

# An Underwater Image Enhancement Benchmark Dataset and Beyond

Chongyi Li, Chunle Guo, Wenqi Ren, *Member, IEEE*, Runmin Cong, *Member, IEEE*, Junhui Hou, *Member, IEEE*, Sam Kwong, *Fellow, IEEE*, and Dacheng Tao, *Fellow, IEEE*

**Abstract**—Underwater image enhancement has been attracting much attention due to its significance in marine engineering and aquatic robotics. Numerous underwater image enhancement algorithms have been proposed in the last few years. However, these algorithms are mainly evaluated using either synthetic datasets or few selected real-world images. It is thus unclear how these algorithms would perform on images acquired in the wild and how we could gauge the progress in the field. To bridge this gap, we present the first comprehensive perceptual study and analysis of underwater image enhancement using large-scale real-world images. In this paper, we construct an *Underwater Image Enhancement Benchmark (UIEB)* including 950 real-world underwater images, 890 of which have the corresponding reference images. We treat the rest 60 underwater images which cannot obtain satisfactory reference images as challenging data. Using this dataset, we conduct a comprehensive study of the state-of-the-art underwater image enhancement algorithms qualitatively and quantitatively. In addition, we propose an underwater image enhancement network (called Water-Net) trained on this benchmark as a baseline, which indicates the generalization of the proposed UIEB for training Convolutional Neural Networks (CNNs). The benchmark evaluations and the proposed Water-Net demonstrate the performance and limitations of state-of-the-art algorithms, which shed light on future research in underwater image enhancement. The dataset and code are available at [https://li-chongyi.github.io/proj\\_benchmark.html](https://li-chongyi.github.io/proj_benchmark.html).

**Index Terms**—underwater image enhancement, real-world underwater images, comprehensive evaluation, deep learning.

This work was supported in part by CCF-Tencent Open Fund, in part by Zhejiang Lab's International Talent Fund for Young Professionals, in part by Fundamental Research Funds for the Central Universities under Grant 2019RC039, in part by National Natural Science Foundation of China under Grant 61802403, Grant 61771334, and Grant 61871342, in part by Hong Kong RGC General Research Funds under Grant 9042038 (CityU 11205314) and Grant 9042322 (CityU 11200116), and in part by Hong Kong RGC Early Career Schemes under Grant 9048123. (*Corresponding author:* Wenqi Ren)

Chongyi Li is with the Department of Computer Science, City University of Hong Kong, Kowloon 999077, Hong Kong SAR, China, and also with State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, China (e-mail: lichongyi25@gmail.com).

Chunle Guo is with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China (e-mail: guochunle@tju.edu.cn).

Wenqi Ren is with State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, China (e-mail: rwq.renwenqi@gmail.com).

Runmin Cong is with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China, and also with the Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing Jiaotong University, Beijing 100044, China (e-mail: rmcong@bjtu.edu.cn).

Junhui Hou and Sam Kwong are with the Department of Computer Science, City University of Hong Kong, Kowloon 999077, Hong Kong SAR, China, and also with the City University of Hong Kong Shenzhen Research Institute, Shenzhen 51800, China (e-mail: jh.hou@cityu.edu.hk; cssamk@cityu.edu.hk).

Dacheng Tao is with the UBTECH Sydney Artificial Intelligence Centre and the School of Information Technologies, the Faculty of Engineering and Information Technologies, the University of Sydney, 6 Cleveland St, Darlington, NSW 2008, Australia (e-mail: dacheng.tao@sydney.edu.au).

## I. INTRODUCTION

DURING the past few years, underwater image enhancement has drawn considerable attention in both image processing and underwater vision [1], [2]. Due to the complicated underwater environment and lighting conditions, enhancing underwater image is a challenging problem. Usually, an underwater image is degraded by wavelength-dependent absorption and scattering including forward scattering and backward scattering [3]–[7]. In addition, the marine snow introduces noise and increases the effects of scattering. These adverse effects reduce visibility, decrease contrast, and even introduce color casts, which limit the practical applications of underwater images and videos in marine biology and archaeology [8], marine ecological [9], to name a few [10]. To solve this problem, earlier methods rely on multiple underwater images or polarization filters, while recent algorithms deal with this problem by using only information from a single image.

Despite the prolific work, both the comprehensive study and insightful analysis of underwater image enhancement algorithms remain largely unsatisfactory due to the lack of a publicly available real-world underwater image dataset. Additionally, it is practically impossible to simultaneously photograph a real underwater scene and the corresponding ground truth image for different water types. Lacking sufficient and effective training data, the performance of deep learning-based underwater image enhancement algorithms does not match the success of recent deep learning-based high-level and low-level vision problems [11]–[15]. To advance the development of underwater image enhancement, we construct a large-scale real-world *Underwater Image Enhancement Benchmark (UIEB)*. Several sampling images and the corresponding reference images from UIEB are presented in Fig. 1. As shown, the raw underwater images in the UIEB have diverse color ranges and degrees of contrast decrease. In contrast, the corresponding reference images are color casts-free (at least relatively genuine color) and have improved visibility and brightness. With the proposed UIEB, we carry out a comprehensive study for several state-of-the-art single underwater image enhancement algorithms both qualitatively and quantitatively, which enables insights into their performance and sheds light on future research. In addition, with the constructed UIEB, CNNs can be easily trained to improve the visual quality of an underwater image. To demonstrate this application, we propose an underwater image enhancement model (Water-Net) trained on the constructed UIEB.

The main contributions of this paper are summarized as



Fig. 1. Sampling images from UIEB. Top row: raw underwater images taken in diverse underwater scenes; Bottom row: the corresponding reference results.

follows.

- We construct a large-scale real-world underwater image enhancement benchmark (*i.e.*, **UIEB**) which contains 950 real underwater images. These underwater images are likely taken under natural light, artificial light, or a mixture of natural light and artificial light. Moreover, the corresponding reference images for 890 images are provided according to laborious, time-consuming, and well-designed pairwise comparisons. UIEB provides a platform to evaluate, at least to some extent, the performance of different underwater image enhancement algorithms. It also makes supervised underwater image enhancement models which are out of the constraints of specific underwater scenes possible.
- With the constructed UIEB, we conduct a comprehensive study of the state-of-the-art single underwater image enhancement algorithms ranging from qualitative to quantitative evaluations. Our evaluation and analysis provide comprehensive insights into the strengths and limitations of current underwater image enhancement algorithms, and suggest new research directions.
- We propose a CNN model (*i.e.*, **Water-Net**) trained by the UIEB for underwater image enhancement, which demonstrates the generalization of the constructed UIEB and the advantages of our Water-Net, and also motivates the development of deep learning-based underwater image enhancement.

## II. EXISTING METHODOLOGY, EVALUATION METRIC, AND DATASET: AN OVERVIEW

### A. Underwater Image Enhancement Method

Exploring underwater world has become an active issue in recent years [16]–[18]. Underwater image enhancement as an indispensable step to improve the visual quality of recorded images has drawn much attention. A variety of methods have been proposed and can be organized into four groups: supplementary information-based, non-physical model-based, physical model-based, and data-driven methods.

**Supplementary Information-based Methods.** In the earlier stage, supplementary information from multiple images [19] or specialized hardware devices (*e.g.*, polarization filtering [20]–[23], range-gated imaging [24], [25], and fluorescence imaging [26], [27]) were utilized to improve the visibility of

underwater images. Compared to supplementary information-based methods, single underwater image enhancement has been proven to be more suitable for challenging situations such as dynamic scenes, and thus, gains extensive attention.

**Non-physical Model-based Methods.** Non-physical model-based methods aim to modify image pixel values to improve visual quality. Iqbal *et al.* [28] stretched the dynamic pixel range in RGB color space and HSV color space to improve the contrast and saturation of an underwater image. Chani and Isa [29], [30] modified the work of [28] to reduce the over-/under-enhanced regions by shaping the stretching process following the Rayleigh distribution. Ancuti *et al.* [31] proposed an underwater image enhancement method by blending a contrast-enhanced image and a color-corrected image in a multi-scale fusion strategy. In [32], a two-step approach for underwater image enhancement was proposed, which includes a color correction algorithm and a contrast enhancement algorithm.

Another line of research tries to enhance underwater images based on the Retinex model. Fu *et al.* [33] proposed a retinex-based method for underwater image enhancement, which consists of color correction, layer decomposition, and enhancement. Zhang *et al.* [34] proposed an extended multi-scale retinex-based underwater image enhancement method. In this work, the underwater

**Physical Model-based Methods.** Physical model-based methods regard the enhancement of an underwater image as an inverse problem, where the latent parameters of an image formation model are estimated from a given image. These methods usually follow the same pipeline: 1) building a physical model of the degradation; 2) estimating the unknown model parameters; and 3) addressing this inverse problem.

One line of research is to modify the Dark Channel Prior (DCP) [35] for underwater image enhancement. In [36], DCP was combined with the wavelength-dependent compensation algorithm to restore underwater images. In [37], an Underwater Dark Channel Prior (UDCP) was proposed based on the fact that the information of the red channel in an underwater image is undependable. Based on the observation that the dark channel of the underwater image tends to be a zero map, Liu and Chau [38] formulated a cost function and minimized it so as to find the optimal transmission map, which is able to maximize the image contrast. Instead of the DCP, Li *et al.* [39] employed the random forest regression model to

estimate the transmission of the underwater scenes. Recently, Peng *et al.* [40] proposed a Generalized Dark Channel Prior (GDCP) for image restoration, which incorporates adaptive color correction into an image formation model.

Another line of research try to employ the optical properties of underwater imaging. Carlevaris-Bianca *et al.* [41] proposed a prior that exploits the difference in attenuation among three color channels in RGB color space to predict the transmission of an underwater scene. The idea behind this prior is that the red light usually attenuates faster than the green light and the blue light in an underwater scenario. Galdran *et al.* [42] proposed a Red Channel method, which recovers the lost contrast of an underwater image by restoring the colors associated with short wavelengths. According to the findings that the background color of underwater images has relations with the inherent optical properties of water medium, Zhao *et al.* [43] enhanced the degraded underwater images by deriving inherent optical properties of water from the background color. Li *et al.* [44], [45] proposed an underwater image enhancement method based on the minimum information loss principle and histogram distribution prior. Peng *et al.* [46] proposed a depth estimation method for underwater scenes based on image blurriness and light absorption, which is employed to enhance underwater images. Berman *et al.* [47] took multiple spectral profiles of different water types into account and reduced the problem of underwater image restoration to single image dehazing. Wang *et al.* [48] combined the adaptive attenuation-curve prior with the characteristics of underwater light propagation for underwater image restoration. More recently, Akkaynak and Treibitz [49] proposed an underwater image color correction method based on a revised underwater image formation model [6] which is physically accurate. The existing physical model-based methods, except for the recent work [49], follow the simplified image formation models that assume the attenuation coefficients are only properties of the water and are uniform across the scene per color channel. This assumption leads to the unstable and visually unpleasing results as demonstrated in [6], [49].

**Data-driven Methods.** Recent years have witnessed the significant advance of deep learning in low-level vision problems. These methods can be trained using synthetic pairs of degraded images and high-quality counterparts. However, underwater image formation models depend on specific scenes and lighting conditions, and even are related to temperature and turbidity. Thus, it is difficult to synthesize realistic underwater images for CNNs training. Further, the learned distribution by CNNs trained on synthetic underwater images does not always generalize to real-world cases. Therefore, the performance and the amount of deep learning-based underwater image enhancement methods do not match the success of recent deep learning-based low-level vision problems [50].

Recently, Li *et al.* [51] proposed a deep learning-based underwater image enhancement model, called WaterGAN. WaterGAN first simulates underwater images from the in-air image and depth pairings in an unsupervised pipeline. With the synthetic training data, the authors use a two-stage network for underwater image restoration, especially for color casts removal. Underwater image enhancement models (*i.e.*,

UWCNNs) trained by ten types of underwater images was proposed in [52], where underwater images are synthesized based on a revised underwater image formation model [6] and the corresponding underwater scene parameters. More recently, a weakly supervised underwater color transfer model [53] (*i.e.*, Water CycleGAN) was proposed based on Cycle-Consistent Adversarial Networks [54]. Benefiting from the adversarial network architecture and multi-term loss function, this network model relaxes the need for paired underwater images for training and allows the underwater images being taken in unknown locations. However, it tends to produce inauthentic results in some cases due to the nature of multiple possible outputs. Guo *et al.* [55] proposed a multiscale dense GAN (*i.e.*, Dense GAN) for underwater image enhancement. The authors combined the non-saturating GAN loss with the  $\ell_1$  loss and gradient loss to learn the distribution of ground truth images in the feature domain. However, this method still cannot avoid the limitations of multiple possible outputs from GANs. Therefore, the robustness and generalization of deep learning-based underwater enhancement methods still fall behind conventional state-of-the-art methods.

## B. Underwater Image Quality Evaluation

In the following, we will give a brief introduction of the image quality evaluation metrics which are widely used for underwater image enhancement methods.

**Full-reference Metrics.** For an underwater image with ground truth image, the full-reference image quality evaluation metrics (*e.g.*, MSE, PSNR, and SSIM [56]) were employed for evaluations. Such underwater images usually are a few color checker images or color image patches taken in the simulated or real underwater environment. For example, Zhao *et al.* [43] treated a plastic color disk as ground truth image and captured its underwater image in a water pool as the testing image.

**Non-reference Metrics.** Different from other low-level vision problems where the ground truth images can be easily obtained (*e.g.*, image super-resolution), it is challenging to achieve a large amount of paired underwater images and the corresponding ground truth images. For a real-world underwater image where the ground truth image was unavailable, non-reference image quality evaluation metrics, such as image entropy, visible edges [57], and dynamic range independent image quality assessment [58], were utilized. In addition, some authors employed specific applications, like feature point matching, edge detection, and image segmentation, to evaluate their results. Besides, several specific non-reference metrics were proposed for underwater image quality evaluation. Yang and Sowmya [59] proposed an underwater color image quality evaluation metric (*i.e.*, UCIQE). UCIQE first quantifies the non-uniform color casts, blurring, and low contrast, and then combines these three components in a linear manner. In [60], the authors proposed a non-reference underwater image quality measure, called UIQM, which comprises three attribute measures: colorfulness measure, sharpness measure, and contrast measure. Each presented attribute measure is inspired by the properties of human visual system.

### C. Underwater Image Datasets

There are several real-world underwater image datasets such as Fish4Knowledge dataset for underwater target detection and recognition<sup>1</sup>, underwater images in SUN dataset for scene recognition and object detection<sup>2</sup> [61], MARIS dataset for marine autonomous robotics<sup>3</sup>, Sea-thru dataset including 1100 underwater image with range maps<sup>4</sup> [49], and Haze-line dataset providing raw images, TIF files, camera calibration files, and distance maps<sup>5</sup> [62]. However, existing datasets usually have monotonous content and limited scenes, few degradation characteristics, and insufficient data. Moreover, these datasets did not provide the corresponding ground truth images or reference results since it is difficult or even impractical to simultaneously obtain a real underwater image and the corresponding ground truth image of the same scene due to the diverse water types and lighting conditions as well as expensive and logically complex imaging devices. In recent years, several underwater image synthesis methods have been proposed. Li *et al.* [51] proposed a GAN-based method<sup>6</sup> while Duarte *et al.* [63] simulated underwater image degradation using milk, chlorophyll, or green tea in a tank. Blasinski *et al.* [64] provided an open-source underwater image simulation tool and a three parameter underwater image formation model [65]. Li *et al.* [52] proposed a synthetic underwater image dataset including ten subsets for different types of water<sup>7</sup>. However, there still exists a gap between synthetic and real-world underwater images. Therefore, it is challenging to evaluate the state-of-the-art methods fairly and comprehensively, and is hard to develop effective deep learning-based models.

### III. PROPOSED BENCHMARK DATASET

After systematically reviewing previous work, we found the main issue existing in the community of underwater image enhancement is lacking a large-scale real-world underwater image dataset with reference images. In what follows, we introduce the constructed dataset in detail, including data collection and reference image generation.

#### A. Data Collection

There are three objectives for underwater image collection:

- 1) a diversity of underwater scenes, different characteristics of quality degradation, and a broad range of image content should be covered;
- 2) the amount of underwater images should be large; and
- 3) the corresponding high-quality reference images should be provided so that pairs of images enable fair image quality evaluation and end-to-end learning.

To achieve the first two objectives, we first collect a large number of underwater images, and then refine them.

<sup>1</sup><http://groups.inf.ed.ac.uk/f4k/>

<sup>2</sup><http://groups.csail.mit.edu/vision/SUN/>

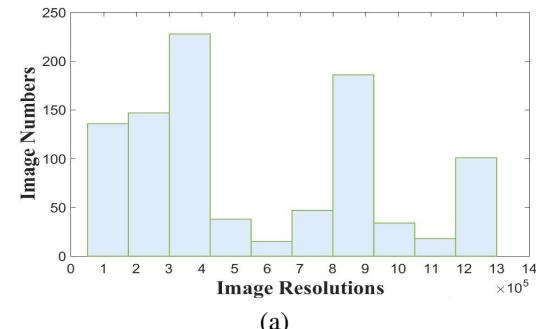
<sup>3</sup><http://rimlab.ce.unipi.it/Marisi.html>

<sup>4</sup>[http://csmms.haifa.ac.il/profiles/Treibitz/datasets/sea\\_thru/index.html](http://csmms.haifa.ac.il/profiles/Treibitz/datasets/sea_thru/index.html)

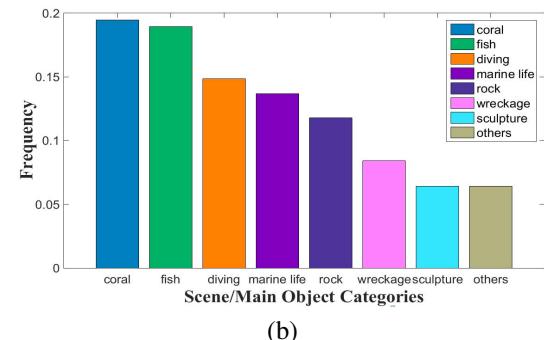
<sup>5</sup>[http://csmms.haifa.ac.il/profiles/Treibitz/datasets/ambient\\_forwardlooking/index.html](http://csmms.haifa.ac.il/profiles/Treibitz/datasets/ambient_forwardlooking/index.html)

<sup>6</sup><https://github.com/kskin/WaterGAN>

<sup>7</sup>[https://li-chongyi.github.io/proj\\_underwater\\_image\\_synthesis.html](https://li-chongyi.github.io/proj_underwater_image_synthesis.html)



(a)



(b)

Fig. 2. Statistics of the constructed UIEB. (a) Image resolutions. (b) Scene/main object categories.

These underwater images are collected from Google, YouTube, related papers [23], [31]–[33], [42], and our self-captured videos. We mainly retain the underwater images which meet the first objective. After data refinement, most of the collected images are weeded out, and about 950 candidate images are remaining. We provide a statistic of image resolutions and the scene/main object categories of the UIEB in Fig. 2 and present some examples of the images in Fig. 3.

As shown in Figs. 2 and 3, the UIEB contains a large range of image resolutions and spans diverse scene/main object categories including coral (*e.g.*, fringing reefs and barrier reefs), marine life (*e.g.*, turtles and sharks), *etc.* To achieve the third objective, we introduce a high-quality reference image generation method in the next section.

#### B. Reference Image Generation

With the candidate underwater images, the potential reference images are generated by 12 image enhancement methods, including 9 underwater image enhancement methods (*i.e.*, fusion-based [31], two-step-based [32], retinex-based [33], UDCP [37], regression-based [39], GDCP [40], Red Channel [42], histogram prior [45], and blurriness-based [46]), 2 image dehazing methods (*i.e.*, DCP [35] and MSCNN [66]), and 1 commercial application for enhancing underwater images (*i.e.*, dive+<sup>8</sup>). We exclude the recent deep learning-based methods due to their limited generalization capability to the diverse real-world underwater images and the fixed size of network output [53], [55]. The source codes of all the employed methods except the fusion-based [31] are provided by their authors.

<sup>8</sup><https://itunes.apple.com/us/app/dive-video-color-correction/id1251506403?mt=8>



Fig. 3. Examples of the images in UIEB. These images have obvious characteristics of underwater image quality degradation (*e.g.*, color casts, decreased contrast, and blurring details) and are taken in a diversity of underwater scenes.

We reimplement the fusion-based [31] since the source code is unavailable. For the dive+, we tune its parameter settings to generate satisfactory results. At last, we totally generate  $12 \times 950$  enhanced results.



Fig. 4. An example of pairwise comparisons. Treating the raw underwater image as a reference, a volunteer needs to independently decide which one is better between the results of method A and method B.

With raw underwater images and the enhanced results, we invite 50 volunteers (25 volunteers with image processing experience; 25 volunteers without related experience) to perform pairwise comparisons among the 12 enhanced results of each raw underwater image under the same monitor.

Specifically, each volunteer is shown a raw underwater image and a set of enhanced result pairs. The enhanced image pairs are drawn from all the competitive methods randomly, and the result winning the pairwise comparisons will be compared again in the next round, until the best one is selected. There is no time constraint for volunteers and zoom-in operation is allowed. An example of pairwise comparisons is shown in Fig. 4. For each pair of enhanced results, taking the raw underwater image as a reference, a volunteer first needs to independently decide which one is better than the other. For each volunteer, the best result will be selected after 11 pairwise comparisons.

Additionally, the volunteer needs to inspect the best result again and then label the best result as being satisfactory or dissatisfaction. The reference image for a raw underwater image is first selected by majority voting after pairwise comparisons.

TABLE I  
PERCENTAGE OF THE REFERENCE IMAGES FROM THE RESULTS OF DIFFERENT METHODS.

Method	Percentage (%)
fusion-based [31]	<b>24.72</b>
two-step-based [32]	7.30
retinex-based [33]	0.22
DCP [35]	2.58
UDCP [37]	0.00
regression-based [39]	1.80
GDCP [40]	0.34
Red Channel [42]	0.90
histogram prior [45]	13.37
blurriness-based [46]	3.48
MSCNN [66]	0.90
dive+	<b>43.93</b>

After that, if the selected reference image has greater than half the number of votes labeled dissatisfaction, its corresponding raw underwater image is treated as a challenging image and the reference image is discarded. We totally achieve 890 available reference images which have higher quality than any individual methods and a challenging set including 60 underwater images. To visualize the process of reference image generation, we present some cases that the results of some methods are shown and indicate which one is the final reference image in Fig. 5. Furthermore, the percentage of the reference images from the results of different methods is presented in Table I. In this paper, we highlight the top one performance in red, whereas the second top one is in blue.

In summary, the results with improved contrast and genuine color are most favored by observers while the over-/under-enhancement, artifacts, and color casts lead to visually unpleasing results. Finally, the constructed UIEB includes two subsets: 890 raw underwater images with the corresponding high-quality reference images; 60 challenging underwater images. To the best of our knowledge, it is the first real-world underwater image dataset with reference images so far. The UIEB has various potential applications, such as performance evaluation and CNNs training. Next, we will introduce these two applications.

#### IV. EVALUATION AND DISCUSSION

A comprehensive and fair evaluation of underwater image enhancement methods has long been missing from the literatures. Using the constructed UIEB, we evaluate the state-of-the-art underwater image enhancement methods (*i.e.*, fusion-based [31], two-step-based [32], retinex-based [33], UDCP [37], regression-based [39], GDCP [40], Red Channel [42], histogram prior [45], blurriness-based [46]) both qualitatively and quantitatively.

##### A. Qualitative Evaluation

We first select several underwater images from the UIEB, and then divide these images into five categories: greenish and bluish images, downward looking images, forward looking images, low backscatter scenes, and high backscatter scenes. The results of different methods and the corresponding reference

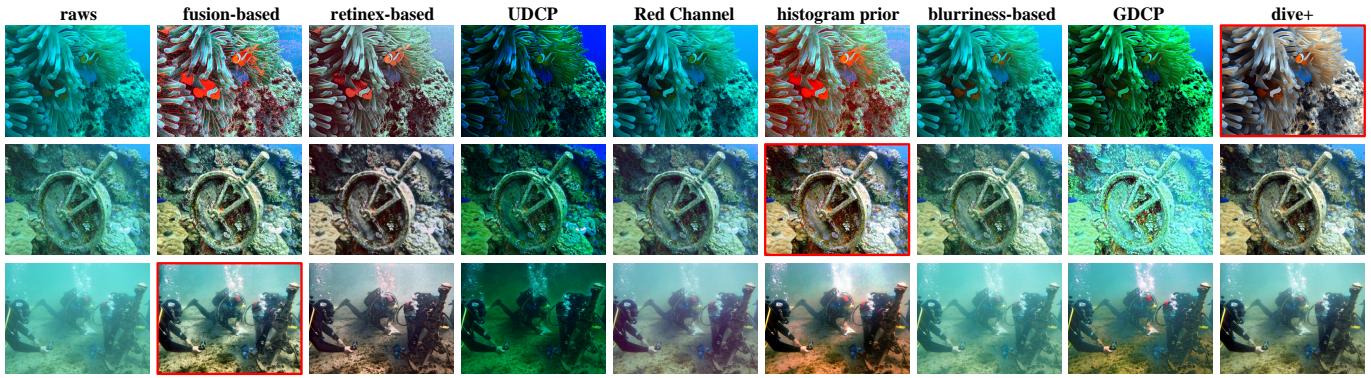


Fig. 5. Results generated by different methods. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], UDCP [37], Red Channel [42], histogram prior [45], blurriness-based [46], GDCP [40] and dive+. Red boxes indicate the final reference images.

images are shown in Figs. 6-10. Best viewed with zoom-in on a digital display. Note that these underwater images cannot cover the entire UIEB.

In open water, the red light first disappears because of its longest wavelength, followed by the green light and then the blue light [7]. Such selective attenuation in open water results in bluish or greenish underwater images, such as the raw underwater images in Fig. 6. Color deviation seriously affects the visual quality of underwater images and is difficult to be removed. As shown, the fusion-based [31], histogram prior [45], and regression-based [39] introduce reddish color deviation due to the inaccurate color correction algorithms used in [31], [39] and the histogram distribution prior [45]. The retinex-based [33] removes the color deviation well, while the UDCP [37] and GDCP [40] aggravate the effect of color casts. The two-step-based [32] can effectively increase the contrast of underwater images. The Red Channel [42] and blurriness-based [46] have less positive effect on these two images on account of the limitations of the priors used in these two methods.

For the downward looking images and forward looking images shown in Figs. 7 and 8, the fusion-based [31], retinex-based [33], and histogram prior [45] significantly remove the effect of haze on the underwater images while the two-step-based [32], Red Channel [42], and blurriness-based [46] remain some haze in the results. We note that UDCP [37], regression-based [39], and GDCP [40] tend to bring in color deviation in the enhanced results. In terms of the physical-model based methods (*e.g.*, UDCP [37], Red Channel [42], regression-based [39], blurriness-based [46], and GDCP [40]), it is hard to estimate the veiling light from the downward looking images accurately. In addition, the physical-model based methods may incorrectly estimate the veiling light from the RGB values of water in textures regions of forward looking images

Backscatter reduces the contrast and produces the foggy veiling in an underwater image. The effect of backscatter is related to the distance between the camera and the scene. For the low backscatter scenes (short distance between the camera and the scene) in Fig. 9, the effect of backscatter is relatively easy to be removed. In contrast, the high backscatter (long distance) significantly degrades the visual quality of

underwater images. As shown in Fig. 10, all the physical model-based methods (*e.g.*, UDCP [37], Red Channel [42], regression-based [39], blurriness-based [46], and GDCP [40]) cannot remove the high backscatter due to the inaccurate physical models and assumptions used in these methods.

In summary, the fusion-based [31] has relatively decent performance on a variety of underwater images. The method of UDCP [37] tends to produce artifacts on enhanced results in some cases. Other competitors are effective to some extent. In fact, it is almost impossible for a color correction algorithm or a kind of prior effective for all types of underwater images. Moreover, the effect of high backscatter is challenging for underwater image enhancement.

### B. Quantitative Evaluation

To quantitatively evaluate the performance of different methods, we perform the full-reference evaluation, non-reference evaluation, and runtime evaluation.

1) *Full-reference Evaluation*: We first conduct a full-reference evaluation using three commonly-used metrics (*i.e.*, MSE, PSNR, and SSIM). The results of full-reference image quality evaluation by using the reference images can provide realistic feedback of the performance of different methods to some extent, although the real ground truth images might be different from the reference images. A higher PSNR score and a lower MSE score denote the result is closer to the reference image in terms of image content, while a higher SSIM score means the result is more similar to the reference image in terms of image structure and texture. We present the average scores of different methods on the 890 images with reference images in the UIEB. As shown in Table II, the dive+ stands out as the best performer across all metrics. In addition, the fusion-based [31] ranks the second best in terms of the full-reference metrics. It is reasonable for such results, since most reference images are selected from the results generated by dive+ and fusion-based [31].

2) *Non-reference Evaluation*: We employ two non-reference metrics (*i.e.*, UCIQE [59] and UIQM [60]) which are usually used for underwater image quality evaluation [39], [40], [45], [46]. A higher UCIQE score indicates the result has better balance among the chroma, saturation, and

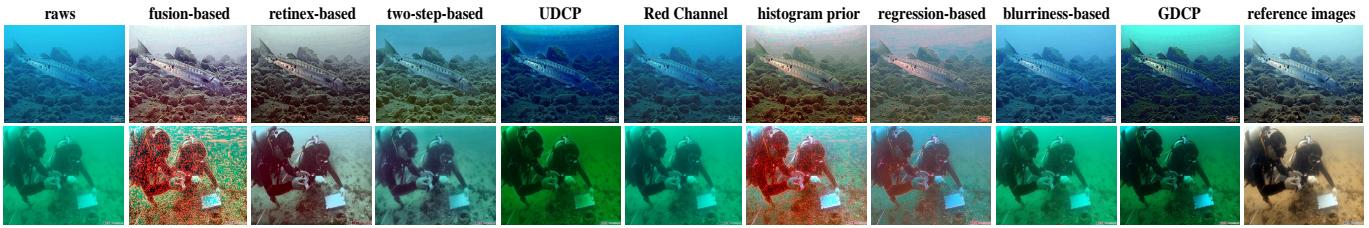


Fig. 6. Subjective comparisons on bluish and greenish underwater images. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], two-step-based [32], UDCP [37], Red Channel [42], histogram prior [45], regression-based [39], blurriness-based [46], GDCP [40], and reference images.

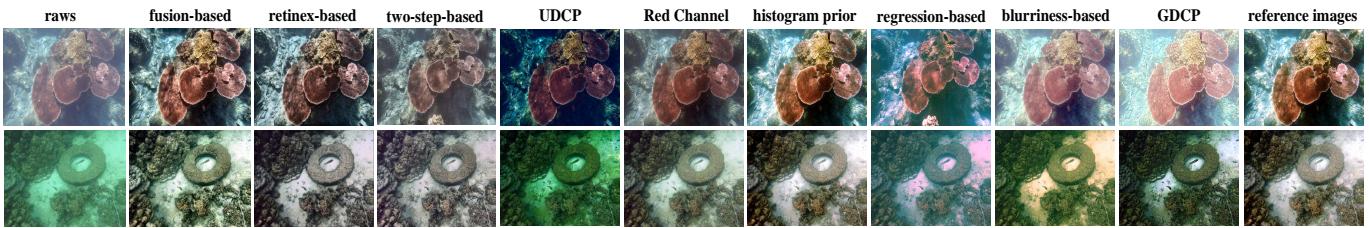


Fig. 7. Subjective comparisons on downward looking images. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], two-step-based [32], UDCP [37], Red Channel [42], histogram prior [45], regression-based [39], blurriness-based [46], GDCP [40], and reference images.

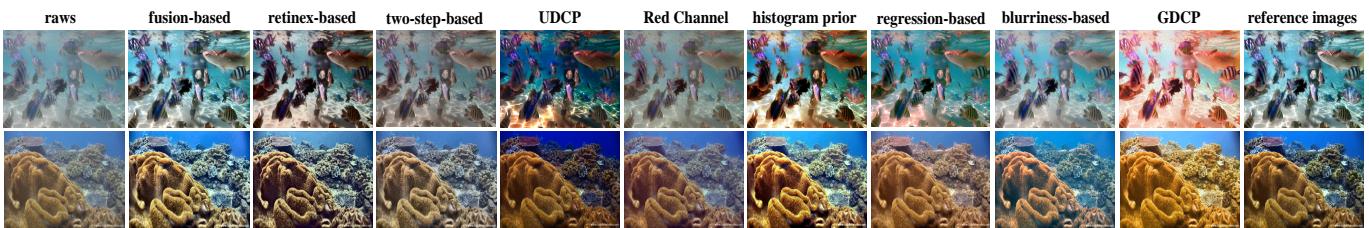


Fig. 8. Subjective comparisons on forward looking images. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], two-step-based [32], UDCP [37], Red Channel [42], histogram prior [45], regression-based [39], blurriness-based [46], GDCP [40], and reference images.

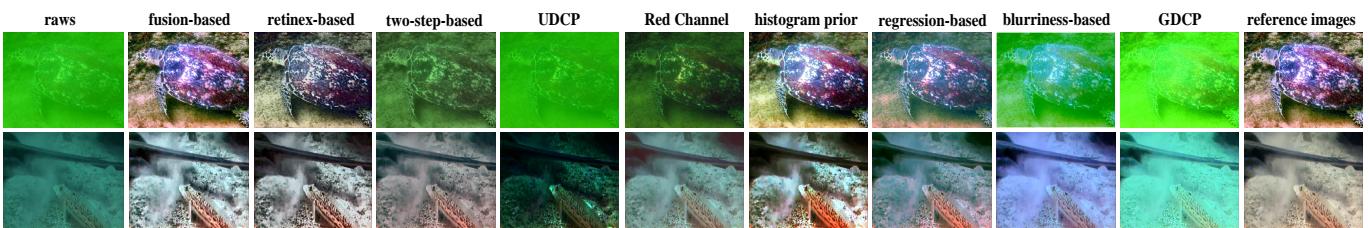


Fig. 9. Subjective comparisons on low backscatter scenes. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], two-step-based [32], UDCP [37], Red Channel [42], histogram prior [45], regression-based [39], blurriness-based [46], GDCP [40], and reference images.

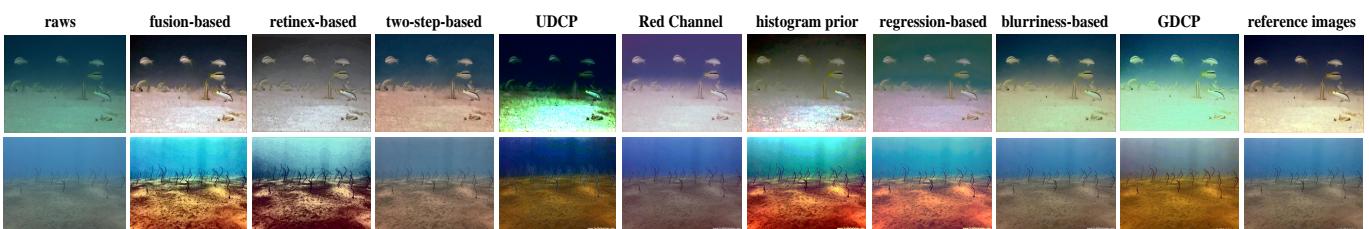


Fig. 10. Subjective comparisons on high backscatter scenes. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], two-step-based [32], UDCP [37], Red Channel [42], histogram prior [45], regression-based [39], blurriness-based [46], GDCP [40], and reference images.

contrast, while a higher UIQM score indicates the result is more consistent with human visual perception. The average scores are shown in Table III.

In Table III, the histogram prior [45] and UDCP [37] obtain the highest scores of UCIQE and UIQM, respectively. The dive+ and fusion-based [31] are no more the best performers. It

TABLE II  
FULL-REFERENCE IMAGE QUALITY EVALUATION.

Method	MSE ( $\times 10^3$ ) $\downarrow$	PSNR (dB) $\uparrow$	SSIM $\uparrow$
fusion-based [31]	<b>0.8679</b>	<b>18.7461</b>	<b>0.8162</b>
two-step-based [32]	1.1146	17.6596	0.7199
retinex-based [33]	1.3531	16.8757	0.6233
UDCP [37]	5.1300	11.0296	0.4999
regression-based [39]	1.1365	17.5751	0.6543
GDCP [40]	3.6345	12.5264	0.5503
Red Channel [42]	2.1073	14.8935	0.5973
histogram prior [45]	1.6282	16.0137	0.5888
blurriness-based [46]	1.5826	16.1371	0.6582
dive+	<b>0.5358</b>	<b>20.8408</b>	<b>0.8705</b>

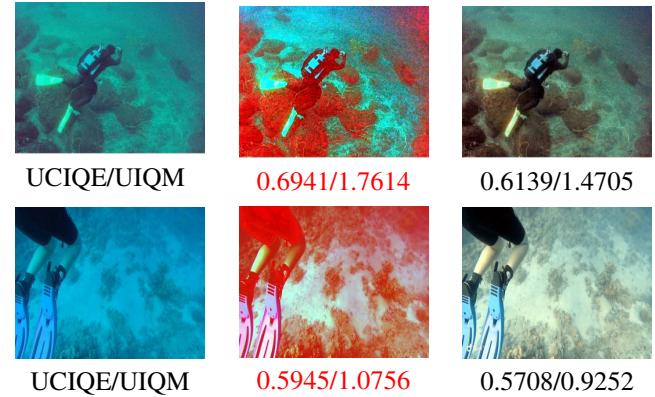
TABLE III  
NO-REFERENCE IMAGE QUALITY EVALUATION.

Method	UCIQE [59] $\uparrow$	UIQM [60] $\uparrow$
fusion-based [31]	<b>0.6414</b>	1.5310
two-step-based [32]	0.5776	1.4002
retinex-based [33]	0.6062	1.4338
UDCP [37]	0.5852	<b>1.6297</b>
regression-based [39]	0.5971	1.2996
GDCP [40]	0.5993	1.4301
Red Channel [42]	0.5421	1.2147
histogram prior [45]	<b>0.6778</b>	<b>1.5440</b>
blurriness-based [46]	0.6001	1.3757
dive+	0.6227	1.3410

is interesting that the good performers in terms of UCIQE and UIQM metrics are not consistent with the subjective pairwise comparisons, though both UCIQE and UIQM claim that they take the human visual perception into account. Such a result provides evidence that the current image quality evaluation metrics designed for underwater image are inconsistent with human visual perception in some cases. This is because humans have not evolved to see in aquatic habitats. When they are shown an underwater photo, they are most likely to pay attention to objects in the center of the scene, or whatever seems to be colorful or interesting (e.g., diver, fish, coral, etc). Thus, human visual perception may be a totally inaccurate way of color correcting underwater images. It might be fine for visually pleasing images, but not to learn attenuation and backscatter.

Furthermore, Figs. 6-10 show the results generated by histogram prior [45] and UDCP [37] still suffer from color casts and over-enhancement. Through further analyzing, we found these two non-reference metrics might be biased to some characteristics (not entire image) and did not take the color shift and artifacts into account. For example, the results with high contrast (e.g., the results of histogram prior [45]) are usually favored by the UICQE metric. To illustrate this phenomenon, we present an example in Fig. 11.

In Fig. 11, the results generated by histogram prior [45] have obvious reddish color shift and artifacts; however, they obtain better quantitative scores in terms of UCIQE and UIQM metrics than the results of dive+. Thus, we believe there is a gap between the quantitative scores of non-reference metrics and the subjectively visual quality. In other words, the current image quality evaluation metrics designed for underwater



(a) Raws      (b) histogram prior [45]      (c) dive+

Fig. 11. Visual comparisons in terms of UCIQE and UIQM metrics. Higher scores are in red. It is obvious that higher quantitative scores do not lead to better subjectively quality.

image have limitations in some cases.

3) *Runtime Evaluation*: We compare the average runtime for the images of different sizes. Experiments are conducted by using MATLAB R2014b on a PC with an Intel(R) i7-6700 CPU, 32GB RAM. The average runtime is shown in Table IV. The two-step-based [32] is the fastest across different image sizes, while the retinex-based [33] ranks the second fastest. The regression-based [39] is the slowest method due to the time-consuming random forest-based transmission prediction, especially for images with large sizes.

TABLE IV  
AVERAGE RUNTIME FOR DIFFERENT IMAGE SIZES (IN SECOND).

Method	500 $\times$ 500	640 $\times$ 480	1280 $\times$ 720
fusion-based [31]	<b>0.6044</b>	<b>0.6798</b>	<b>1.8431</b>
two-step-based [32]	<b>0.2978</b>	<b>0.4391</b>	<b>1.0361</b>
retinex-based [33]	0.6975	0.8829	2.1089
UDCP [37]	2.2688	3.3185	9.9019
regression-based [39]	138.6138	167.1711	415.4935
GDCP [40]	3.2676	3.8974	9.5934
Red Channel [42]	2.7523	3.2503	9.7447
histogram prior [45]	4.6284	5.8289	16.9229
blurriness-based [46]	37.0018	47.2538	146.0233

After reviewing and evaluating the state-of-the-art underwater image enhancement methods, we found that the fusion-based [31] is the relatively best performer in most cases, while other compared methods have obvious disadvantages. However, there is no method which always wins when facing a large-scale real-world underwater image dataset (*i.e.*, UIEB). All in all, due to neglecting the underwater imaging physical models, the non-physical model-based methods, such as two-step-based [32] and retinex-based [33], produce over-/under-enhanced results. Physical model-based methods, such as UDCP [37], employ an outdoor haze formation based model to predict the medium transmission which is not well-suited for the underwater scenario. Inaccurate physical models and assumptions result in color casts and remaining haze in the results such as regression-based [39], GDCP [40], Red Channel [42], histogram prior [45], and blurriness-based [46]. Some methods, such as retinex-based [33] and histogram

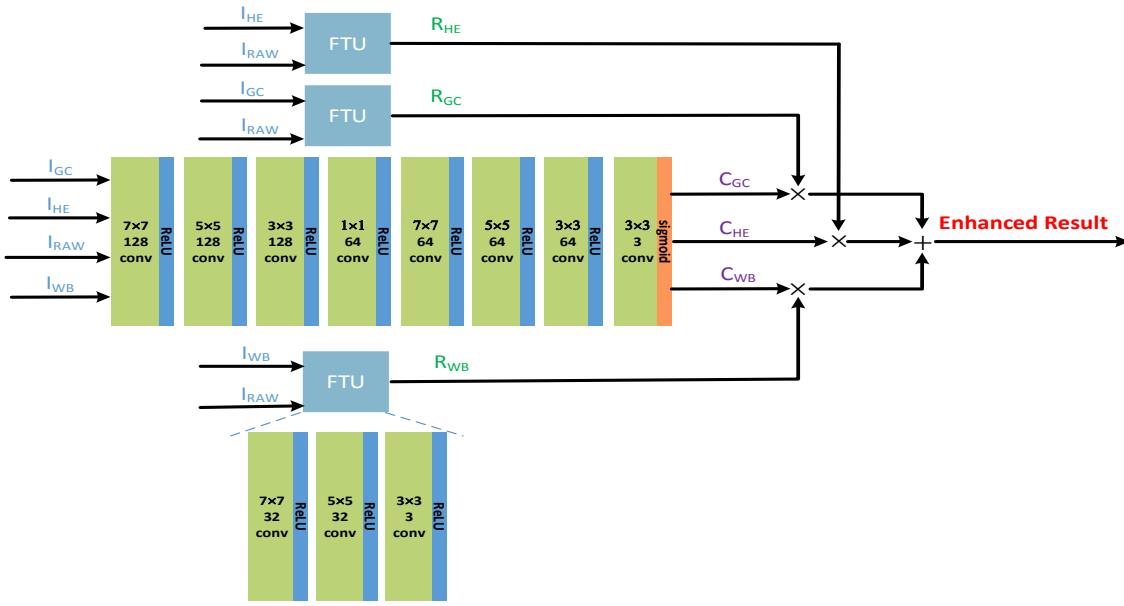


Fig. 12. An overview of the proposed Water-Net architecture. Water-Net is a gated fusion network, which fuses the inputs with the predicted confidence maps to achieve the enhanced result. The inputs are first transferred to the refined inputs by the Feature Transformation Units (FTUs) and then the confidence maps are predicted. At last, the enhanced result is achieved by fusing the refined inputs and the corresponding confidence maps.

prior [45], tend to introduce noise and artifacts, which leads to visually unpleasing results. The runtime of some methods seriously limits their practical applications.

In the future, a comprehensive method that can robustly deal with a variety of underwater image degradation is expected. The non-reference metrics which are more effective and consistent with human visual perception are desired in the community of underwater image enhancement. More discussions will be provided in Sec. VI.

## V. PROPOSED MODEL

Despite the remarkable progress of underwater image enhancement methods, the generalization of deep learning-based underwater image enhancement models still falls behind the conventional state-of-the-art methods due to the lack of effective training data and well-designed network architectures. With the UIEB, we propose a CNN model for underwater image enhancement, called Water-Net. The purpose of the proposed Water-Net as a baseline is to call for the development of deep learning-based underwater image enhancement, and demonstrate the generalization of the UIEB for training CNNs. Note that the proposed Water-Net is only a baseline model which can be further improved by well-designed network architectures, task-related loss functions, and the like.

In this section, we first present input generation from an underwater image and the architecture of the proposed Water-Net. Then we present the training and implementation details. At last, we perform experiments to demonstrate its advantages.

### A. Input Generation

As discussed in Sec. IV, there is no algorithm generalized to all types of underwater images due to the complicated underwater environment and lighting conditions. In general, the

fusion-based [31] achieves decent results, which benefits from the inputs derived by multiple pre-processing operations and a fusion strategy. In the proposed Water-Net, we also employ such a manner. Based on the characteristics of underwater image degradation, we generate three inputs by respectively applying White Balance (WB), Histogram Equalization (HE) and Gamma Correction (GC) algorithms to an underwater image. Specifically, WB algorithm is used to correct the color casts, while HE and GC algorithms aim to improve the contrast and lighten up dark regions, respectively. We directly employ the WB algorithm proposed in [31], whose effectiveness has been turned out. For the HE algorithm, we apply the *adapthisteq* function [67] provided by MATLAB to the L component in Lab color space, and then transform back into RGB color space. We set the Gamma value of GC algorithm to 0.7 empirically.

### B. Network Architecture

Water-Net employs a gated fusion network architecture to learn three confidence maps which will be used to combine the three input images into an enhanced result. The learned confidence maps determine the most significant features of inputs remaining in the final result. The impressive performance of the fusion-based underwater image enhancement method [31] also encourages us to explore the fusion-based networks.

The architecture of the proposed Water-Net and parameter settings are shown in Fig. 12. As a baseline model, the Water-Net is a plain fully CNN. We believe that the widely used backbones such as the U-Net architecture [68] and the residual network architecture [69] can be incorporated to improve the performance. We feed the three derived inputs and original input to the Water-Net to predict the confidence maps. Before performing fusion, we add three Feature Transformation Units

(FTUs) to refine the three inputs. The purpose of the FTU is to reduce the color casts and artifacts introduced by the WB, HE, and GC algorithms. At last, the refined three inputs are multiplied by the three learned confidence maps to achieve the final enhanced result:

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

where  $I_{en}$  is the enhanced result;  $\odot$  indicates the element-wise production of matrices;  $R_{WB}$ ,  $R_{HE}$ , and  $R_{GC}$  are the refined results of input after processing by WB, HE, and GC algorithms, respectively;  $C_{WB}$ ,  $C_{HE}$ , and  $C_{GC}$  are the learned confidence maps.

### C. Implementations

A random set of 800 pairs of the images extracted from the UIEB is used to generate the training set. We resize the training data to size  $112 \times 112$  due to our limited memory. Flipping and rotation are used to obtain 7 augmented versions of original training data. Resizing the training data to a fixed size is widely used in deep learning, such as image dehazing [66], salient object detection [70], etc. In contrast to image super-resolution (e.g., [71]) that has the same degradation in different regions of an input image, different regions of an underwater image may have different degradation. Thus, the contextual information of an underwater image is important for network optimization. This is the main reason why we do not use image patches to train our network. The rest 90 pairs of the images in the UIEB are treated as the testing set.

To reduce the artifacts induced by pixel-wise loss functions such as  $\ell_1$  and  $\ell_2$ , we minimize the perceptual loss function to learn the mapping function of underwater image enhancement. The perceptual loss can produce visually pleasing and realistic results, which has been widely used in image restoration and synthesis networks, such as image super-resolution [72], photographic image synthesis [73], etc.

Inspired by [72], we define the perceptual loss based on the ReLU activation layers (*i.e.*, layer `relu5_4`) of the pre-trained 19 layers VGG network [74]. Let  $\phi_j(x)$  be the  $j$ th convolution layer (after activation) of the VGG19 network  $\phi$  pretrained on the ImageNet dataset [75]. The perceptual loss is expressed as the distance between the feature representations of the enhanced image  $I_{en}$  and the reference image  $I_{gt}$ :

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

where  $N$  is the number of each batch in the training procedure;  $C_j H_j W_j$  represents the dimension of the feature maps of the  $j$ th convolution layer within the VGG19 network.  $C_j$ ,  $H_j$ , and  $W_j$  are the number, height, and width of the feature map.

We implemented the proposed Water-Net with TensorFlow on a PC with an Nvidia 1080Ti GPU. During training, a batch-mode learning method with a batch size of 16 was applied. The filter weights of each layer were initialized by standard Gaussian distribution. Bias was initialized as a constant. We used ADAM with default parameters for our network optimization. We initialized the learning rate to  $1e^{-3}$  and decreased the learning rate by 0.1 every 10,000 iterations

until the Water-Net converges. Our Water-Net can process an image with a size of  $640 \times 480$  within 0.128s (8FPS).

### D. Experiments

To demonstrate the advantages achieved by the proposed Water-Net, we compare it against several state-of-the-art underwater image enhancement methods. The experiments are conducted on the testing set which includes 90 underwater images and the challenging set including 60 underwater images. We show several results in Figs. 13 and 14.

In Fig. 13, the proposed Water-Net effectively removes the haze on the underwater images and remits color casts, while the competing methods introduce unexpected colors (*e.g.*, fusion-based [31], GDCP [40], histogram prior [45], Water CycleGAN [53], and Dense GAN [55]) and artifacts (*e.g.*, fusion-based [31], retinex-based [33], histogram prior [45], Water CycleGAN [53], and Dense GAN [55]) or have little effect on inputs (*e.g.*, blurriness-based [46]). In addition, it is interesting that our results even achieve better visual quality than the corresponding reference images (*e.g.*, more natural appearance and better details). This is because that the perceptual loss optimized Water-Net can learn the potential attributes of good visual quality from the large-scale real-world underwater image dataset. For the results on challenging set shown in Fig. 14, the proposed Water-Net produces visually pleasing results. By contrast, other methods tend to introduce artifacts, over-enhancement (*e.g.*, foregrounds), and color casts (*e.g.*, reddish or greenish color).

TABLE V  
FULL-REFERENCE IMAGE QUALITY ASSESSMENT IN TERMS OF MSE,  
PSNR, AND SSIM ON TESTING SET.

Method	MSE ( $\times 10^3$ ) $\downarrow$	PSNR (dB) $\uparrow$	SSIM $\uparrow$
fusion-based [31]	<b>1.1280</b>	<b>17.6077</b>	<b>0.7721</b>
retinex-based [33]	1.2924	17.0168	0.6071
GDCP [40]	4.0160	12.0929	0.5121
histogram prior [45]	1.7019	15.8215	0.5396
blurriness-based [46]	1.9111	15.3180	0.6029
Water CycleGAN [53]	1.7298	15.7508	0.5210
Dense GAN [55]	1.2152	17.2843	0.4426
Water-Net	<b>0.7976</b>	<b>19.1130</b>	<b>0.7971</b>

Table V reports the quantitative results of different methods in terms of MSE, PSNR, and SSIM on the testing set. The quantitative results are obtained by comparing the result of each method with the corresponding reference image. We discard the non-reference metrics designed for underwater image enhancement based on the conclusion drawn in Sec. IV. Our Water-Net achieves the best performance in terms of full-reference image quality assessment. In addition, instead of pairwise comparison, we conduct a user study to score the visual quality of the results on challenging set. This is because some images in the challenging set are too difficult to obtain satisfactory results following the procedure of reference image generation. Thus, we invited 50 participants (the same volunteers with the reference image generation) to score results. The scores have five scales ranging from 5 to 1 which represent “Excellent”, “Good”, “Fair”, “Poor” and “Bad”, respectively. The average scores of the results by each

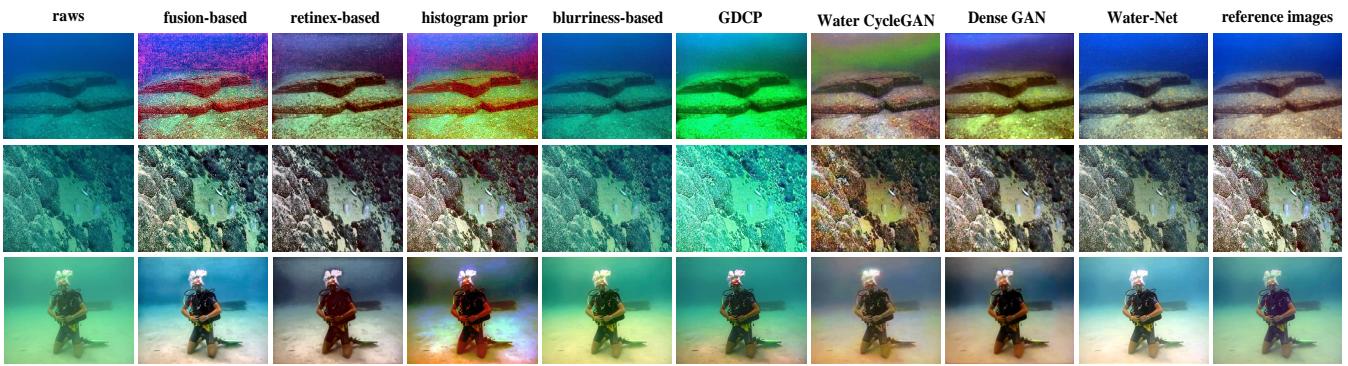


Fig. 13. Subjective comparisons on underwater images from testing set. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], histogram prior [45], blurriness-based [46], GDCP [40], Water CycleGAN [53], Dense GAN [55], the proposed Water-Net, and reference images.

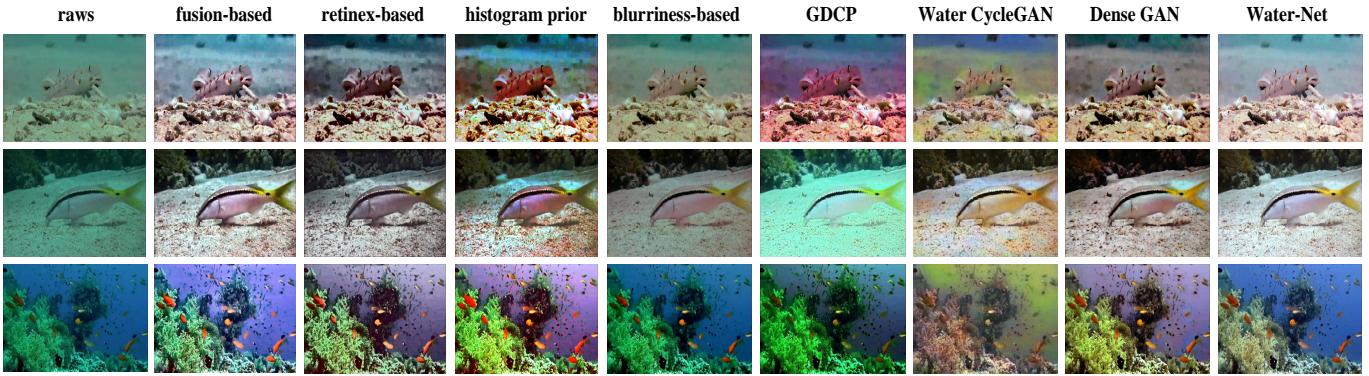


Fig. 14. Subjective comparisons on underwater images from challenging set. From left to right are raw underwater images, and the results of fusion-based [31], retinex-based [33], histogram prior [45], blurriness-based [46], GDCP [40], Water CycleGAN [53], Dense GAN [55], and the proposed Water-Net.

method on challenging set are shown in Table VI. Besides, we also provide the standard deviation of the results by each method on challenging set. We exclude the scores of Water CycleGAN [53] and Dense GAN [55] due to their obviously unpleasing results as shown in Fig. 14. Our Water-Net receives the highest average score and lowest standard deviation, which indicates our method produces better results from a subjective perspective and has more robust performance.

TABLE VI  
THE AVERAGE SCORES AND STANDARD DEVIATION OF DIFFERENT METHODS ON CHALLENGING SET.

Method	Average Score $\uparrow$	Standard Deviation $\downarrow$
fusion-based [31]	2.28	0.8475
retinex-based [33]	2.23	0.8720
GDCP [40]	1.90	0.8099
histogram prior [45]	2.08	0.7897
blurriness-based [46]	2.02	0.7762
Water-Net	2.57	0.7280

Qualitative and quantitative experiments demonstrate the effectiveness of the proposed Water-Net and also indicate the constructed dataset can be used for training CNNs. However, there is room for the improvement of underwater image enhancement. Besides, underwater images in the challenging set still cannot be enhanced well.

## VI. CONCLUSION, LIMITATIONS, AND FUTURE WORK

In this paper, we have constructed an underwater image enhancement benchmark dataset which offers large-scale real underwater images and the corresponding reference images. This benchmark dataset enables us to comprehensively study the existing underwater image enhancement methods, and easily train CNNs for underwater image enhancement. As analyzed in qualitative and quantitative evaluations, there is no method which always wins in terms of full- and no-reference metrics. In addition, effective non-reference underwater image quality evaluation metrics are highly desirable. To promote the development of deep learning-based underwater image enhancement methods, we proposed an underwater image enhancement CNN trained by the constructed dataset. Experimental results demonstrate the proposed CNN model performs favorably against the state-of-the-art methods, and also verify the generalization of the constructed dataset for training CNNs.

Although our reference image generation strategy can select visually pleasing results, there is a problem that affects the selection of reference images and further limits the performance of our network. Specifically, the effect of backscatter is difficult to be completely removed, especially for the backscatter in far distances. We use the state-of-the-art image enhancement algorithms to process the raw underwater images; however, the backscatter still cannot be completely removed in some cases. Despite our constructed dataset is not dominated by

such underwater images, the backscatter is always discouraged in the reference images. Moreover, some invited volunteers cannot identify the exponentially increasing effect of backscatter in scenes with large ranges. In summary, the main reasons for the shortcomings of our reference image generation are: 1) the existing algorithms follow inaccurate image formation models or assumptions which inherently limit the performance of underwater image enhancement; and 2) some volunteers do not understand the underwater imaging physical model, thus they may ignore the effect of the presence of backscatter in far ranges. Note that the use of inaccurate imaging models is a major problem which keeps the field of underwater computer vision at standstill. Fortunately, the recent work [49] sheds light on the future research of underwater vision.

In future work, we will extend the constructed dataset towards more challenging underwater images and underwater videos. Moreover, we will try to design a range map estimation network. The provided 1100 underwater images with range maps in [49] could be used for the range map estimation network training. With the estimated range maps, we will make full use of such key prior information to further improve the performance of underwater image enhancement network. Besides, inspired by recent work [6], [49], we believe that more physically reasonable underwater image enhancement algorithms will arise. At that time, we will re-organize the selection of the reference images from more reliable results and also further train the volunteers on what the degrading effects of attenuation and backscatter are, and what it looks like when either is improperly corrected. Additionally, the main purpose of constructing the real-world underwater dataset in this paper is to evaluate the state-of-the-art underwater image enhancement methods and provide paired training data for deep models. Since the full-reference metrics and training a deep model only need a single reference, we do not select multiple references or define the image quality level. However, the image quality level of multiple reference images does help in underwater image enhancement. Thus, we will provide multiple reference images for an underwater image and define the image quality level of their reference images when we re-organize the selection of the reference images.

## REFERENCES

- [1] J. Jaffe, "Underwater optical imaging: The past, the present, and the prospects," *IEEE J. Oceanic Eng.*, vol. 40, no. 3, pp. 683-700, 2015. 1
- [2] M. Sheinin and Y. Schechner, "The next best underwater view," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2016, pp. 3764-3773. 1
- [3] B. McGlamery, "A computer model for underwater camera systems," in *Ocean Optics VI*, pp. 221-231, 1980. 1
- [4] J. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE J. Oceanic Eng.*, vol. 15, no. 2, pp. 101-111, 1990. 1
- [5] W. Hou, S. Woods, E. Jarosz, *et al.*, "Optical turbulence on underwater image degradation in natural environments," *Appl. Opt.*, vol. 15, no. 14, pp. 2678-2686, 2012. 1
- [6] D. Akkaynak and T. Treibitz, "A revised underwater image formation model," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2018, pp. 6723-6732. 1, 3, 12
- [7] D. Akkaynak, T. Treibitz, T. Shlesinger, *et al.*, "What is the space of attenuation coefficients in underwater computer vision," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2017, pp. 568-577. 1, 6
- [8] M. Ludvigsen, B. Sortland, G. Johnsen, *et al.*, "Applications of georeferenced underwater photo mosaics in marine biology and archaeology," *J. Oceanography*, vol. 20, no. 4, pp. 140-149, 2007. 1
- [9] N. Strachan, "Recognition of fish species by colour and shape," *J. Image Vis. Comput.*, vol. 11, no. 1, pp. 2-10, 1993. 1
- [10] M. Yang, J. Hu, C. Li, *et al.*, "An in-depth survey of underwater image enhancement and restoration," *IEEE Access*, vol. 7, pp. 123638-123657, 2019. 1
- [11] C. Li, C. Guo, J. Guo, *et al.*, "PDR-Net: Perception-inspired single image dehazing network with refinement," *IEEE Trans. Multimedia*, 2019. 1
- [12] W. Wang, Q. Lai, H. Fu, *et al.*, "Salient object detection in the deep learning era: An in-depth survey," in *arXiv preprint arXiv: 1904.09146*, 2019. 1
- [13] C. Li, R. Cong, J. Hou, *et al.*, "Nested network with two-stream pyramid for salient object detection in optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9156-9166, 2019. 1
- [14] R. Cong, J. Lei, H. Fu, *et al.*, "An iterative co-saliency framework for RGBD images," *IEEE Trans. Cybernetics.*, no. 49, no. 1, pp. 233-249, 2019. 1
- [15] C. Guo, C. Li, J. Guo, *et al.*, "Hierarchical features driven residual learning for depth map super-resolution," *IEEE Trans. Image Process.*, vol. 28, no. 5, pp. 2545-2557, 2019. 1
- [16] R. Schettini and S. Corchs, "Underwater image processing: State of the art of restoration and image enhancement methods," *J. Advan. Signal Process.*, vol. 2010, pp. 1-14, 2010. 2
- [17] M. Han, Z. Lyu, and T. Qiu, "A review on intelligence dehazing and color restoration for underwater images," *IEEE Trans. Syst., Man, Cybern.*, Syst., pp. 1-13, 2018 (Early Access). 2
- [18] R. Cui, L. Chen, C. Yang, *et al.*, "Extended state observer-based integral sliding mode control for an underwater robot with unknown disturbances and uncertain nonlinearities," *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6785-6795, 2017. 2
- [19] S. Narasimhan and S. Nayar, "Contrast restoration of weather degraded images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25 no. 6, pp. 713-724, 2003. 2
- [20] N. Shashar, R. Hanlon, and A. deM. Petz, "Polarization vision helps detect transparent prey," *Nature*, vol. 393 no. 6682, pp. 222-223, 1998. 2
- [21] Y. Schechner and N. Karpel, "Clear underwater vision," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2004, pp. 536-543. 2
- [22] Y. Schechner and N. Karpel, "Recovery of underwater visibility and structure by polarization analysis," *IEEE J. Oceanic Eng.*, vol. 30, no. 3, pp. 570-587, 2005. 2
- [23] T. Treibitz and Y. Schechner, "Active polarization descattering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 3, pp. 385-399, 2009. 2, 4
- [24] C. Tan, G. Sluzek, and D. He, "A novel applications of range-gated underwater laser imaging system in near target turbid medium," *Optica and Lasers Eng.*, vol. 43, no. 9, pp. 995-1009, 2005. 2
- [25] H. Li, X. Wang, T. Bai, *et al.*, "Speckle noise suppression of range gated underwater imaging system," *Applid Optics*, vol. 38, no. 18, pp. 3937-3944, 2009. 2
- [26] T. Treibitz and Y. Schechner, "Turbid scene enhancement using multi-directional illumination fusion," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4662-4667, 2012. 2
- [27] Z. Murez, T. Treibitz, R. Ramamoorthi, *et al.*, "Photometric stereo in a scattering medium," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2015, pp. 3415-3423. 2
- [28] K. Iqbal, M. Odetayo, and A. James, "Enhancing the low quality images using unsupervised colour correction method," in *Proc. of IEEE Int. Conf. Syst., Man, Cybern.*, 2010, pp. 1703-1709. 2
- [29] A. Ghani and N. Isa, "Underwater image quality enhancement through integrated color model with Rayleigh distribution," *Appl. Soft Comput.*, vol. 27, pp. 219-230, 2015. 2
- [30] A. Ghani and N. Isa, "Enhancement of low quality underwater image through integrated global and local contrast correction," *Appl. Soft Comput.*, vol. 37, pp. 332-344, 2015. 2
- [31] C. Ancuti, C. O. Ancuti, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2012, pp. 81-88. 2, 4, 5, 6, 7, 8, 9, 10, 11
- [32] X. Fu, Z. Fan, and M. Ling, "Two-step approach for single underwater image enhancement," in *Symposium. of IEEE Intell. Signal Process. Commun. Syst.*, 2017, pp. 789-794. 2, 4, 5, 6, 7, 8
- [33] X. Fu, P. Zhang, Y. Huang, *et al.*, "A retinex-based enhancing approach for single underwater image," in *Proc. of IEEE Int. Conf. Image Process. (ICIP)*, 2014, pp. 4572-4576. 2, 4, 5, 6, 7, 8, 10, 11

- [34] S. Zhang, T. Wang, J. Dong, *et al.*, "Underwater image enhancement via extended multi-scale Retinex," *Neurocomput.*, vol. 245, no. 5, pp. 1-9, 2017. 2
- [35] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341-2353, 2011. 2, 4, 5
- [36] J. Chiang and Y. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1756-1769, 2012. 2
- [37] P. Drews-Jr, E. Nascimento, S. Botelho, *et al.*, "Underwater depth estimation and image restoration based on single images," *IEEE Comput. Graph. Appl.*, vol. 36, no. 2, pp. 24-35, 2016. 2, 4, 5, 6, 7, 8
- [38] H. Liu and L. Chau, "Underwater image restoration based on contrast enhancement," in *Proc. of IEEE Int. Digit. Signal Process.*, 2016, pp. 584-588. 2
- [39] C. Li, J. Guo, C. Guo, *et al.*, "A hybrid method for underwater image correction," *Pattern Rec. Lett.*, vol. 94, pp. 62-67, 2017. 2, 4, 5, 6, 7, 8
- [40] Y. Peng, T. Cao, and P. Cosman, "Generalization of the dark channel prior for single image restoration," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2856-2868, 2018. 3, 4, 5, 6, 7, 8, 10, 11
- [41] N. Carlevaris-Bianco, A. Mohan, and R. Eustice, "Initial results in underwater single image dehazing," in *Proc. of IEEE Int. Conf. Oceans*, 2010, pp. 1-8. 3
- [42] A. Galdran, D. Pardo, and A. Picn, "Automatic Red-Channel underwater image restoration," *J. Vis. Commu. and Image Repre.*, vol. 26, pp. 132-145, 2015. 3, 4, 5, 6, 7, 8
- [43] X. Zhao, T. Jin, and S. Qu, "Deriving inherent optical properties from background color and underwater image enhancement," *Ocean Eng.*, vol. 94, pp. 163-172, 2015. 3
- [44] C. Li, J. Guo, S. Chen, *et al.*, "Underwater image restoration based on minimum information loss principle and optical properties of underwater imaging," in *Proc. of IEEE Int. Conf. Image Process. (ICIP)*, 2016, pp. 1993-1997. 3
- [45] C. Li, J. Guo, R. Cong, *et al.*, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," *IEEE Trans. Image Process.*, vol. 25, no. 12, pp. 5664-5677, 2016. 3, 4, 5, 6, 7, 8, 9, 10, 11
- [46] Y. Peng and P. Cosman, "Underwater image restoration based on image blurriness and light absorption," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1579-1594, 2017. 3, 4, 5, 6, 7, 8, 10, 11
- [47] D. Berman, T. Treibitz, and S. Avidan, "Diving into haze-lines: Color restoration of underwater images," in *Proc. of Briti. Mach. Vis. Conf. (BMVC)*, 2017, pp. 1-12. 3
- [48] Y. Wang, H. Liu, and L. Chau, "Single underwater image restoration using adaptive attenuation-curve prior," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 3, pp. 992-1002, 2018. 3
- [49] D. Akkaynak and T. Treibitz, "Sea-thru: A method for removing water from underwater images," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2019, pp. 1682-1691. 3, 4, 12
- [50] S. Anwar and C. Li, "Diving deeper into underwater image enhancement: A survey," in *arXiv preprint arXiv: 1907.07863*, 2019. 3
- [51] J. Li, K. Sinner, R. Eustice, *et al.*, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 387-394, 2018. 3, 4
- [52] C. Li, S. Anwar, and F. Porikli, "Underwater scene prior inspired deep underwater image and video enhancement," *Pattern Rec.*, vol. 98, pp. 107038-107049, 2020. 3, 4
- [53] C. Li, J. Guo, and C. Guo, "Emerging from water: Underwater image color correction based on weakly supervised color transfer," *IEEE Signal Process. Lett.*, vol. 25, no. 3, pp. 323-327, 2018. 3, 4, 10, 11
- [54] Y. Zhu, T. Par, P. Isola, *et al.*, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. of IEEE Int. Conf. Comput. Vis.(ICCV)*, 2017, pp. 2242-2251. 3
- [55] Y. Guo, H. Li, and P. Zhuang, "Underwater image enhancement using a multiscale dense generative adversarial network," *IEEE J. Oceanic Eng.*, pp. 1-9, 2019. 3, 4, 10, 11
- [56] Z. Wang, A. Bovik, H. Sheikh, *et al.*, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, 2004. 3
- [57] N. Hautiere, J. Tarel, D. Aubert, *et al.*, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *Image Anal. Stereo.*, vol. 27, no. 2, pp. 87-95, 2011. 3
- [58] T. Aydin, R. Mantiuk, K. Myszkowski, *et al.*, "Dynamic range independent image quality assessment," *ACM Trans. Graphics*, vol. 27, no. 3, pp. 69-79, 2008. 3
- [59] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6062-6071, 2015. 3, 6, 8
- [60] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE J. Ocean Eng.*, vol. 41, no. 3, pp. 541-551, 2016. 3, 6, 8
- [61] J. Xiao, J. Hays, K. Ehinger, *et al.*, "Sun database: Large-scale scene recognition from abbey to zoo," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2010, pp. 3485-3492. 4
- [62] D. Berman, D. Levy, S. Avidan, *et al.*, "Underwater single image color restoration using haze-lines and a new quantitative dataset," in *arXiv preprint arXiv:1811.01343*, 2018. 4
- [63] A. Duarte, F. Codevilla, J. Gaya, *et al.*, "A dataset to evaluate underwater image restoration methods," in *Proc. of IEEE Int. Conf. Oceans*, 2016, pp. 1-6. 4
- [64] H. Blasinski, T. Lian, and E. Farrell, "Underwater image systems simulation," in *Imaging and Applied Optics*, 2017, pp. 1-2. 4
- [65] H. Blasinski and E. Farrell, "A three parameter underwater image formation model," *Electron. Imag.*, vol. 18, pp. 1-8, 2016. 4
- [66] W. Ren, S. Liu, H. Zhang, *et al.*, "Single image dehazing via multi-scale convolutional neural networks," in *Proc. of Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 154-169. 4, 5, 10
- [67] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics Gems*, pp. 474-485, 1994. 9
- [68] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. of Med. Image Comput. Comput. Assist. Inter. (MICCAI)*, 2015, pp. 234-241. 9
- [69] K. He, X. Zhang, S. Ren, *et al.*, "Deep residual learning for image recognition," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2016, pp. 770-778. 9
- [70] H. Chen and Y. Li, "Progressively complementarity-aware fusion network for RGB-D salient object detection," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2018, pp. 3051-3060. 10
- [71] C. Dong, C. Loy, K. He, *et al.*, "Learning a deep convolutional network for image super-resolution," in *Proc. of Eur. Conf. Comput. Vis. (ECCV)*, 2014, pp. 184-199. 10
- [72] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. of Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 694-711. 10
- [73] Q. Chen and V. Koltun, "Photographic image synthesis with cascaded refinement networks," in *Proc. of IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 1511-1520. 10
- [74] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. of Int. Conf. Machine Learn. (ICML)*, 2015, pp. 1-14. 10
- [75] J. Deng, W. Dong, R. Socher, *et al.*, "Imagenet: A large-scale hierarchical image database," in *Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR)*, 2009, pp. 248-255. 10



**Chongyi Li** received his Ph.D. degree with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China in June 2018. During 2016 to 2017, he was a joint-training Ph.D student in Australian National University, Australia.

He is currently a Postdoc Research Fellow at the Department of Computer Science, City University of Hong Kong (CityU), Hong Kong SAR, China. His current research focuses on image processing, computer vision, and deep learning, particularly in the domains of image restoration and enhancement.

He is a Guest Editor of special issues for the Signal Processing: Image Communication.



**Chunle Guo** received his B.S. degree from School of Electronic Information Engineering in Tianjin University. He is currently pursuing his Ph.D. degree with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China. His current research focuses on image processing and computer vision, particularly in the domains of deep learning-based image restoration and enhancement.



**Sam Kwong** (M'93-SM'04-F'13) received the B.S. degree in electrical engineering from The State University of New York at Buffalo, Buffalo, NY, USA, in 1983, the M.S. degree from the University of Waterloo, Waterloo, ON, Canada, in 1985, and the Ph.D. degree from the University of Hagen, Germany, in 1996. From 1985 to 1987, he was a Diagnostic Engineer with Control Data Canada, Mississauga, ON, Canada. He was a Member of Scientific Staff with Bell Northern Research Canada, Ottawa. In 1990, he became a Lecturer with the Department of Electronic Engineering, City University of Hong Kong, where he is currently a Chair Professor with the Department of Computer Science. His research interests are video and image coding and evolutionary algorithms.



**Wengi Ren** is an assistant professor in Institute of Information Engineering, Chinese Academy of Sciences, China. He received his Ph.D. degree from Tianjin University in 2017. During 2015 to 2016, he was a joint-training Ph.D. student in Electrical Engineering and Computer Science at the University of California, Merced, CA, USA. His research interest includes image/video analysis and enhancement, and related vision problems.



interests include computer vision and intelligent video analysis, multimedia information processing, saliency detection and segmentation, remote sensing image interpretation, and deep learning.

He is a Reviewer for the IEEE TIP, TMM, and TCSVT, etc, and is a Guest Editor of special issues for the Signal Processing: Image Communication. Dr. Cong was a recipient of the Best Student Paper Runner-Up at IEEE ICME in 2018, the First Prize for Scientific and Technological Progress Award of Tianjin Municipality, and the Excellent Doctoral Degree Dissertation Award from BSIG.



**Dacheng Tao** (F'15) is Professor of Computer Science and ARC Laureate Fellow in the School of Computer Science and the Faculty of Engineering, and the Inaugural Director of the UBTECH Sydney Artificial Intelligence Centre, at The University of Sydney. His research results in artificial intelligence have expounded in one monograph and 200+ publications at prestigious journals and prominent conferences, such as IEEE T-PAMI, IJCV, JMLR, AAAI, IJCAI, NIPS, ICML, CVPR, ICCV, ECCV, ICDM, and KDD, with several best paper awards.

He received the 2018 IEEE ICDM Research Contributions Award and the 2015 Australian Scopus-Eureka prize. He is a Fellow of the IEEE and the Australian Academy of Science.



**Junhui Hou** (S'13-M'16) received the B.Eng. degree in information engineering (Talented Students Program) from the South China University of Technology, Guangzhou, China, in 2009, the M.Eng. degree in signal and information processing from Northwestern Polytechnical University, Xian, China, in 2012, and the Ph.D. degree in electrical and electronic engineering from the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2016. He has been an Assistant Professor with the Department of Computer Science, City University of Hong Kong, since 2017. His research interests fall into the general areas of visual signal processing, such as image/video/3D geometry data representation, processing and analysis, semi-supervised/unsupervised data modeling for clustering/classification, and data compression and adaptive transmission.

Dr. Hou was the recipient of several prestigious awards, including the Chinese Government Award for Outstanding Students Study Abroad from China Scholarship Council in 2015, and the Early Career Award from the Hong Kong Research Grants Council in 2018. He is serving/served as an Associate Editor for The Visual Computer, an Area Editor for Signal Processing: Image Communication, the Guest Editor for the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, the Journal of Visual Communication and Image Representation, and Signal Processing: Image Communication, and an Area Chair of ACM International Conference on Multimedia (ACM MM) 2019 and IEEE International Conference on Multimedia & Expo (IEEE ICME) 2020.