✓	✓	✓
HLP _{cb}											✓	✓	
DPATN	✓	✓						✓	✓	✓	✓	✓	✓

three types of color deviations. When dealing with task-driven problems, DPATN, CAP_{cb}, and DCP_{cb} are more dominant especially in difficult situations. Furthermore, the experimental results demonstrate that no one single UIE algorithm can work the best for all tasks upon all criteria. Besides, no strong correlation exists between the image quality assessment and detection accuracy (mAP). Therefore, tremendous efforts are highly demanded to more effective quality assessment, more elaborate paradigms simultaneously improving visibility and correcting color cast, and designated deep networks with more accurate detection/classification for underwater objects.

We advocate to develop data-driven non-reference assessments, joint learning frameworks for visibility enhancement and color correction, and end-to-end networks for underwater detection/classification in the future. All these possible directions would benefit from the proposed benchmark having a large number of real world images as well as showing a wide range of diversities. We expect this study to draw more interest from the computer vision community to work toward the challenging underwater tasks.

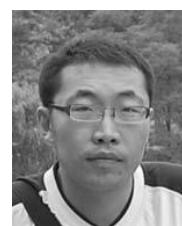
ACKNOWLEDGMENTS

Dalian Zhangzidao Group and Zhongji Oceanic Engineering Group provide great support for setting up the undersea image capturing system.

REFERENCES

- [1] D. J. Jobson, Z. Rahman, and G. A. Woodell, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” *IEEE TIP*, vol. 6, no. 7, pp. 965–976, 2002.
- [2] S. M. Pizer, R. E. Johnston, J. P. Erickson, B. C. Yankaskas, and K. E. Muller, “Contrast-limited adaptive histogram equalization: speed and effectiveness,” in *Visualization in Biomedical Computing*, 1990.
- [3] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, “Enhancing underwater images and videos by fusion,” in *CVPR*, 2012.
- [4] K. B. Gibson, “Preliminary results in using a joint contrast enhancement and turbulence mitigation method for underwater optical imaging,” in *OCEANS’15 MTS*, 2015.
- [5] N. Carlevaris-Bianco, A. Mohan, and R. M. Eustice, “Initial results in underwater single image dehazing,” in *Oceans*, 2010.
- [6] J. Y. Chiang and Y. Chen, “Underwater image enhancement by wavelength compensation and dehazing,” *IEEE TIP*, vol. 21, no. 4, pp. 1756–1769, 2012.
- [7] D. Berman, T. Treibitz, and S. Avidan, “Diving into haze-lines: Color restoration of underwater images,” in *BMVC*, 2017.
- [8] H. Wen, Y. Tian, T. Huang, and W. Gao, “Single underwater image enhancement with a new optical model,” in *ISCAS*, 2013.
- [9] P. D. Jr, E. D. Nascimento, F. Moraes, and S. Botelho, “Transmission estimation in underwater single images,” in *ICCV Workshops*, 2013.
- [10] A. Galdran and A. Alvarez-Gila, “Automatic red-channel underwater image restoration,” *JVCIR*, vol. 26, no. C, pp. 132–145, 2015.
- [11] Y. Li, H. Lu, J. Li, X. Li, Y. Li, and S. Serikawa, “Underwater image de-scattering and classification by deep neural network,” *Computers and Electrical Engineering*, vol. 54, pp. 68–77, 2016.
- [12] Y. T. Peng and P. C. Cosman, “Underwater image restoration based on image blurriness and light absorption.” *IEEE TIP*, vol. 26, no. 4, pp. 1579–1594, 2017.
- [13] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, “Watergan: Unsupervised generative network to enable real-time color correction of monocular underwater images,” *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 387–394, 2018.
- [14] X. Chen, J. Yu, S. Kong, Z. Wu, X. Fang, and L. Wen, “Towards real-time advancement of underwater visual quality with gan,” *IEEE TIE*, vol. 66, no. 12, pp. 9350–9359, 2019.
- [15] D. Xueyan, W. Yafei, W. Jun, and Z. Xianping, “Underwater image dehaze using scene depth estimation with adaptive color correction,” *OCEANS 2017 - Aberdeen*, pp. 1–5, 2017.
- [16] K. Cao, Y. T. Peng, and P. C. Cosman, “Underwater image restoration using deep networks to estimate background light and scene depth,” in *SSIAI*, 2018.
- [17] M. Hou, R. Liu, X. Fan, and Z. Luo, “Joint residual learning for underwater image enhancement,” in *ICIP*, 2018.
- [18] R. Liu, X. Fan, M. Hou, Z. Jiang, Z. Luo, and L. Zhang, “Learning aggregated transmission propagation networks for haze removal and beyond,” *IEEE Transactions on Neural Networks and Learning Systems*, no. 99, pp. 1–14, 2018.
- [19] R. Liu, L. Ma, Y. Wang, and L. Zhang, “Learning converged propagations with deep prior ensemble for image enhancement,” *IEEE TIP*, vol. 28, no. 3, pp. 1528–1543, 2019.

- [20] J. Perez, P. J. Sanz, M. Bryson, and S. B. Williams, "A benchmarking study on single image dehazing techniques for underwater autonomous vehicles," *OCEANS 2017 - Aberdeen*, pp. 1–9, 2017.
- [21] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE SPL*, vol. 22, no. 12, pp. 2387–2390, 2015.
- [22] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE TIP*, vol. 21, no. 12, 2012.
- [23] K. Iqbal, M. O. Odetayo, A. E. James, R. A. Salam, and A. Z. Talib, "Enhancing the low quality images using unsupervised colour correction method," in *SMC*, 2010.
- [24] M. S. Hitam, E. A. Awalludin, N. J. H. W. Y. Wan, and Z. Bachok, "Mixture contrast limited adaptive histogram equalization for underwater image enhancement," in *ICCA*, 2013.
- [25] A. S. A. Ghani and N. A. M. Isa, "Underwater image quality enhancement through integrated color model with rayleigh distribution," *ASC*, vol. 27, pp. 219–230, 2015.
- [26] H. Y. Yang, P. Y. Chen, C. C. Huang, Y. Z. Zhuang, and Y. H. Shiao, "Low complexity underwater image enhancement based on dark channel prior," in *IBICA*, 2012.
- [27] H. Lu, Y. Li, L. Zhang, and S. Serikawa, "Contrast enhancement for images in turbid water," *Journal of the Optical Society of America A*, vol. 32, no. 5, p. 886, 2015.
- [28] C. Li, J. Guo, R.-M. Cong, Y. Pang, and B. Wang, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," *IEEE TIP*, vol. 25, no. 12, pp. 5664–5677, 2016.
- [29] R. Hummel, "Image enhancement by histogram transformation," *CGIP*, vol. 6, no. 2, pp. 184–195, 1977.
- [30] G. Buchsbaum, "A spatial processor model for object colour perception," *J Franklin I*, vol. 310, no. 1, pp. 1–26, 1980.
- [31] Y. C. Liu, W. H. Chan, and Y. Q. Chen, "Automatic white balance for digital still camera," *IEEE Transactions on Consumer Electronics*, vol. 41, no. 3, pp. 460–466, 2004.
- [32] van de Weijer J, G. T, and G. A, "Edge-based color constancy." *IEEE TIP*, vol. 16, no. 9, pp. 2207–2214, 2010.
- [33] D. H. Foster, "Color constancy," *Vision Research*, vol. 51, no. 7, pp. 674–700, 2011.
- [34] C. O. Ancuti, C. Ancuti, V. C. De, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE TIP*, vol. 27, no. 1, pp. 379–393, 2017.
- [35] T. Treibitz and Y. Y. Schechner, "Active polarization descattering," *IEEE TPAMI*, vol. 31, no. 3, pp. 385–399, 2009.
- [36] G. Singh, N. Jaggi, S. Vasamsetti, H. K. Sardana, S. Kumar, and N. Mittal, "Underwater image/video enhancement using wavelet based color correction method," *IEEE Underwater Technology*, pp. 1–5, 2015.
- [37] C. J. Prabhakar and P. U. P. Kumar, "An image based technique for enhancement of underwater images," *IJMI*, vol. 3.
- [38] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE Journal of Oceanic Engineering*, vol. 15, no. 2, pp. 101–111, 1990.
- [39] B. L. Mcglamery, "A computer model for underwater camera systems," *Ocean Optics VI*, vol. 208, no. 208, pp. 221–231, 1980.
- [40] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *CVPR*, 2009.
- [41] N. Wang, H. Zheng, and B. Zheng, "Underwater image restoration via maximum attenuation identification," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2017.
- [42] Y. Wang, H. Liu, and L. P. Chau, "Single underwater image restoration using adaptive attenuation-curve prior," *IEEE TCS*, vol. PP, no. 99, pp. 1–11, 2017.
- [43] Y. Zhou, Q. Wu, K. Yan, L. Feng, and W. Xiang, "Underwater image restoration using color-line model," *IEEE TCSVT*, vol. PP, no. 99, pp. 1–1, 2018.
- [44] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *CVPR*, 2016.
- [46] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE TIP*, vol. 24, no. 12, pp. 6062–6071, 2015.
- [47] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [48] G. Meng, W. Ying, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *ICCV*, 2014.
- [49] Q. Z, J. M, and L. S, "A fast single image haze removal algorithm using color attenuation prior," *IEEE TIP*, vol. 24, no. 11, pp. 3522–3533, 2015.
- [50] D. Berman and S. Avidan, "Non-local image dehazing," in *CVPR*, 2016.
- [51] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 3, pp. 541–551, 2016.
- [52] K. Le, Y. Peng, L. Yi, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *CVPR*, 2014.
- [53] S. Bianco, L. Celona, P. Napoletano, and R. Schettini, "On the use of deep learning for blind image quality assessment," *Signal Image and Video Processing*, no. 3, pp. 1–8, 2016.
- [54] D. Pan, P. Shi, M. Hou, Z. Ying, S. Fu, and Y. Zhang, "Blind predicting similar quality map for image quality assessment," in *CVPR*, 2018.
- [55] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "Aod-net: All-in-one dehazing network," in *ICCV*, 2017.
- [56] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE TIP*, vol. 28, no. 1, pp. 492–505, 2019.
- [57] Y. Pei, Y. Huang, Q. Zou, Y. Lu, and S. Wang, "Does haze removal help cnn-based image classification?" in *ECCV*, 2018.



Risheng Liu Risheng Liu was born in 1984. He received the B.E. degree in mathematics and Ph.D. degree in computational mathematics from Dalian University of Technology, Dalian, China, in 2007 and 2012, respectively. He was engaged in joint training research at the Robotics Institute at Carnegie Mellon University, from 2010 to 2012. He has been a lecturer and associate professor at School of Software Technology, Dalian University of Technology, since 2012. His current research interests include machine learning, deep learning, computer vision, multimedia and optimization.



Xin Fan Xin Fan was born in 1977. He received the B.E. and Ph.D. degrees in information and communication engineering from Xian Jiaotong University, Xian, China, in 1998 and 2004, respectively. He was with Oklahoma State University, Stillwater, from 2006 to 2007, as a post-doctoral research Fellow. He joined the School of Software, Dalian University of Technology, Dalian, China, in 2009. He is also the duty dean of the DUTRU International School of Information and Science Technology. His current research interests include computational geometry and their applications to low-level image processing and DTI-MR image analysis.



Ming Zhu Ming Zhu received the bachelor and master also doctor degree in communication and information system school of software both from Dalian University of Technology, China. His research interests include wireless sensor networks and wireless networks communications.



Minjun Hou Minjun Hou was born in 1995. She received the B.E. degree in Software Engineering from Dalian University of Technology, China, in 2017. Now she is studying as a postgraduate student in Dalian University of Technology. Her research interests include computer vision, image processing and machine learning.



Zhongxuan Luo Zhongxuan Luo received the B.S. degree in Computational Mathematics from Jilin University, China, in 1985, the M.S. degree in Computational Mathematics from Jilin University in 1988, and the PhD degree in Computational Mathematics from Dalian University of Technology, China, in 1991. He has been a full professor of the School of Mathematical Sciences at Dalian University of Technology since 1997. His research interests include computational geometry and computer vision.