Jian et al.
Intelligent Marine Technology and Systems (2024) 2:9
https://doi.org/10.1007/s44295-024-00023-6

Intelligent Marine Technology and Systems

REVIEW

Open Access

Underwater object detection and datasets: a survey

Muwei Jian[1,2*†], Nan Yang[2†], Chen Tao[2], Huixiang Zhi[2] and Hanjiang Luo[3]

Abstract

The rapidly growing exploitation and utilization of marine resources by humans has sparked considerable interest in underwater object detection tasks. Targets captured in underwater environments differ significantly from those captured in general images owing to various factors, such as water turbidity, complex background conditions, and lighting variations. These adverse factors pose a host of challenges, such as high intensity noise, texture distortion, uneven illumination, low contrast, and limited visibility in underwater images. To address the specific difficulties encountered in underwater environments, numerous underwater object detection methods have been developed in recent years in response to these challenges. Furthermore, there has been a significant effort in constructing diverse and comprehensive underwater datasets to facilitate the development and evaluation of these methods. This paper outlines 14 traditional methods used in underwater object detection based on three aspects that rely on handmade features. Thirty-four more advanced technologies based on deep learning were presented from eight aspects. Moreover, this paper conducts a comprehensive study of seven representative datasets used in underwater object detection missions. Subsequently, the challenges encountered in current underwater object detection tasks were analyzed from five directions. Based on the findings, potential research directions are expected to promote further progress in this field and beyond.

Keywords Underwater images, Object detection, Underwater dataset, Marine internet of things

1 Introduction

The twenty-first century has been widely recognized as the 'century of the ocean', representing a pivotal era in which humanity will extensively exploit the vast resources that can be derived from the ocean. According to statistical data, China's ocean area comprises approximately 14% of the world's total ocean area, while the global ocean area accounts for around 71% of the Earth's total surface area. These figures underscore the significant presence of the ocean on our planet. As a vast reservoir, the ocean harbors abundant natural resources, making it a subject of great interest to humanity. The rapid advancement of science and technology, along with the urgent resource demands of our society, has driven the utilization and exploitation of marine resources, thus amplifying the importance of underwater object detection.

Over the past decade, the field of underwater object detection has witnessed the development of numerous distinctive methods, leading to remarkable achievements. The innovative achievements of these methods in the field of underwater object detection continue to drive the development of underwater object detection. These methods have not only made significant contributions but also continue to drive the advancement of underwater object detection tasks. More importantly, these achievements provide essential technical support

[†]Muwei Jian and Nan Yang contributed equally to this work.

*Correspondence:
Muwei Jian
jianmuweihk@163.com
[1] School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250000, China
[2] School of Information Science and Engineering, Linyi University, Linyi 276000, China
[3] College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

for scientific research, marine resource development, and various other fields.

However, the intricate nature of the marine environment presents considerable challenges in the detection and analysis of underwater objects. Underwater images are frequently affected by several factors, such as water flow, lighting variations, limited visibility, and substantial changes in pose and spatial position information, resulting in high noise and low contrast. Furthermore, the construction of underwater datasets for underwater object detection has proven to be a challenging task-one that has affected the progress of underwater object detection work and severely limited the practical applications of underwater object detection and recognition.

The main structure of this paper is as follows. Section 2 introduces the different methods for underwater object detection proposed in recent years, including methods based on traditional artificial features and those based on deep learning. Section 3 summarizes and introduces representative datasets used for underwater object detection tasks. Section 4 provides a brief analysis of the challenges faced in the current development of underwater object detection and the prospects for future research. Finally, a summary of this study is provided in Section 5.

2 Overview of underwater object detection methods

In the field of underwater computer vision and image processing, the primary objective of object detection is to enable computers to comprehend underwater scenes. This capability is crucial not only for understanding the underwater environment but also for underlining its significant role in the exploration and use of resources. In recent years, research progress in underwater object detection has experienced a notable transformation from relying on traditional manual features to embracing deep learning techniques. Initially, traditional manual features were predominantly used in the early stages of research. However, these approaches faced significant limitations when applied in practical underwater environments. Furthermore, most detection algorithms for underwater object detection rely on manually designed feature extraction, which is a process requiring professional expertise and complex algorithm debugging. However, this approach has limited universality and detection accuracy, which hinders its development in related fields.

Recently, the development of artificial intelligence (AI) technology has attracted the attention of scholars from universities and research institutes dedicated to underwater object detection research. Numerous methods have been developed in this field, which can be broadly categorized into two main categories: methods based on traditional manual features and those based on deep learning. For instance, Duan et al. (2015) conducted a comprehensive analysis of research progress on fish size, shape, color, and other aspects from the perspective of computer vision. They covered various stages, such as image acquisition, contour extraction, feature calibration, and calculation, and discussed the application of computer vision in diagnosing, detecting, and classifying aquatic animal diseases. Peng et al. (2021) examined deep learning methods for underwater image preprocessing and discussed their advantages and disadvantages. They also discussed enhancements made to deep learning methods and practical application challenges. Wu et al. (2019) studied the impact of lighting conditions on underwater image characteristics. They employed different image processing algorithms to extract invariant features from underwater images and conducted underwater red ball experiments to verify the feasibility of underwater object detection. In addition, Yu (2020) conducted a comprehensive review of studies that covered data collection techniques for aquatic animals, such as fish, shrimp, and sea cucumbers; comparison of underwater object detection datasets; preprocessing methods for underwater image data; different underwater object detection technologies; and the application of deep learning in detection and tracking.

Therefore, based on the abovementioned information, the current section primarily reviews underwater object detection methods based on traditional artificial features and deep learning, incorporating the contributions of various researchers in the field. In addition, when selecting a method for inclusion, we consider its relevance to the topic of this paper, the clarity and accessibility of the method, and its contribution to the field. At the same time, we also consider the diversity of methods to better highlight the topic content of this paper.

2.1 Methods based on traditional artificial features

Traditional underwater object detection methods depend on manually designed feature extraction and classification algorithms. These methods encompass various techniques, such as sonar and optical imaging (Mukherjee et al. 2011; Tucker and Azimi-Sadjadi 2011; Ghafoor and Noh 2019; Gillis 2020; Jian et al. 2021). They also involve extracting and combining traditional artificial features, such as texture, shape, color, and motion of targets, which are then used in conjunction with classical machine learning algorithms to achieve underwater object detection. These methods are summarized in Table 1.

Texture features are valuable indicators of the surface properties of an image. Shi et al. (2019) introduced a method based on grayscale co-occurrence matrix (GLCM) using a support vector machine (SVM) classifier to automatically identify underwater cage boundaries.

Table 1 Summary of underwater object detection methods based on traditional artificial features

Category	Reference
Texture features	Han and Choi 2011; Beijbom et al. 2012; Nagaraja et al. 2015; Fatan et al. 2016; Srividhya and Ramya 2017; Shi et al. 2019
Color and motion features	Gordan et al. 2006; Chen and Chen 2010; Singh et al. 2015; Komari Alaie and Farsi 2018; Susanto et al. 2018
Saliency detection	Wang et al. 2014; Zhu et al. 2016; Jian et al. 2018b

This technique extracts and computes GLCM features from underwater images by using rich texture information for precise detection of underwater cage boundaries. Nagaraja et al. (2015) employed robust local binary pattern descriptors to extract texture features from underwater images. Similarly, Fatan et al. (2016) proposed an underwater cable detection method based on texture information for image edge classification. They used a multilayer perceptron (MLP) neural network (Taud and Mas 2018) and a texture-based SVM to extract image edges. The detected edges were further refined by removing background information using morphological operators followed by Hough transform-based detection. Srividhya and Ramya (2017) proposed a method that combines learning algorithms with texture features for the accurate detection and recognition of underwater objects. In an earlier study, Beijbom et al. (2012) developed a novel algorithm employing multiscale texture and color descriptors, surpassing other methods in verifying underwater coral reef data. Han and Choi (2011) proposed an efficient and accurate method for detecting and tracking texture-less objects in underwater environments. Their proposed method addresses the challenges posed by the absence of distinctive texture features in certain underwater objects.

The abovementioned studies emphasize the significance of texture features in underwater object detection and demonstrate the effectiveness of various texture-based algorithms in different underwater scenarios. Apart from texture, color and motion features play a crucial role in underwater image analysis, and these have been studied in previous works. For example, Chen and Chen (2010) proposed a new color edge detection algorithm in 2010, which used the Kuwahara filter (Bartyzel 2016) to smoothen the original image. They incorporated adaptive thresholding and edge sparsity algorithms to enhance detection efficiency and performance. Gordan et al. (2006) introduced an architecture specifically designed for underwater scene analysis using SVM classifiers. Their method detects and recognizes underwater objects by extracting color pixel features and using threshold comparison techniques. Singh et al. (2015) presented a method for the automatic real-time detection and tracking of moving objects in video frames using color and motion features. Susanto et al. (2018) developed a color-based detection system that distinguishes and detects objects based on selected colors. Similarly, Komari Alaie and Farsi (2018) designed a novel method for detecting underwater sonar targets using adaptive thresholds. To improve target/object detection, their approach combines detection points with techniques such as Bayesian classification, maximum likelihood estimation, and minimum mean square adaptive filtering.

Saliency object detection technology (Jian et al. 2014, 2018a) has also found extensive application in underwater image object detection. For example, Jian et al. (2018b) proposed a new framework for detecting salient objects in underwater images using the quaternion distance Weber descriptor, mode clarity, and local contrast. Their proposed framework combines the quaternion system and principal component analysis to achieve superior detection performance. Zhu et al. (2016) introduced an automatic detection method using saliency-based region merging. To achieve more accurate automatic detection of underwater objects, they incorporated prominent object detection, background prior methods, and an improved interactive image segmentation method based on region merging. Similarly, Wang et al. (2014) proposed a region saliency calculation model for underwater object detection that combines saliency regions and prior knowledge. This model adopts a target/object detection method based on regional saliency and underwater optical priors, thereby reducing algorithm complexity and enhancing detection accuracy at the same time.

While traditional underwater object detection methods rely on manual feature extraction, which is time-consuming and lacks robustness, the emergence of deep learning and convolutional neural networks has ushered in a new phase in underwater object detection algorithms.

2.2 Methods based on deep learning

In recent years, the field of underwater object detection has witnessed significant advancements thanks to the development of deep learning techniques. Deep learning methods, such as convolutional neural networks (CNNs), have gained widespread popularity because of their ability to automatically learn and extract features from underwater images, thus leading to improved accuracy in detection and recognition tasks. Compared with traditional methods, deep learning approaches exhibit enhanced robustness and performance. Table 2 summarizes the methods based on deep learning that have been proposed in recent years.

Current research in underwater object detection revolves around deep learning methods and strives to enhance the universality and accuracy of established algorithms. The introduction of R-CNN (Girshick et al. 2014) marked a pivotal moment in the rapid progress of deep learning in object detection and recognition. At present, numerous scholars have increasingly embraced deep learning and applied it to underwater object detection, resulting in notable and innovative research outcomes.

The existing object detection algorithms can be categorized into two main types: two-stage and single-stage algorithms. On the one hand, two-stage algorithms, such as the R-CNN series algorithms, including Fast R-CNN (Girshick 2015) and Faster R-CNN (Ren et al. 2015), involve generating region proposals and then performing classification and regression tasks on these proposals. Although these algorithms have demonstrated improved detection performance, they tend to have low processing efficiency. On the other hand, single-stage algorithms, such as single-shot multiBox detector (Liu et al. 2016) and the you only look once (YOLO) series of algorithms (Bochkovskiy et al. 2020) focus on achieving high detection speed while maintaining good detection performance. These algorithms use direct regression methods to predict the category and position of targets in a single pass. For instance, Hu et al. (2021) modified the network connections and replaced the feature mapping responsible for large features in YOLO-v4 with finer-grained feature maps to address the issue of high ammonia nitrogen levels in aquaculture caused by the nonconsumption of feed particles in water. Their approach eliminated redundant operations and significantly improved detection and recognition accuracy in real breeding environments. Another example is the research by Ge et al. (2022a, b), who proposed a single-level underwater object detection method based on feature anchor frames and feature double enhancement. They designed a composite connected backbone network to leverage the advantages of different backbone networks, thereby improving contextual relevance and multiscale detection capabilities. Furthermore, Lei et al. (2022) made enhancements to the YOLO-v5 algorithm specifically for underwater object detection. To enhance the algorithm's performance in underwater environments, they incorporated the twin transformer as the backbone network and improved the multiscale feature fusion method and confidence loss function.

Due to turbidity, absorption, and scattering in the underwater environment, underwater images often suffer from challenges, including high noise and low contrast (Yuan et al. 2022). Researchers have developed various methods to address these issues and improve the accuracy and performance of underwater object detection. For example, Chen et al. (2017) designed a detection method for underwater object recognition using monocular visual sensors. Their approach focused on enhancing the detection accuracy of underwater scenes by removing background noise. Chen et al. (2018) developed an effective model using adversarial networks for super-resolution generation to enhance the visual impact of underwater images in target/object detection and recognition tasks. Sun et al. (2018) introduced an underwater object detection model based on CNNs. In particular, they were able to discriminate targets in low-contrast underwater images by incorporating a weighted probability decision mechanism.

Target/object state changes and occlusion also significantly impact the target detection process. Lin et al. (2020) proposed a method called RoIMix, which exhibited improved generalization performance, especially for detecting underwater images with overlap, occlusion, and blur. To achieve underwater object detection, Lau and Lai (2021) focused on the selection and

Table 2 Summary of the deep learning methods for underwater object detection

Category	Reference
Single-stage algorithm	Liu et al. 2016; Bochkovskiy et al. 2020; Hu et al. 2021; Ge et al. 2022a, b; Lei et al. 2022
Two-stage algorithm	Girshick et al. 2014; Girshick 2015; Ren et al. 2015
High noise and low contrast	Chen et al. 2017, 2018; Sun et al. 2018
State changes and occlusion	Yang et al. 2019; Lin et al. 2020; Lau and Lai 2021; Zhang et al. 2021
Shadows and uneven illumination	Song et al. 2014; Cao et al. 2016; Li et al. 2016a; Ding et al. 2017; Yu et al. 2019; Fan et al. 2020; Wei et al. 2021; Chen et al. 2023
Weak lighting and low quality	Rashwan et al. 2019; Chen et al. 2020a; Han et al. 2020; Ge et al. 2022a, b; Liu et al. 2022
Low data volume	Zurowietz and Nattkemper 2020; Zeng et al. 2021
Saliency detection	Li et al. 2016b; Mou et al. 2017; Zhou et al. 2019; Chen et al. 2020c

enhancement of the basic network architecture in Faster R-CNN. They performed preprocessing on the obtained images and tested the performance of different network architectures to identify the most suitable one for training object detection in turbid media. Yang et al. (2019) combined a deep short-term memory network (DLSTM) with a deep autoencoder neural network to effectively identify targets at different depths and reduce radiated noise. They used a pretrained DLSTM model and a SoftMax classifier to detect and classify ship-radiated noise. Zhang et al. (2021) developed a lightweight underwater object detection method based on MobileNet v2, YOLO-v4 algorithms, and attention feature fusion. Their proposed method reduces the number of parameters, resulting in a lighter model that significantly improves the speed and accuracy of detection.

During the detection and recognition of underwater objects, the intensity of underwater light decreases with depth, thereby leading to challenges such as shadows and uneven illumination. Thus, to address these issues and improve the accuracy of target/object detection and recognition in underwater environments, researchers have proposed several approaches. For example, Song et al. (2014) used underwater vehicles equipped with visual imaging devices to compensate for targets with varying light intensities. The algorithm reduces the impact of uneven lighting on target/object detection by extracting image and color features from the target image. Li et al. (2016a) introduced an effective defogging model to restore visibility, color, and natural appearance in underwater images. This model improves the quality of underwater images and enhances the detection and recognition accuracy of underwater objects. Ding et al. (2017) designed an underwater image enhancement strategy that combines model-based defogging and adaptive color correction. Their proposed strategy helps reveal more features by enhancing the original underwater image, thus effectively improving the quality of images and increasing the accuracy of target detection and recognition.

Yu et al. (2019) proposed a redesigned framework for underwater generative adversarial network (GAN) image restoration. This framework uses GAN classifiers to learn structural losses and generates more realistic images through simulation via the underwater image generation model. Their proposed approach improves target/object recognition accuracy by reducing the impact of abnormal image contrast. Chen et al. (2023) invented a comprehensive object detection algorithm based on a lightweight transformer that incorporates cross-scale feature fusion and enhanced multiscale feature fusion. This algorithm improves feature fusion, reduces model parameters, and enhances local feature correlation, thus leading to improved detection accuracy.

Wei et al. (2021) proposed an object detection algorithm that integrates attention mechanisms and scale enhancement. Furthermore, to enhance feature extraction capabilities, they added compression and excitation modules after the deep convolutional layer. Combining shallow and deep features with more positional information helped improve the detection performance of small target models. Cao et al. (2016) devised an underwater object recognition and classification framework that combines stacked automatic encoders and Softmax. In particular, this framework learns invariant features and extracts advanced features from the spectral data of underwater objects by employing sparse and stacked autoencoders. Fan et al. (2020) proposed a framework for underwater object detection based on feature enhancement and anchor refinement. This framework incorporates a composite connection backbone to enhance feature representation and introduces a receptive field enhancement module to exploit multiscale contextual features.

In addition, given the low lighting and quality issues in underwater environments, researchers have proposed various methods and architectures to enhance the original underwater images and improve their visual perception and applicability. Han et al. (2020) combined the max-RGB and grayscale methods to improve underwater vision. Then, by training mapping relationships to obtain illumination maps, they further introduced a CNN method to address the issue of weak illumination in underwater images. Chen et al. (2020a) developed a neural network structure called sample weighted hypernetwork (SWIPENet) for the specific purpose of detecting small underwater objects. This architecture aims to overcome image blur and improve the accuracy of target/object detection. Rashwan et al. (2019) introduced a deep architecture called a matrix network for object detection. This architecture incorporates a scaling and aspect ratio sensing mechanism to enhance keypoint-based object detection.

Recently, Ge et al. (2022a, b) proposed a GAN-based underwater image enhancement method to tackle the problem of underwater image degradation. They successfully created an underwater style dataset and made lightweight improvements to the model by combining the multiscale retinex with color restoration and DehazeNet (Cai et al. 2016), resulting in significant improvements in detection accuracy. Liu et al. (2022) redesigned an underwater enhancement method based on object-guided dual adversarial contrastive learning, and their approach achieved both visual friendliness and task-oriented enhancement. They employed comparative prompts during the training phase and embedded a task perception feedback module in the enhancement process to make the restored image more realistic.

In underwater object detection tasks, the limited amount of underwater image data poses a significant challenge. In response, researchers have proposed several approaches to address this problem and improve the detection capability of underwater object detection algorithms. For example, Zeng et al. (2021) proposed an underwater object detection algorithm based on Faster R-CNN and adversarial networks. By incorporating adversarial networks into the standard Faster R-CNN detection network for joint training, they increased the number of training samples and improved the network's detection capability. Zurowietz and Nattkemper (2020) introduced unsupervised knowledge transfer (UnKnoT) as a more effective method for training with limited data. This approach uses a data augmentation technique called 'scale transfer' to reuse existing training data and detect the same object classes in a new image dataset.

Inspired by saliency detection, Chen et al. (2020c) designed an underwater saliency object detection model that considers both two-dimensional (2D) and three-dimensional (3D) depth cues. Their proposed model improves the detection of underwater objects by leveraging saliency detection principles. Meanwhile, to address the issues of low contrast and low-quality images, Li et al. (2016b) proposed a foreground extraction-based underwater image saliency detection framework. This framework focuses on extracting salient foreground regions and enhancing the detection of underwater objects. Mou et al. (2017) used the Harris angle detection operator to locate geometric centers and designed a simple linear iterative clustering method. Their approach achieves the effective detection of underwater objects by highlighting foreground targets while attenuating background areas. Zhou et al. (2019) introduced a composite convolutional neural network based on shared latent sparse features (SLS) and deep belief networks (DBN), thereby overcoming the lack of CNN training data by using texture images and optimizing and interfering with textures using SLS and DBN. Their proposed method enhances the performance of underwater object detection and classification.

3 Datasets for underwater object/target detection

In recent years, underwater object detection has emerged as a prominent area of research with the rapid advancement of AI technology. However, the complexity and extensive demands of underwater environments have presented significant challenges in constructing underwater datasets. To overcome this obstacle and offer more comprehensive data support for advancements in underwater image processing, numerous research teams have successfully developed unique underwater datasets using methods such as underwater robots and simulation labs (Chen et al. 2020b). In this section, we provide a concise overview of noteworthy datasets in the field of underwater object detection. These datasets encompass several underwater imageries, thereby providing researchers with valuable data for algorithm validation and performance evaluation. Table 3 summarizes some representative underwater datasets used for underwater object detection tasks in recent years.

The Brackish dataset was first proposed and made publicly available by Pedersen et al. (2019). This dataset consists of 14518 frames of images, with the original data being video data. The Brackish dataset contains 25613 annotations belonging to six categories: big fish, small fish, crab, jellyfish, shrimp, and starfish. Figure 1 shows the sample frame image for each category in the Brackish dataset.

Saliency is typically generated by 'contrast', usually due to the contrast between an item and its adjacent items. The detection of underwater saliency targets is

Table 3 Representative datasets for underwater object detection tasks

Dataset	Number	Application	Challenge	Year
Brackish	14518	Object/target detection	Poor image quality, too blurry, contains few underwater targets and categories	2019
MUED	8600	Object/target detection Saliency detection	Large dataset is not conducive to model training and validation, and the overall background is too uniform	2019
RUIE	4230	Target detection Target enhancement Target classification	The amount of data is small, and each subset has strong specificity, resulting in poor generalization	2020
UWD	10000	Target detection	Few data categories and uneven sample size	2020
TrashCan	7212	Target detection Target segmentation	The types of underwater targets are complex and difficult to distinguish	2020
UDD	2227	Target detection	Insufficient data volume and uneven distribution of categories, resulting in poor generalization ability due to insufficient samples in certain categories	2021
DUO	7782	Target detection	There are few underwater target categories, and the sample size of each category is uneven	2021

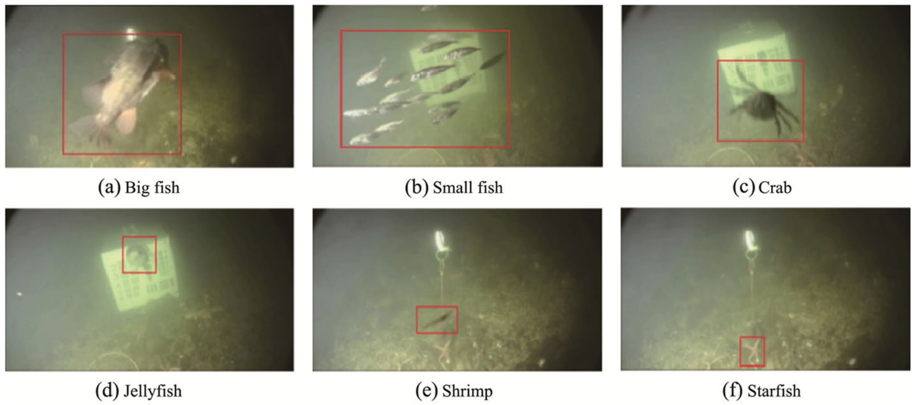

Fig. 1 Diagram of an example frame from the Brackish dataset category (Pedersen et al. 2019)

often difficult due to the diversity of the underwater environment and the lack of underwater datasets. To address this challenge, the marine underwater environment database (MUED; Jian et al. 2019) provides 8600 underwater images with 430 different categories of salient objects. These images have complex backgrounds and multiple prominent objects and show complex changes in posture, spatial position, lighting, and other aspects. Figure 2 shows six examples, including—from left to right—posture changes, spatial position changes, lighting changes, water turbidity changes, background changes, and target/object number changes.

Unlike most datasets, the real-time underwater image enhancement (RUIE; Liu et al. 2020) dataset consists of three subsets: underwater image quality set, underwater color cast set, and underwater high-level task-driven set (UHTS). They each target three challenging underwater tasks: visibility reduction, color deviation, and higher-level detection/classification. Of these, UHTS is most commonly used for underwater object detection, and this subset contains 300 underwater images. Figure 3 shows an example of underwater images from three subsets of the RUIE dataset.

To better validate the generalization of the underwater object detection framework, Fan et al. (2020) collected and integrated relevant underwater images from the Internet, after which they constructed an underwater dataset (UWD) for object detection through manual annotation. UWD contains 10000 training and test images classified into four categories: sea cucumber, octopus, scallop, and starfish. Notably, despite the large number of images in the UWD, this dataset does not specifically divide the number of images in the training and testing sets. Figure 4 is an example image of UWD.

Currently, most datasets used for underwater object detection tasks focus on marine organisms. The Trash-Can (Hong et al. 2020) dataset consists mainly of

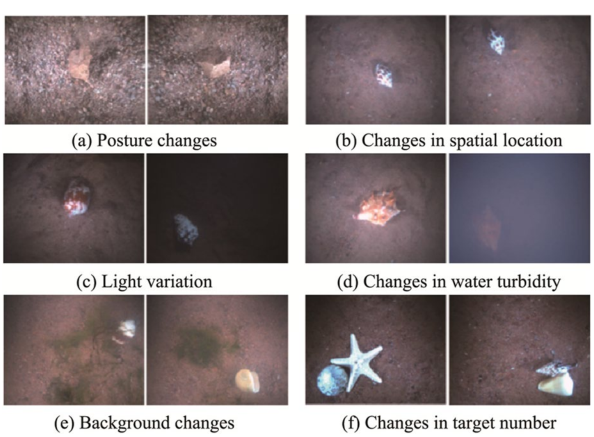

Fig. 2 Six types of sample images from the MUED dataset (Jian et al. 2019)

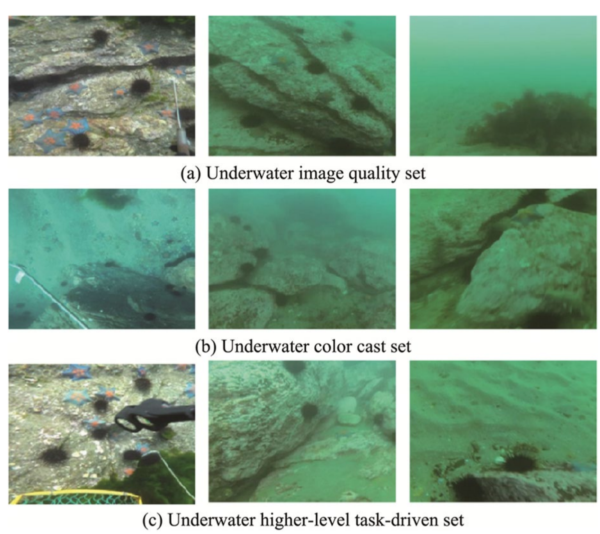

Fig. 3 Image samples of three subsets in the RUIE dataset (Liu et al. 2020)

underwater garbage, with annotations in the form of instance-segmented annotations containing bitmaps with masks. TrashCan consists of 7212 annotated images, including images of underwater debris, underwater robots, and various underwater animals and plants. This dataset uses bounding boxes and segmentation labels for annotation, which can be used for underwater object detection and segmentation tasks. Figure 5 shows the original image of the TrashCan dataset.

In deep water environments, various factors, such as water current strength, water turbidity levels, and benthic organism activity, can significantly impact the clarity of underwater images and consequently affect their quality. Among these factors, water turbidity is a crucial element. In particular, existing underwater image datasets often suffer from the influence of water turbidity, which leads to subpar image quality within the datasets. Liu et al. (2021a) tackled this challenge by collecting and constructing a high-resolution underwater detection dataset (UDD) for open-sea farm objects in the seafloor environment. The UDD comprises a total of 2227 underwater images, with 1827 images dedicated to training and 400 images for testing. The dataset encompasses three distinct categories: sea cucumbers, sea urchins, and scallops, encompassing 15022 categorized objects, including 1148 sea cucumbers, 13592 sea urchins, and 282 scallops. As a complement to the UDD, the research team also constructed an augmented underwater farm object detection dataset (AUDD), a large-scale dataset consisting of 18661 images based on the UDD. Figure 6 presents examples of raw images from the UDD.

The detection of underwater objects (DUO) dataset (Liu et al. 2021b) underwent a reorganization that involved collecting and reannotating various existing underwater datasets, including URPC2017, URPC2018, URPC2019, and other datasets previously published in underwater robot competitions. After eliminating excessively similar images, the resulting dataset consists of 7782 underwater images, comprising 6671 images for training and 1111 images for testing. The DUO dataset specifically focuses on four categories of marine organisms: sea cucumbers, thorns, scallops, and starfish. The annotations have been refined to improve the accuracy of underwater object detection. A selection of original images from the DUO dataset's training and testing sets can be found in Fig. 7. As shown in the figure, (a) represents an original image from the training set, and (b) represents an original image from the testing set.

Fig. 4 UWD sample images (Fan et al. 2020)

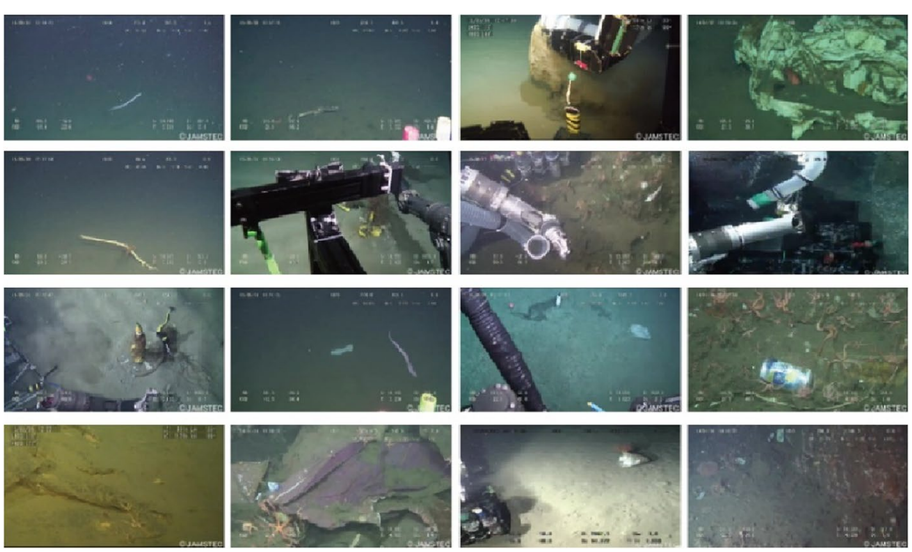

Fig. 5 Original images from the TrashCan dataset (Hong et al. 2020)

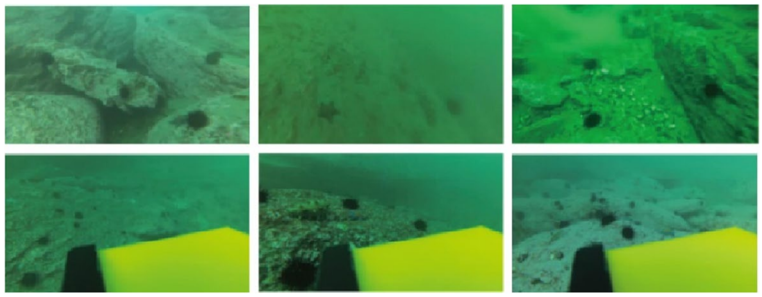

Fig. 6 Examples of raw images from UDD (Liu et al. 2021a)

4 Challenges and future prospects

In this study, we provide a comprehensive overview of recent research advancements in the field of underwater object detection. Rapid progress in AI technology has facilitated the emergence of numerous effective methods for underwater object detection, leading to significant achievements in this field. Undoubtedly, underwater object detection remains a highly active area of research that has attracted the attention of many scholars. However, in recent years, this field of study has continued to confront substantial challenges. In this section, we briefly analyze the existing challenges and outline potential research directions for future research. Our aim is to draw the attention of relevant researchers and foster the continued growth and advancement of underwater object detection. Overall, we summarize the existing challenges as follows:

First, although numerous models for underwater object detection have been proposed, traditional artificial feature-based methods and deep learning-based methods often concentrate on a single perspective. In the future, it will be crucial to emphasize the diverse characteristics exhibited by underwater objects. In particular, researchers can achieve a more holistic understanding of underwater scenes by incorporating different perspectives, thus leading to improved detection performance across diverse underwater environments.

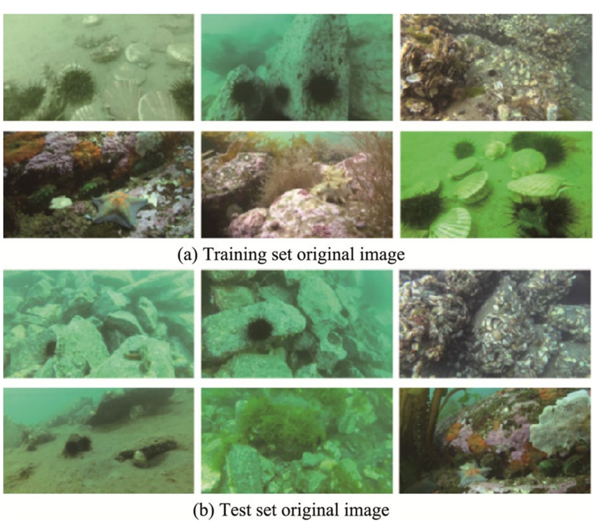

Fig. 7 Original images from the DUO training and testing sets (Liu et al. 2021b)

Second, the detection of small underwater objects poses a significant challenge for deep learning-based models because these targets are often characterized by small size and high levels of camouflaging properties. Existing deep learning models generally exhibit limited robustness in accurately detecting small targets. In the future, there should be a heightened focus on intensifying research efforts dedicated to small underwater object detection.

Third, in underwater environments, the degree of turbidity and refraction in the water are still key factors affecting the quality of underwater images. To reduce their adverse effects and improve the quality of underwater images, researchers should maximize the development and progress of related technologies and use as much advanced equipment and technologies as possible in the future.

Fourth, owing to the complexity of underwater environments, underwater images often suffer from problems such as low contrast, texture loss, and color distortion, making underwater recognition tasks more difficult. Therefore, to minimize the impact of background information and improve the accuracy of underwater object detection, the issue of similarities between the foreground and background of underwater images should also be considered in future works.

Fifth, a major bottleneck in current underwater object detection research is data. At present, existing underwater datasets are not sufficient to meet research needs. While scientists have begun to build their own datasets to better validate the effectiveness of underwater object detection methods, these datasets tend to focus on a particular research direction, have poor generalizability, and have significant limitations. Thus, in the future, it will be necessary to develop large underwater datasets with greater diversity and complexity to support research in underwater object detection.

5 Conclusions

In this paper, we begin with a comprehensive review of the recent research and methodologies of underwater object detection tasks. We highlight the strengths and limitations of each approach and provide insights into their respective contributions to the field. Next, we summarize and present representative datasets that have been used for underwater object detection in recent years. Moreover, we delve into the current challenges faced in underwater object detection. To conclude this work, we outline future research directions in the field of underwater object detection. These directions address challenges and promote advancements in the field. Overall, this paper provides a comprehensive overview of recent research achievements, datasets, challenges, and future directions in underwater object detection.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) (Grant No. 61976123), the Taishan Young Scholars Program of Shandong Province, and the Key Development Program for Basic Research of Shandong Province (Grant No. ZR2020ZD44).

Additional information

Edited by: Wenwen Chen.

Authors' contributions

All authors have contributed to the conceptualization and design of the research. Muwei Jian conceived the idea of this review paper on underwater object detection and datasets and made critical modifications. Nan Yang and Chen Tao drafted the manuscript. Huixiang Zhi conducted the literature search and data analysis. Hanjiang Luo made critical revisions and corrections to the manuscript.

Availability of data and materials

The data and references presented in this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Informed consent for publication was obtained from all participants.

Competing interests

The authors declare that they have no competing interests.

Received: 19 October 2023 Revised: 4 January 2024 Accepted: 8 February 2024
Published online: 04 March 2024

References

Bartyzel K (2016) Adaptive Kuwahara filter. Signal Image Video Proc 10:663–670. https://doi.org/10.1007/s11760-015-0791-3

Beijbom O, Edmunds PJ, Kline DI, Mitchell BG, Kriegman D (2012) Automated annotation of coral reef survey images. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, pp 1170–1177. https://doi.org/10.1109/CVPR.2012.6247798

Bochkovskiy A, Wang CY, Liao HYM (2020) YOLOv4: optimal speed and accuracy of object detection. Preprint at arXiv: 2004.10934

Cai BL, Xu XM, Jia K, Qing CM, Tao DC (2016) DehazeNet: an end-to-end system for single image haze removal. IEEE Trans Image Proc 25(11):5187–5198. https://doi.org/10.1109/TIP.2016.2598681

Cao X, Zhang XM, Yu Y, Niu LT (2016) Deep learning-based recognition of underwater target. In: 2016 IEEE International Conference on Digital Signal Processing, Beijing, pp 89–93. https://doi.org/10.1109/ICDSP.2016.7868522

Chen L, Liu ZH, Tong L, Jiang ZH, Wang SK, Dong JY et al (2020a) Underwater object detection using Invert Multi-Class Adaboost with deep learning. In: 2020 International Joint Conference on Neural Networks, Glasgow, pp 1–8. https://doi.org/10.1109/IJCNN48605.2020.9207506

Chen L, Tong L, Zhou FX, Jiang ZH, Li ZY, Lv JL et al (2020b) A benchmark dataset for both underwater image enhancement and underwater object detection. Preprint at arXiv:2006.15789

Chen L, Yang YY, Wang ZH, Zhang J, Zhou SW, Wu LH (2023) Underwater target detection lightweight algorithm based on multi-scale feature fusion. J Mar Sci Eng 11(2):320. https://doi.org/10.3390/jmse11020320

Chen X, Chen HJ (2010) A novel color edge detection algorithm in RGB color space. In: IEEE 10th International Conference on Signal Processing Proceedings, Beijing, pp 793–796. https://doi.org/10.1109/ICOSP.2010.5655926

Chen Z, Gao HM, Zhang Z, Zhou HL, Wang X, Tian Y (2020c) Underwater salient object detection by combining 2D and 3D visual features. Neurocomputing 391:249–259. https://doi.org/10.1016/j.neucom.2018.10.089

Chen Z, Zhang Z, Dai FZ, Bu Y, Wang HB (2017) Monocular vision-based underwater object detection. Sensors 17(8):1784. https://doi.org/10.3390/s17081784

Chen ZY, Zhao TT, Cheng N, Sun XD, Fu XP (2018) Towards underwater object recognition based on supervised learning. In: 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans, Kobe, pp 1–4. https://doi.org/10.1109/OCEANSKOBE.2018.8559050

Ding XY, Wang YF, Zhang J, Fu XP (2017) Underwater image dehaze using scene depth estimation with adaptive color correction. In: OCEANS 2017-Aberdeen, Aberdeen, pp 1–5. https://doi.org/10.1109/OCEANSE.2017.8084665

Duan YE, Li DL, Li ZB, Fu ZT (2015) Review on visual attributes measurement research of aquatic animals based on computer vision. Trans Chin Soc Agric Eng 31(15):1–11. https://doi.org/10.11975/j.issn.1002-6819.2015.15.001 (in Chinese with English abstract)

Fan BJ, Chen W, Cong Y, Tian JD (2020) Dual refinement underwater object detection network. In: 16th European Conference on Computer Vision, Glasgow, pp 275–291. https://doi.org/10.1007/978-3-030-58565-5_17

Fatan M, Daliri MR, Shahri AM (2016) Underwater cable detection in the images using edge classification based on texture information. Measurement 91:309–317. https://doi.org/10.1016/j.measurement.2016.05.030

Ge HL, Dai YW, Zhu ZY, Liu RB (2022a) A deep learning model applied to optical image target detection and recognition for the identification of underwater biostructures. Machines 10(9):809. https://doi.org/10.3390/machines10090809

Ge HL, Dai YW, Zhu ZY, Zang X (2022b) Single-stage underwater target detection based on feature anchor frame double optimization network. Sensors 22(20):7875. https://doi.org/10.3390/s22207875

Ghafoor H, Noh Y (2019) An overview of next-generation underwater target detection and tracking: an integrated underwater architecture. IEEE Access 7:98841–98853. https://doi.org/10.1109/ACCESS.2019.2929932

Gillis DB (2020) An underwater target detection framework for hyperspectral imagery. IEEE J Sel Top Appl Earth Observ Remote Sens 13:1798–1810. https://doi.org/10.1109/JSTARS.2020.2969013

Girshick R (2015) Fast R-CNN. In: 2015 IEEE International Conference on Computer Vision, Santiago, pp 1440–1448. https://doi.org/10.1109/ICCV.2015.169

Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, pp 580–587. https://doi.org/10.1109/CVPR.2014.81

Gordan M, Dancea O, Stoian I, Georgakis A, Tsatos O (2006) A new SVM-based architecture for object recognition in color underwater images with classification refinement by shape descriptors. In: 2006 IEEE International Conference on Automation, Quality and Testing, Robotics, Cluj-Napoca, pp 327–332. https://doi.org/10.1109/AQTR.2006.254654

Han FL, Yao JZ, Zhu HT, Wang CH (2020) Underwater image processing and object detection based on deep CNN method. J Sens 2020:6707328. https://doi.org/10.1155/2020/6707328

Han KM, Choi HT (2011) Shape context based object recognition and tracking in structured underwater environment. In: 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, pp 617–620. https://doi.org/10.1109/IGARSS.2011.6049204

Hong J, Fulton M, Sattar J (2020) Trashcan: a semantically-segmented dataset towards visual detection of marine debris. Preprint at arXiv:2007.08097

Hu XL, Liu Y, Zhao ZX, Liu JT, Yang XT, Sun CH et al (2021) Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO-V4 network. Comput Electron Agric 185:106135. https://doi.org/10.1016/j.compag.2021.106135

Jian MW, Lam KM, Dong JY, Shen LL (2014) Visual-patch-attention-aware saliency detection. IEEE Trans Cybern 45(8):1575–1586. https://doi.org/10.1109/TCYB.2014.2356200

Jian MW, Liu XY, Luo HJ, Lu XW, Yu H, Dong JY (2021) Underwater image processing and analysis: a review. Signal Proc: Image Commun 91:116088. https://doi.org/10.1016/j.image.2020.116088

Jian MW, Qi Q, Dong JY, Yin YL, Lam KM (2018a) Integrating QDWD with pattern distinctness and local contrast for underwater saliency detection. J vis Commun Image Represent 53:31–41. https://doi.org/10.1016/j.jvcir.2018.03.008

Jian MW, Qi Q, Yu H, Dong JY, Cui CR, Nie XS et al (2019) The extended marine underwater environment database and baseline evaluations. Appl Soft Comput 80:425–437. https://doi.org/10.1016/j.asoc.2019.04.025

Jian MW, Zhang WY, Yu H, Cui CR, Nie XS, Zhang HX et al (2018b) Saliency detection based on directional patches extraction and principal local color contrast. J vis Commun Image Represent 57:1–11. https://doi.org/10.1016/j.jvcir.2018.10.008

Komari Alaie H, Farsi H (2018) Passive sonar target detection using statistical classifier and adaptive threshold. Appl Sci 8(1):61. https://doi.org/10.3390/app8010061

Lau PY, Lai SC (2021) Localizing fish in highly turbid underwater images. In: International Workshop on Advanced Imaging Technology (IWAIT), pp 294–299. https://doi.org/10.1117/12.2590995

Lei F, Tang FF, Li SH (2022) Underwater target detection algorithm based on improved YOLOv5. J Mar Sci Eng 10(3):310. https://doi.org/10.3390/jmse10030310

Li CY, Guo JC, Cong RM, Pang YW, Wang B (2016a) Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. IEEE Trans Image Proc 25(12):5664–5677. https://doi.org/10.1109/TIP.2016.2612882

Li X, Hao J, Shang M, Yang Z (2016b) Saliency segmentation and foreground extraction of underwater image based on localization. In: OCEANS 2016-Shanghai, Shanghai, pp 1–4. https://doi.org/10.1109/OCEANSAP.2016.7485498

Lin WH, Zhong JX, Liu S, Li T, Li G (2020) ROIMIX: proposal-fusion among multiple images for underwater object detection. In: 2020 IEEE International Conference on Acoustics, Speech and Signal Processing, Barcelona, pp 2588–2592. https://doi.org/10.1109/ICASSP40776.2020.9053829

Liu CW, Li HJ, Wang SC, Zhu M, Wang D, Fan X et al (2021a) A dataset and benchmark of underwater object detection for robot picking. In: 2021 IEEE International Conference on Multimedia & Expo Workshops, Shenzhen, pp 1–6. https://doi.org/10.1109/ICMEW53276.2021.9455997

Liu CW, Wang ZH, Wang SJ, Tang T, Tao YL, Yang CF et al (2021b) A new dataset, Poisson GAN and AquaNet for underwater object grabbing. IEEE Trans Circuits Syst Video Technol 32(5):2831–2844. https://doi.org/10.1109/TCSVT.2021.3100059

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. https://doi.org/10.1109/TCSVT.2019.2963772

Liu RS, Jiang ZY, Yang SZ, Fan X (2022) Twin adversarial contrastive learning for underwater image enhancement and beyond. IEEE Trans Image Proc 31:4922–4936. https://doi.org/10.1109/TIP.2022.3190209

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. https://doi.org/10.1007/978-3-319-46448-0_2

Mou L, Zhang XW, Zhang JJ, Shen XH, Xu XL (2017) Saliency detection of underwater target based on spatial probability. In: 2017 International Conference on Computer Systems, Electronics and Control, Dalian, pp 630–632. https://doi.org/10.1109/ICCSEC.2017.8446733

Mukherjee K, Gupta S, Ray A, Phoha S (2011) Symbolic analysis of sonar data for underwater target detection. IEEE J Ocean Eng 36(2):219–230. https://doi.org/10.1109/JOE.2011.2122590

Nagaraja S, Prabhakar CJ, Kumar PUP (2015) Extraction of texture based features of underwater images using RLBP descriptor. In: Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications, Bhubaneswar, pp 263–272. https://doi.org/10.1007/978-3-319-12012-6_29

Pedersen M, Bruslund Haurum J, Gade R, Moeslund TB (2019) Detection of marine animals in a new underwater dataset with varying visibility. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Long Beach, pp 18–26

Peng XH, Liang ZX, Zhang J, Chen RF (2021) Review of underwater image preprocessing based on deep learning. Comput Eng Appl 57(13):43–54 (in Chinese with English abstract)

Rashwan A, Kalra A, Poupart P (2019) Matrix Nets: a new deep architecture for object detection. In: 2019 IEEE/CVF International Conference on Computer Vision Workshops, Seoul, pp 2025–2028. https://doi.org/10.1109/ICCVW.2019.00252

Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Shi XT, Huang H, Wang B, Pang S, Qin HD (2019) Underwater cage boundary detection based on GLCM features by using SVM classifier. In: 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Hong Kong, pp 1169–1174. https://doi.org/10.1109/AIM.2019.8868517

Singh P, Deepak BBVL, Sethi T, Murthy MDP (2015) Real-time object detection and tracking using color feature and motion. In: 2015 International Conference on Communications and Signal Processing, Melmaruvathur, pp 1236–1241. https://doi.org/10.1109/ICCSP.2015.7322705

Song DL, Sun WC, Ji ZH, Hou GJ, Li XF, Liu L (2014) Color model selection for underwater object recognition. In: 2014 International Conference on Information Science, Electronics and Electrical Engineering, Sapporo, pp 1339–1342. https://doi.org/10.1109/InfoSEEE.2014.6947890

Srividhya K, Ramya MM (2017) Accurate object recognition in the underwater images using learning algorithms and texture features. Multimed Tools Appl 76:25679–25695. https://doi.org/10.1007/s11042-017-4459-6

Sun X, Shi JY, Liu LP, Dong JY, Plant C, Wang XH et al (2018) Transferring deep knowledge for object recognition in Low-quality underwater videos. Neurocomputing 275:897–908. https://doi.org/10.1016/j.neucom.2017.09.044

Susanto T, Mardiyanto R, Purwanto D (2018) Development of underwater object detection method base on color feature. In: 2018 International Conference on Computer Engineering, Network and Intelligent Multimedia, Surabaya, pp 254–259. https://doi.org/10.1109/CENIM.2018.8711290

Taud H, Mas JF (2018) Multilayer perceptron (MLP). In: Geomatic approaches for modeling land change scenarios. Springer, Cham, pp 451–455. https://doi.org/10.1007/978-3-319-60801-3_27

Tucker JD, Azimi-Sadjadi MR (2011) Coherence-based underwater target detection from multiple disparate sonar platforms. IEEE J Ocean Eng 36(1):37–51. https://doi.org/10.1109/JOE.2010.2094230

Wang HB, Zhang Q, Wang X, Chen Z (2014) Object detection based on regional saliency and underwater optical priors. Chin J Sci Instrum 35(2):387–397. https://doi.org/10.19650/j.cnki.cjsi.2014.02.021 (in Chinese with English abstract)

Wei XY, Yu L, Tian SW, Feng PC, Ning X (2021) Underwater target detection with an attention mechanism and improved scale. Multimed Tools Appl 80:33747–33761. https://doi.org/10.1007/s11042-021-11230-2

Wu Y, Cai YB, Tang RH (2019) Research on the underwater optical imaging processing and identification. Ship Electron Eng 39(5):93–96 (in Chinese with English abstract)
Yang HH, Xu GH, Yi SZ, Li YQ (2019) A new cooperative deep learning method for underwater acoustic target recognition. In: OCEANS 2019-Marseille, Marseille, pp 1–4. https://doi.org/10.1109/OCEANSE.2019.8867490
Yu H (2020) Research progresson object detection and tracking techniques utilization in aquaculture: a review. J Dalian Ocean Univ 35(6):793–804 (in Chinese with English abstract)
Yu XL, Qu YY, Hong M (2019) Underwater-GAN: underwater image restoration via conditional generative adversarial network. In: 24th International Conference on Pattern Recognition (ICPR), Beijing, pp 66–75. https://doi.org/10.1007/978-3-030-05792-3_7
Yuan X, Guo LX, Luo CT, Zhou XT, Yu CL (2022) A survey of target detection and recognition methods in underwater turbid areas. Appl Sci 12(10):4898. https://doi.org/10.3390/app12104898
Zeng LC, Sun B, Zhu DQ (2021) Underwater target detection based on Faster R-CNN and adversarial occlusion network. Eng Appl Artif Intell 100:104190. https://doi.org/10.1016/j.engappai.2021.104190
Zhang MH, Xu SB, Song W, He Q, Wei QM (2021) Lightweight underwater object detection based on YOLO v4 and multi-scale attentional feature fusion. Remote Sens 13(22):4706. https://doi.org/10.3390/rs13224706
Zhou XY, Yang KD, Duan R (2019) Deep learning based on striation images for underwater and surface target classification. IEEE Signal Proc Lett 26(9):1378–1382. https://doi.org/10.1109/LSP.2019.2919102
Zhu YF, Chang L, Dai JL, Zheng HY, Zheng B (2016) Automatic object detection and segmentation from underwater images via saliency-based region merging. In: OCEANS 2016-Shanghai, Shanghai, pp 1–4. https://doi.org/10.1109/OCEANSAP.2016.7485598
Zurowietz M, Nattkemper TW (2020) Unsupervised knowledge transfer for object detection in marine environmental monitoring and exploration. IEEE Access 8:143558–143568. https://doi.org/10.1109/ACCESS.2020.3014441

Publisher's Note