


RESEARCH PAPER

Open Access



WaterPairs: a paired dataset for underwater image enhancement and underwater object detection

Long Chen^{1†}, Xirui Dong^{2†}, Yunzhou Xie³ and Sen Wang^{1*}

Abstract

Due to its importance in marine engineering and aquatic robotics, underwater image enhancement works as a pre-processing step to improve the performance of high-level vision tasks such as underwater object detection and recognition. Although several studies exhibit that underwater image enhancement algorithms can boost the detection accuracy of detectors, no work has focused on studying the relationship between these two tasks. This is mainly because current underwater datasets lack either bounding box annotations or high-quality reference images, based on which detection accuracy or image quality assessment metrics are calculated. To examine how underwater image enhancement methods affect underwater object detection tasks, we provide a large-scale underwater object detection dataset with both bounding box annotations and high-quality reference images, namely, the WaterPairs dataset. The WaterPairs dataset offers a platform for researchers to comprehensively study the influence of underwater image enhancement algorithms on underwater object detection tasks. We will release our dataset at <https://github.com/lanDragon/WaterPairs> once this paper is accepted.

Keywords Underwater datasets, Underwater object detection, Underwater image enhancement, Reference image generation

1 Introduction

In the past few years, underwater object detection (UOD) (Foresti and Gentili 2000) has drawn considerable attention in marine engineering and aquatic robotics. Because of the complex underwater environment and frequently changing illumination conditions, object detection in underwater scenes is a demanding task. The underwater images suffer from severe

wavelength-dependent absorption and scattering, which reduces visibility, decreases contrast, and even introduces color casts (Zhang et al. 2022; Zhuang et al. 2022). These adverse effects restrict many practical applications of underwater images and videos in marine biology, archaeology, and ecology. Thus, many underwater image enhancement (UIE) algorithms are used as a preprocessing step for UOD tasks to enhance the detection accuracy of detectors by increasing the quality of underwater images (Bazeille et al. 2006; Schettini and Corchs 2010).

Despite the prolific literature, comprehensive studies and insightful analyses of the relationship between UIE and UOD tasks are still scant, mainly because of the lack of publicly available underwater image datasets with both bounding box annotations and reference images (i.e., the underwater images without degradation). Since there are no reference images, previous literature (Liu et al. 2020) only investigated how UIE algorithms affect

[†]Long Chen and Xirui Dong contributed equally to this work.

*Correspondence:

Sen Wang
sen.wang@imperial.ac.uk

¹ Faculty of Engineering, Imperial College London, Electrical Engineering, South Kensington Campus, London SW7 2AZ, UK

² School of Professional and Continuing Education, King's College London, Strand Campus, London WC2B 4BG, UK

³ Faculty of Natural Sciences, Imperial College London, Faculty Building, South Kensington Campus, London SW7 2AZ, UK

UOD tasks by studying the relationships between non-reference image quality assessment metrics (Panetta et al. 2015; Yang and Sowmya 2015) and the detection accuracy. However, nonreference image quality evaluation metrics can only explain some characteristics of image quality and are not always consistent with human subjective perception (Liu et al. 2020). A comprehensive investigation of the relationship between the two tasks should also focus on the relationship between detection accuracy and full-reference image evaluation metrics (Wang et al. 2004, 2015), which can extensively assess the characteristics of image quality with respect to color, texture, image content, and structure. However, reference images are necessary when conducting full-reference image quality evaluations. In recent years, several underwater image synthesis (UIS) algorithms (Fabbri et al. 2018; Li et al. 2020a) have been proposed to synthesize underwater images from high-quality in-air images. Then another UIE model is trained on the image pairs to improve the visibility of underwater images. However, the synthetic images are not realistic enough and greatly affect the performance of late UIE models. In contrast, Li et al. (2019) used 11 different UIE algorithms to improve the underwater images and selected high-quality reference images from the 11 enhanced results using human subjective perception, i.e., the perception of the human visual system. Nevertheless, subjective perception can be ambiguous and tendentious because different people may have different preferences and biases (Pronin 2007). In addition, human perception cannot perceive minor differences that are present in two visually similar images. To compensate for the deficiency of subjective perception, we incorporate objective assessment to select high-quality reference images, which is more robust and dependable than subjective perception alone.

In this work, we construct a paired underwater image dataset called WaterPairs, which provides underwater images and corresponding high-quality reference images. More importantly, bounding box annotations

are also provided for the objects in the underwater images. The underwater images come from the real underwater dataset OUC-VISION (Jian et al. 2017), which provides only bounding box-level annotations without high-quality reference images. To produce high-quality reference images, we propose a novel hybrid reference image generation algorithm that combines human subjective perception and computational objective assessment. Figure 1 exhibits several sampling underwater images and the corresponding reference images produced by our hybrid reference image generation algorithm. The raw underwater images in the WaterPairs dataset suffer from diverse degrees of haze and contrast decrease. On the other hand, the corresponding reference images are characterized by natural color, enhanced visibility, and suitable brightness. With this dataset, we perform a comprehensive qualitative and quantitative study of the state-of-the-art UIE, UOD, and UIS algorithms. Most importantly, we investigate how UIE algorithms influence UOD tasks to gain insights into their performance and shed light on future research. The main contributions of this study are summarized as follows:

- (1) We offer a large-scale underwater dataset called WaterPairs for training and assessing UIE, UOD, and UIS algorithms. To the best of our knowledge, this is the first underwater dataset that provides both underwater images and corresponding high-quality reference images together with object-level bounding box annotations.
- (2) To produce high-quality reference images for the underwater images, we propose a novel reference image generation method that integrates both subjective human perception and computational objective assessment.
- (3) We perform extensive experiments to examine the relationships between UIE and UOD and obtain some interesting conclusions that may provide

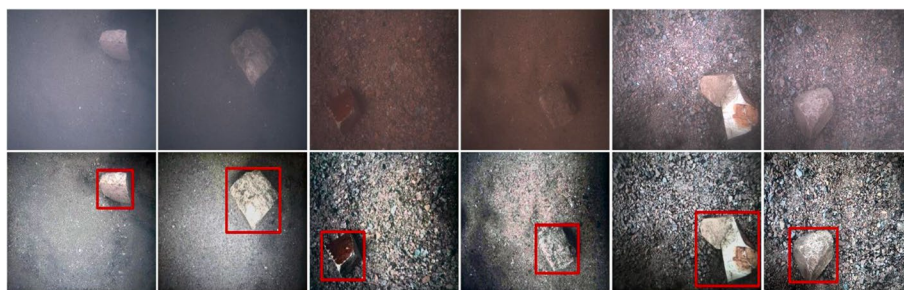


Fig. 1 Sampling images from our WaterPairs dataset. The top row shows the raw underwater images taken in various underwater scenes, and the bottom row presents the corresponding high-quality reference images and bounding box annotations

meaningful insights for the future development of this research field.

The rest of the paper is structured as follows: Section 2 summarizes the related works. Section 3 describes the proposed channel-weighted skip connection network. Section 4 reports and discusses the experimental results.

2 Related work

2.1 Underwater image enhancement

UIE plays an important role in practical applications that explore and develop the underwater world, such as autonomous underwater vehicle (Clark et al. 2013), unmanned underwater vehicle (Xu et al. 2015), and remotely operated vehicle (Bogue 2015) navigation. Various UIE methods have been proposed and can be categorized into three. The first line of research is to modify the image pixel values to enhance the image contrast, remove haze and correct color casts. It can be divided into spatial domain adjustment and transform domain adjustment. The spatial domain methods (Ancuti et al. 2012; Fu et al. 2017) perform adjustment directly in captured underwater images. The transform domain methods (Singh et al. 2015) first transform the captured underwater image into a specific domain and then perform adjustments for haze removal and color correction. These methods can enhance the visual quality to some extent, but they may degrade details, accentuate noise, introduce artifacts, and cause color distortions.

The second line is physical model-based methods (Galdran et al. 2015; Li et al. 2017a), which takes UIE as the inverse problem of underwater image degradation. It first constructs and estimates a physical image degradation process and then recovers the potential high-quality image from the estimated physical degradation model. To estimate the parameters of the underwater image degradation model, many UIE algorithms (Drews et al. 2013; Peng et al. 2018) adapted the classic dark channel prior (DCP) (He et al. 2010), which was designed for dehazing in natural scenes to underwater scenes. However, these priors do not always work in some cases. For instance, DCP-based UIE algorithms showed limited improvement in visual quality or even aggravated degradation in underwater images containing white objects or regions. This is because the DCP is no longer valid in scenes where white objects or regions are present.

The third line is deep learning-based UIE algorithms, which can be trained using a large-scale underwater dataset. Li et al. (2022a) proposed a feature pyramid attention network to remove motion blur and restore blurry underwater images. In the work of Li et al. (2022b), a two-stage generative adversarial network

was put forward to remove the blur effects. Zhang et al. (2023) proposed a weighted wavelet-based UIE framework to address quality degradation issues, while Li et al. (2022c) constructed a UIE framework using an adaptive color restoration module and a haze-line-based dehazing module to restore color distribution and remove the haze effects. Due to the lack of training pairs, Li et al. (2017b) proposed an underwater image synthesis model called WaterGAN to convert high-quality in-air images and corresponding depth images into underwater-like images. Then, these synthetic image pairs were used to train another deep UIE network. Motivated by Cycle-Consistent Adversarial Networks (Zhu et al. 2017), which allows learning the mutual mappings between two domains from unpaired data, Fabbri et al. (2018) proposed a weakly supervised underwater image synthesis model to synthesize underwater images from high-quality in-air images and then utilized these synthetic image pairs to train another deep UIE network. In contrast, Li et al. (2020a) generated training data by harnessing a physically underwater image degradation model and a fixed set of predefined parameters. However, the performance of the deep UIE network heavily depends on the quality of the synthetic images, which cannot be perfectly solved by previous underwater image synthesis methods. Thus, the availability and volume of training data are major bottlenecks in the development of deep learning-based UIE methods. To obtain reliable high-quality training data, Li et al. (2019) collected a real-world underwater dataset and processed it using multiple image enhancement methods. Afterward, they invited volunteers to select satisfactory reference images by conducting pairwise comparisons. This method produced, at least to some extent, trustworthy reference images by applying human subjective perception.

2.2 Underwater image quality evaluation

Image quality assessment techniques have important applications in UIE tasks and are especially beneficial for the development of UIE algorithms. They can be divided into subjective assessment and objective assessment. Subjective assessment is the most reliable method for quantifying the perceptual quality of content because, in most cases, such content is meant to be viewed by humans (Seshadrinathan et al. 2010; Mohammadi et al. 2014). However, subjective assessment, which depends on human judgment, can be ambiguous and tendentious because subjective perceptions of different observers are inconsistent.

The objective image quality assessment metrics are utilized to measure some important characteristics of the images using statistical numbers, and they can be further divided into full-reference image quality assessment

metrics (Wang et al. 2004) and nonreference image quality assessment metrics (Yang and Sowmya 2015). Most previous works (Drews et al. 2013; Peng et al. 2018) only used the nonreference metrics to assess UIE algorithms because underwater datasets do not offer reference images. The underwater color image quality evaluation metric (UCIQE) (Yang and Sowmya 2015) and underwater image quality measure (UIQM) (Panetta et al. 2015) are two widely used nonreference metrics. UCIQE quantifies nonuniform color casts, blurring, and low contrast and then integrates these three components in a linear manner. UIQM includes three attribute measures: colorfulness measure, sharpness measure, and contrast measure. Full-reference metrics are commonly used in cases where reference images are present. For instance, the peak signal-to-noise ratio (PSNR) measures the similarity between the enhanced underwater images and the reference images in terms of content, and the SSIM measures the structure and texture similarity of the enhanced images and the reference images.

One major limitation of contemporary objective assessment metrics is that they are usually sensitive to only one or limited types of distortions while ignoring the evaluation of distortions of other types, e.g., color distortion, blurry appearance, or decreasing contrast in the underwater images. Thus, tremendous efforts are highly demanded to develop more effective image quality assessment methods.

3 Reference image generation

In this section, we construct a large-scale underwater image dataset called WaterPairs, which provides underwater images, corresponding reference images, and bounding box annotations. First, we introduce the selection of underwater images and then present a novel method for producing reference images by integrating subjective perception and objective assessment.

3.1 Selection of underwater images

We aim to build a large-scale underwater dataset that enables researchers to assess different UIE, UOD, and UIS algorithms and, more importantly, explore how UIE algorithms affect UOD algorithms. Thus, the underwater dataset should contain underwater images, high-quality reference images and bounding box annotations. We set three objectives when constructing the underwater dataset: (1) The number of underwater images should be sufficiently large, and bounding box level annotations should be provided. (2) The underwater images should suffer from a diversity of degradation. (3) The quality of the reference images should be assured so that the image pairs allow fair evaluation of different UIE and UIS algorithms.

To realize the first two objectives, we choose a large real underwater dataset, OUC-VISION (Jian et al. 2017), which provides underwater images and bounding box annotations. This dataset contains 4400 underwater images captured under different illuminations simulated by a specially designed lighting system. Moreover, three degrees of turbidity variations, i.e., limpidity, medium, and turbidity, are simulated by adding soil to the water. Thus, the underwater images of OUC-VISION suffer from a diversity of illumination variations and turbidity variations. The images have a resolution of 486×648 pixels. Figure 2 presents some examples of the raw underwater images in the OUC-VISION dataset. The underwater images in the OUC-VISION dataset have different characteristics, such as different color casts, decreased contrast, and haze levels. To obtain truth-worthy reference images, we propose a novel hybrid reference image generation method that incorporates both subjective perception and objective assessment.

3.2 Hybrid reference image generation

Previous work (Liu et al. 2020) first enhanced underwater images using different UIE algorithms, and then multiple observers were invited to choose high-quality reference images from the enhanced results. However, using only subjective human perception to choose images can be ambiguous and tendentious: (1) In many practical cases, the compared images appear to have the same visual quality that the observers have difficulties in distinguishing and choosing the best one. For instance, as shown in Fig. 3, two observers select the enhanced images of different UIE methods as the final reference images because the visual appearance of the two results is extremely similar or presents respective good characteristics. (2) The subjective perception is associated with the human visual system: different observers may have different preferences and biases, and no universal standards exist. As shown in the top row of Fig. 3, the two observers have different preferences and choose different enhanced images as the reference images. To solve the ambiguity and bias issues, we propose a hybrid reference image generation method that combines subjective human perception and a novel pairwise objective assessment metric.

The pairwise objective assessment metric. In particular, when the observers cannot make a decision according to their subjective perceptions in the pairwise comparison, a novel designed pairwise objective assessment metric is used to help in selecting the better one from the two enhanced results. The pairwise objective assessment metric is calculated on the union scores of UIQM and UCIQE. For the two compared UIE algorithms, their pairwise objective scores P_Score_1 and P_Score_2 are expressed as Eqs. (1) and (2), respectively.



Fig. 2 Examples of the raw underwater images in the OUC-VISION dataset. These images have different illuminations and haze degrees because they were taken under different underwater environments

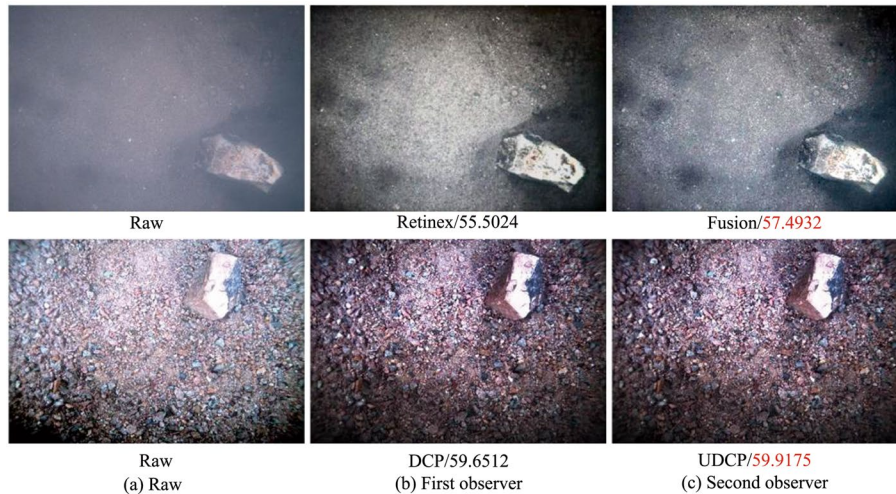


Fig. 3 Inconsistencies of the subjective perceptions of different observers

$$P_Score_1 = UCIQE_1 + normUIQM_1, \tag{1}$$

$$P_Score_2 = UCIQE_2 + normUIQM_2, \tag{2}$$

$normUIQM_1$ and $normUIQM_2$ are the normalized UIQM scores of the two UIE algorithms, which are expressed as Eqs. (3) and (4).

$$normUIQM_1 = \frac{UCIQE_1 + UCIQE_2}{UIQM_1 + UIQM_2} * UIQM_1, \tag{3}$$

$$normUIQM_2 = \frac{UCIQE_1 + UCIQE_2}{UIQM_1 + UIQM_2} * UIQM_2. \tag{4}$$

For the novel objective metric, we suppose UIQM and UCIQE as equally important, so we first normalize UIQM as $normUIQM$ and then combine it with the

UCIQE score to form the final pairwise objective score (Table 1).

Process of reference image generation. We first enhance the underwater images using 11 image enhancement methods, including 7 physical-model-based UIE methods (i.e., DCP (He et al. 2010), UDCP (Drews et al. 2013), GDCP (Peng et al. 2018), Blurriness (Peng and Cosman 2017), Regression (Li et al. 2017a), Red-Channel (Galdran et al. 2015), and Histogram (Li et al. 2016)), 3 model-free UIE methods (i.e., Fusion (Ancuti et al. 2012), Twostep (Fu et al. 2017), and Retinex (Fu et al. 2014)), and 1 commercial application for enhancing underwater images (i.e., Dive+). We do not employ deep learning-based UIE methods because we have no training image pairs. Finally, we totally obtain 11×4400 enhanced results. With the raw underwater images and the enhanced results, we invite 28 observers, all of whom

Table 1 Statistics of observers involved in the reference image generation process, including age, gender, and experience. IE refers to whether the observer has image enhancement experiences

	Group	IE Group	Non-IE Group
Age	18-25	2	4
	26-30	7	12
	>30	1	2
Gender	Male	8	16
	Female	2	2
IE experiments	≤ 1 year	4	18
	> 1 year	6	0

are students with image processing and computer vision experience, to conduct pairwise comparisons. Among all the observers, 10 of them have experience in image enhancement. The detailed statistics of the observers are summarized in Table 1. They are allowed to draw support from the pairwise objective assessment metric when they cannot make a decision on two ambiguous images in the pairwise comparison. There is no time constraint for observers, and zoom-in operation is allowed.

The generation of reference images has three stages: (1) the reference image selection by a single observer; (2) checking the reference images again and removing unsatisfactory images; and (3) combining the results of all the observers to obtain the final reference images. For each raw underwater image, the observer is shown two randomly selected enhanced results for pairwise comparison at one time. The

observer needs to select the preferred one or press the button that helps to select the better image using the pairwise objective metric. The result that wins the pairwise comparison is compared again in the next round until the best one is selected. After the observer finishes the selection work, he/she inspects the reference images set again and removes the unsatisfactory images. Afterward, the reference images of all observers are combined. For each raw underwater image, if more than half the number of observers remove its corresponding reference images, this underwater image and its reference images will be removed from the final dataset. Finally, the enhanced image selected by more than 50% of observers is selected as the final reference image.

We achieve 3698 available reference images that have higher quality than the results of any individual UIE method. To visualize the process of reference image generation, we present some cases in which the results of some methods are shown and indicate which one is the final reference image in Fig. 4. Moreover, the percentage of reference images from the results of different methods is presented in Table 2, which reveals that the best method is the Retinex algorithm, whereas the second best method is the Fusion algorithm.

4 Evaluation of the UIE and UIS algorithms on the WaterPairs dataset

4.1 Evaluation of the UIE algorithms on the WaterPairs dataset

We assess different UIE algorithms on the WaterPairs dataset. We resize all the images into 512×512 pixels and

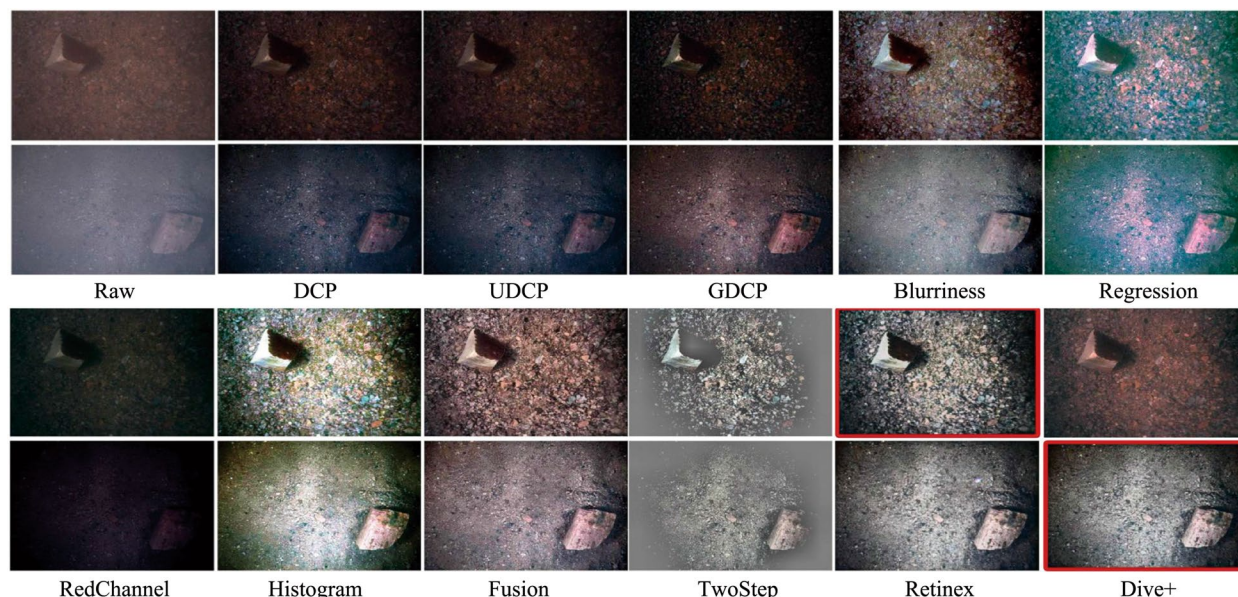


Fig. 4 Results generated by using different methods. The images with red bounding boxes refer to the final selected reference images. From left to right are raw underwater images and the results of DCP, UDCP, GDCP, Blurriness, Regression, RedChannel, Histogram, Fusion, TwoStep, Retinex, and Dive+. Red boxes refer to the final reference images

Table 2 Percentage of the reference images from the results of different methods

Method	DCP	UDCP	GDCP	Blurriness	Regression	RC	Histogram	Fusion	TwoStep	Retinex	Dive+
Percentage (%)	3.68	4.50	1.30	4.40	0.00	0.00	0.70	25.10	0.00	41.72	18.60

divide the WaterPairs dataset into a training set with 2500 image pairs and a testing set with 1198 image pairs. Figure 5 illustrates the qualitative comparisons of different UIE algorithms on underwater images under different light conditions. The top image is captured under light condition 1 and suffers from serious reddish color distortion, the middle image incorporates slight color distortion and haze effect due to light condition 2, and the bottom image incorporates an evident haze effect and minor color distortion due to light condition 3. We observe that none of the physical model-based methods can solve the reddish color distortion. This is because the presence of reddish underwater images violates the physical prior. In water, red light first disappears because it has the longest wavelength, followed

by green light and then blue light. Such selective attenuation in water leads to greenish and bluish underwater images and seldom reddish underwater images. Moreover, among all the physical model-based algorithms, Regression, Histogram, and RedChannel cannot adequately deal with underwater images under all kinds of light conditions. Regression introduces serious bluish color distortion due to its inaccurate color correction algorithm, and Histogram introduces greenish color distortion due to its histogram distribution prior. RedChannel greatly decreases the brightness, which severely smears the details of images. Moreover, TwoStep, a nonphysical model-based algorithm, also fails under all kinds of light conditions. It over enhances the contrast and generates unnatural images. In contrast,

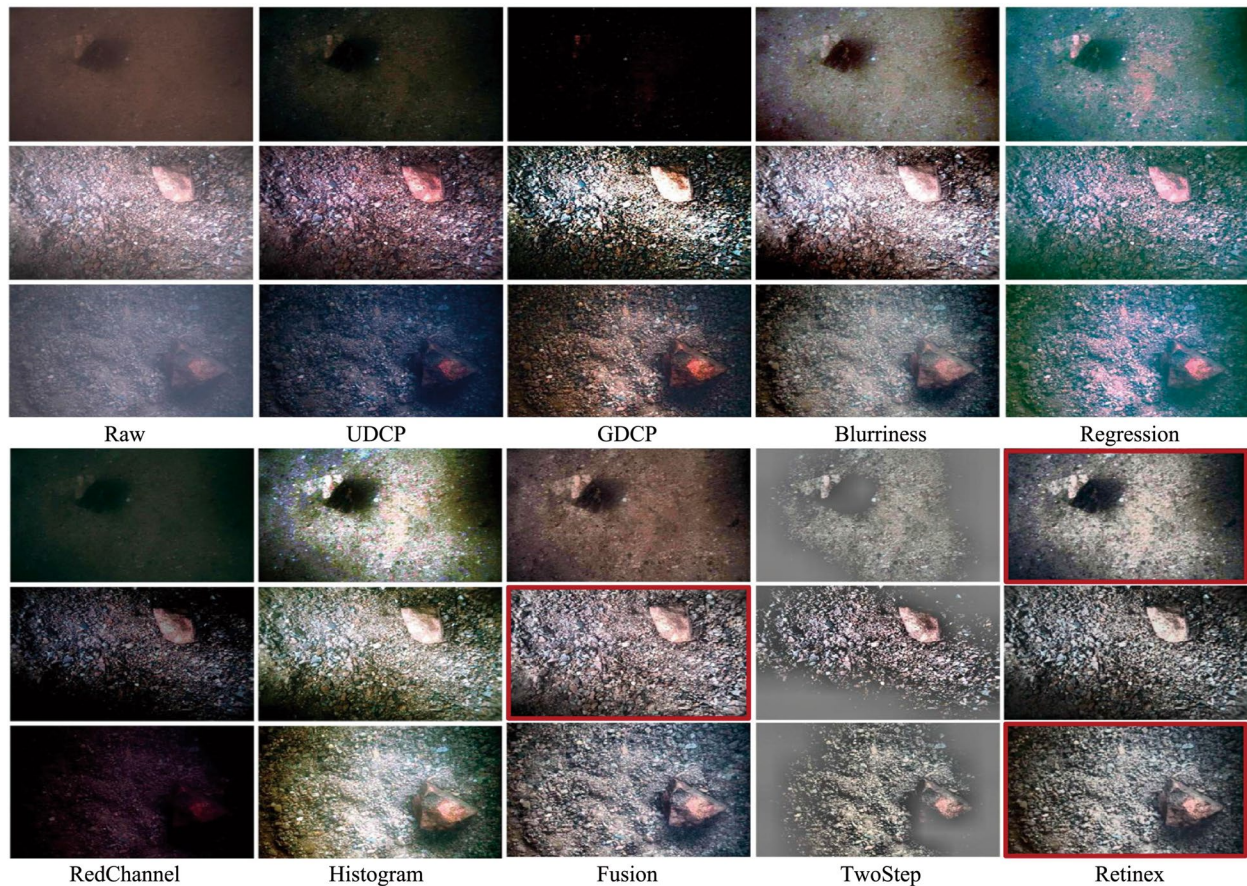


Fig. 5 Results generated by using different methods. The images with red bounding boxes refer to the final selected reference images. From left to right are raw underwater images and the results of UDCP, GDCP, Blurriness, Regression, RedChannel, Histogram, Fusion, TwoStep, and Retinex. Red boxes refer to the final reference images

OurPatch deals well with all kinds of underwater images in terms of both color distortion and haze effect, while the rest only work in special scenes. For instance, GDCP and Fusion remove the haze effects and greatly enhance the visibility of underwater images captured under light conditions 2 and 3. UDCP greatly removes haze; however, it introduces a bluish color tone in the images captured under light condition 2 and a reddish color tone in the images captured under light condition 3. Blurriness greatly removes haze on images captured under all three light conditions but fails to remit the color casts in images captured under light condition 1. These physical model-based methods all fail in some underwater images captured under specific light conditions because of the limitations of the priors used in them. Among the nonphysical model-based methods, Retinex greatly removes haze and color cast in all kinds of underwater images, but its results suffer from limited saturation.

Table 3 summarizes the quantitative scores of the different UIE algorithms in the testing set of WaterPairs. Fusion has the best MSE, PSNR, and PCQI scores,

whereas Retinex has the best SSIM score. Moreover, we train multiple SSD frameworks (Liu et al. 2016) using the enhanced images of different UIE algorithms and report the mean Average Precision (mAP). In terms of mAP, Retinex has the best detection accuracy of 87.2 mAP.

To examine whether all the UIE methods enhance the performance of the detection network, we train multiple detection networks using the enhanced images generated from different UIE methods and test them on the enhanced images. We also train two detection networks using the raw underwater images (denoted as 'Baseline') and high-quality images (denoted as 'OurWaterPairs') in the WaterPairs dataset. Figure 6 shows the performance of detection networks trained using the enhanced images generated from different UIE methods. The detection network trained using our high-quality reference images has the best detection performance because the quality of the reference images is much better than that of the enhanced images produced by other UIE algorithms. Not all UIE algorithms enhance the detection performance.

Table 3 Full-reference image quality and detection accuracy evaluations of different UIE algorithms on the WaterPairs dataset

Methods	UDCP	GDCP	Blurriness	Regression	RedChannel	Histogram	Fusion	TwoStep	Retinex
MSE	3.3769	2.2945	0.6279	0.4150	7.2612	0.5640	0.2508	1.5833	0.4083
PSNR	13.3206	15.9326	20.8400	22.2458	9.6755	20.9896	28.4754	16.2539	26.5359
SSIM	0.5130	0.6407	0.7472	0.5691	0.1696	0.7602	0.8905	0.6002	0.8789
PCQI	0.4259	0.5968	0.6678	0.6741	0.1524	0.8102	0.9270	0.4898	0.8337
mAP	87.1	86.9	86.4	81.6	41.6	81.5	83.9	74.8	87.2

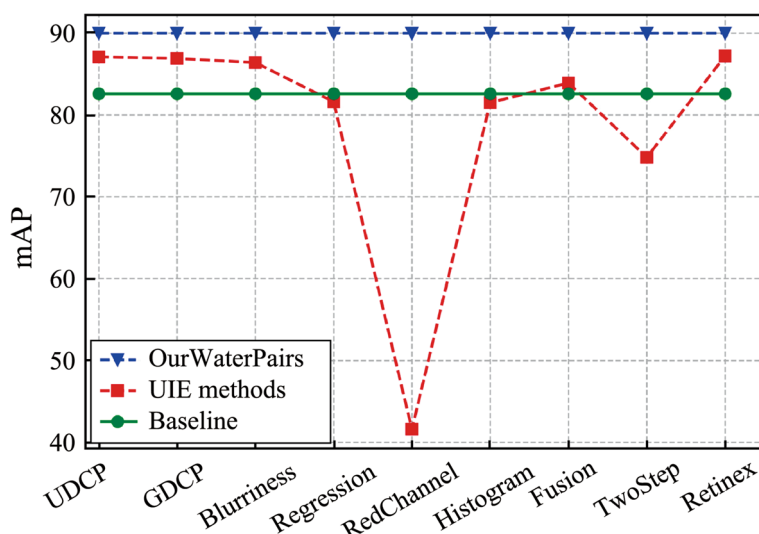


Fig. 6 Performance of detection networks trained using images enhanced by using different UIE methods. 'Baseline' means that the detection network is trained and tested using raw underwater images without enhancement. 'OurWaterPairs' means that the detection network is trained and tested using high-quality reference images in the WaterPairs dataset

For example, RedChannel and TwoStep greatly decrease the detection performance because these two methods degrade the visual quality of the raw underwater images, as depicted in Fig. 5.

4.2 Evaluation of the UIS algorithms on the WaterPairs dataset

We also evaluate different synthesis models on the WaterPairs dataset, namely, Physical (Li et al. 2020b), CycleGAN (Li et al. 2017b), and WaterGAN (Wang et al. 2015). Physical is a physical-based UIS method, whereas the other two are deep learning-based models. Physical (Li et al. 2020b) applies the physical underwater image formation model and 10 groups of predefined parameters to synthesize 10 Jerlov-type underwater images from RGB-D in-air images. The synthetic dataset contains 10 types of underwater images with various color distortions and haze effects. Because the WaterPairs dataset does not provide a depth map, we apply the depth estimation method used in Fu et al. (2014) to obtain the depth maps for all the reference images.

Table 4 lists the average MSE, PSNR, SSIM, and PCQI scores of the three UIS methods, from which we observe that CycleGAN performs better than the other two methods in terms of four full-reference metrics. Thus, we select CycleGAN as the baseline UIS algorithm on the WaterPairs dataset.

Table 4 Full-reference image quality evaluations of different UIS algorithms on the WaterPairs dataset

Methods	MSE	PSNR	SSIM	PCQI
Physical	0.4450	21.7190	0.6849	0.4782
WaterGAN	0.3112	23.2620	0.7378	0.5916
CycleGAN	0.3007	23.3956	0.7549	0.6384

5 Conclusions

In this work, we propose a novel reference image generation method that combines subjective perception and objective assessment. With the generated high-quality reference images for underwater images, we are able to construct a large-scale underwater dataset, named WaterPairs, which offers underwater images, corresponding high-quality reference images, and object-level bounding box annotations. This dataset provides a public platform for researchers to comprehensively compare different UIE and UIS algorithms and to explore the effect of UIE algorithms on underwater object detection tasks.

Acknowledgements

Thanks to the volunteers in the Biomedical Image Processing Lab <https://bipl-uol.github.io/> of the University of Leicester and the Sense Robotics Lab of Imperial College London <https://www.imperial.ac.uk/sense-robotics-lab/> for helping to generate the high-quality reference images.

Additional information

Edited by: Wenwen Chen

Authors' contributions

Long Chen conceived the idea of this study and designed the hybrid reference image generation algorithm. He also organized the volunteers in the Biomedical Image Processing Lab of the University of Leicester to produce high-quality reference images. Xirui Dong and Yunzhou Xie carried out experiments on different underwater image synthesis frameworks. Sen Wang and Xirui Dong organized the volunteers in the Sense Robotics Lab of Imperial College London to produce high-quality reference images, and they also co-wrote the paper.

Availability of data and materials

The WaterPairs dataset will be released at <https://github.com/lanDragon/WaterPairs> once this paper is accepted.

Code availability

Not applicable.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests. Long Chen and Sen Wang are Editorial Board Members, but they were not involved in the journal's review of, or decision related to, this manuscript.

Received: 15 November 2023 Revised: 26 December 2023 Accepted: 3 January 2024

Published online: 18 February 2024

References

- Ancuti C, Ancuti CO, Haber T, Bekaert P (2012) Enhancing underwater images and videos by fusion. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, pp 81–88
- Bazeille S, Quidu I, Jaulin L, Malkasse JP (2006) Automatic underwater image pre-processing. In: CMM'06, Brest, pp 1–8
- Bogue R (2015) Underwater robots: a review of technologies and applications. *Ind Robot* 42(3):186–191
- Clark CM, Forney C, Manii E, Shinzaki D, Gage C, Farris M et al (2013) Tracking and following a tagged leopard shark with an autonomous underwater vehicle. *J Field Robot* 30(3):309–322
- Drews P, Nascimento E, Moraes F, Botelho S, Campos M (2013) Transmission estimation in underwater single images. In: 2013 IEEE International Conference on Computer Vision Workshops, Sydney, pp 825–830
- Fabrizi C, Islam MJ, Sattar J (2018) Enhancing underwater imagery using generative adversarial networks. In: 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, pp 7159–7165
- Foresti GL, Gentili S (2000) A vision based system for object detection in underwater images. *Int J Pattern Recognit Artif Intell* 14(2):167–188
- Fu XY, Fan ZW, Ling M, Huang Y, Ding XH (2017) Two-step approach for single underwater image enhancement. In: 2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Xiamen, pp 789–794

- Fu X, Zhuang P, Huang Y, Liao YH, Zhang XP, Ding XH (2014) A retinex-based enhancing approach for single underwater image. In: 2014 IEEE International Conference on Image Processing (ICIP), Paris, pp 4572–4576
- Galdran A, Pardo D, Picón A, Alvarez-Gila A (2015) Automatic Red-Channel underwater image restoration. *J Vis Commun Image Represent* 26:132–145
- He KM, Sun J, Tang XO (2010) Single image haze removal using dark channel prior. *IEEE Trans Pattern Anal Mach Intell* 33(12):2341–2353
- Jian MW, Qi Q, Dong JY, Yin YL, Zhang WY, Lam KM (2017) The OUC-vision large-scale underwater image database. In: 2017 IEEE International Conference on Multimedia and Expo (ICME), Hong Kong, pp 1297–1302
- Li CY, Anwar S, Porikli F (2020a) Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognit* 98:107038
- Li CY, Guo CL, Ren WQ, Cong RM, Hou JH, Kwong S et al (2019) An underwater image enhancement benchmark dataset and beyond. *IEEE Trans Image Proc* 29:4376–4389
- Li CY, Guo JC, Cong RM, Pang YW, Wang B (2016) Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. *IEEE Trans Image Proc* 25(12):5664–5677
- Li CY, Guo JC, Guo CL, Cong RM, Gong JC (2017a) A hybrid method for underwater image correction. *Pattern Recognit Lett* 94:62–67
- Li J, Skinner KA, Eustice RM, Johnson-Roberson M (2017b) WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robot Autom Lett* 3(1):387–394
- Li TY, Rong SH, Cao XT, Liu YB, Chen L, He B (2020b) Underwater image enhancement framework and its application on an autonomous underwater vehicle platform. *Opt Eng* 59(8):083102
- Li TY, Rong SH, Chen L, Zhou HY, He B (2022a) Underwater motion deblurring based on cascaded attention mechanism. *IEEE J Ocean Eng*. <https://doi.org/10.1109/JOE.2022.3192047>
- Li TY, Rong SH, He B, Chen L (2022b) Underwater image deblurring framework using a generative adversarial network. In: OCEANS 2022, Chennai, pp 1–4
- Li TY, Rong SH, He B, Chen L (2022c) Underwater image enhancement using adaptive color restoration and dehazing. *Opt Express* 30(4):6216–6235
- Liu RS, Fan X, Zhu M, Hou MJ, Luo ZX (2020) Real-world underwater enhancement: challenges, benchmarks, and solutions under natural light. *IEEE Trans Circuits Syst Video Technol* 30(12):4861–4875
- Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY et al (2016) SSD: single shot multibox detector. In: 14th European Conference on Computer Vision (ECCV), Amsterdam, pp 21–37
- Mohammadi P, Ebrahimi-Moghadam A, Shirani S (2014) Subjective and objective quality assessment of image: a survey. Preprint at [arXiv:1406.7799](https://arxiv.org/abs/1406.7799)
- Panetta K, Gao C, Agaian S (2015) Human-visual-system-inspired underwater image quality measures. *IEEE J Ocean Eng* 41(3):541–551
- Peng YT, Cao KM, Cosman PC (2018) Generalization of the dark channel prior for single image restoration. *IEEE Trans Image Proc* 27(6):2856–2868
- Peng YT, Cosman PC (2017) Underwater image restoration based on image blurriness and light absorption. *IEEE Trans Image Proc* 26(4):1579–1594
- Pronin E (2007) Perception and misperception of bias in human judgment. *Trends Cogn Sci* 11(1):37–43
- Schettini R, Corchs S (2010) Underwater image processing: state of the art of restoration and image enhancement methods. *EURASIP J Adv Signal Proc* 2010:746052
- Seshadrinathan K, Soundararajan R, Bovik AC, Cormack LK (2010) Study of subjective and objective quality assessment of video. *IEEE Trans Image Proc* 19(6):1427–1441
- Singh G, Jaggi N, Vasamsetti S, Sardana HK, Kumar S, Mittal N (2015) Underwater image/video enhancement using wavelet based color correction (WBCC) method. In: 2015 IEEE Underwater Technology (UT), Chennai, pp 1–5
- Wang SQ, Ma KD, Yeganeh H, Wang Z, Lin WS (2015) A patch-structure representation method for quality assessment of contrast changed images. *IEEE Signal Proc Lett* 22(12):2387–2390
- Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Proc* 13(4):600–612
- Xu J, Wang M, Qiao L (2015) Dynamical sliding mode control for the trajectory tracking of underactuated unmanned underwater vehicles. *Ocean Eng* 105:54–63
- Yang M, Sowmya A (2015) An underwater color image quality evaluation metric. *IEEE Trans Image Proc* 24(12):6062–6071
- Zhang WD, Zhou L, Zhuang PX, Li GH, Pan XP, Zhao WY et al (2023) Underwater image enhancement via weighted wavelet visual perception fusion. *IEEE Trans Circuits Syst Video Technol*. <https://doi.org/10.1109/TCSVT.2023.3299314>
- Zhang WD, Zhuang PX, Sun HH, Li GH, Kwong S, Li CY (2022) Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. *IEEE Trans Image Proc* 31:3997–4010
- Zhu JY, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV), Venice, pp 2242–2251
- Zhuang PX, Wu JM, Porikli F, Li CY (2022) Underwater image enhancement with hyper-laplacian reflectance priors. *IEEE Trans Image Proc* 31:5442–5455

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.