



OPEN ACCESS

EDITED AND REVIEWED BY
Hervé Claustre,
Centre National de la Recherche Scientifique
(CNRS), France

*CORRESPONDENCE
Haiyong Zheng
✉ zhenghaiyong@ouc.edu.cn

RECEIVED 26 March 2024
ACCEPTED 23 April 2024
PUBLISHED 03 May 2024

CITATION
Zheng H, Bi H, Cheng X and Benfield MC
(2024) Editorial: Deep learning for
marine science.
Front. Mar. Sci. 11:1407053.
doi: 10.3389/fmars.2024.1407053

COPYRIGHT
© 2024 Zheng, Bi, Cheng and Benfield. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Editorial: Deep learning for marine science

Haiyong Zheng^{1*}, Hongsheng Bi², Xuemin Cheng³
and Mark C. Benfield⁴

¹College of Electronic Engineering, Ocean University of China, Qingdao, China, ²Center for Environmental Science, University of Maryland, College Park, Cambridge, MA, United States, ³Shenzhen International Graduate School, Tsinghua University, Shenzhen, China, ⁴College of the Coast & Environment, Louisiana State University, Baton Rouge, LA, United States

KEYWORDS

research survey, marine/underwater image enhancement/restoration/compression, marine/underwater visual recognition/detection, dataset and labeling, marine process/phenomenon prediction/detection, marine physical/biogeochemical variable prediction/reconstruction, marine optics/acoustics

Editorial on the Research Topic Deep learning for marine science

In recent years, Deep Learning (DL) technology has been widely used in marine science and technology research, and provides powerful technical support for related researches and applications. As ocean observation technology continues to advance, the volume of data generated by marine scientific research is steadily increasing. This offers vast potential for data-driven DL to demonstrate its capabilities and has therefore emerged as a valuable technology across multiple research fields, including biology, ecosystems, climate, energy, as well as physical and chemical interactions.

The Research Topic “Deep Learning for Marine Science” aims to provide a research collection to collect relevant research work on the application of DL technology in marine science. A total of 39 papers are published with contributions by 236 authors. The contents in these papers focus on the following aspects: research survey, marine/underwater image enhancement/restoration/compression, marine/underwater visual recognition/detection, dataset and labeling, marine process/phenomenon prediction/detection, marine physical/biogeochemical variable prediction/reconstruction, and marine optics/acoustics. Here, we summarize the contents of these papers and highlight their key contributions to the Research Topic.

1 Research survey

Although machine learning tools hold great promise, they are still not being used to their full potential in several areas, such as species and environmental monitoring, biodiversity surveys, fisheries abundance and size estimation, rare events, and species detection, the study of animal behavior, and citizen science. To help researchers effectively apply image-based machine learning methods in their research problems, [Belcher et al.](#) write a review article that provides an easily approachable end-to-end guide.

In terms of underwater image restoration technology, [Song et al.](#) make a systematic review to bridge the gap between shallow sea and deep-sea image restoration through experimental analysis. The review mainly describes the core concepts and methods of the three types of shallow sea image restoration methods. It also summarizes the research status and main challenges of deep-sea image restoration, discusses potential solutions, conducts experiments and in-depth discussions, and proposes several development directions for deep-sea image restoration in the future.

2 Marine/underwater image enhancement/restoration/compression

It is a challenging task to store and transmit high-quality underwater images. To improve the performance of adaptive sampling and reconstruction of underwater images, [Li et al.](#) combine the advantages of compressed sensing and DL to propose ESPC-BCS-Net. The method obtains parameters (such as sampling matrix, sparse transforms, and shrinkage thresholds) through end-to-end learning. The experimental results are visually and quantitatively evaluated, demonstrating that the proposed method has good compression and reconstruction effects.

[Xin et al.](#) introduce an end-to-end network for Simultaneous Localization And Mapping (SLAM) pre-processing in low-light underwater environments, aiming to address the limitations of visual SLAM systems based on feature point extraction. The proposed network comprises a low-light enhancement branch with a non-reference loss function, a self-supervised feature point detector, and a descriptor extraction branch. Additionally, a unique matrix transformation method is designed to enhance the feature similarity between two adjacent video frames, thereby improving the performance of underwater SLAM.

In order to solve the important problems of blur and color distortion in underwater optical imaging and improve the ability to accurately perceive underwater images, [Zhang et al.](#) propose a multi-scale weighted fusion method. By merging, enhancing, and reconstructing images, the clarity and color fidelity of underwater images are effectively improved, and the quality of underwater images presented is improved. Excellent results have been obtained in many experimental indexes.

[Zheng et al.](#) propose a solution to improve the performance of underwater monocular visual SLAM systems. The existing SLAM algorithms are often impractical or invalid due to the complex aquatic environment and the poor image quality obtained in such conditions. The proposed solution involves using a Generative Adversarial Network (GAN) to enhance the underwater images before SLAM processing. To reduce the inference cost, the GAN is compressed through knowledge distillation. This approach ensures real-time inference and high-fidelity underwater image enhancement.

To improve the quality of underwater images and achieve simultaneous restoration and super-resolution, [Wang et al.](#) propose an end-to-end trainable model named Simultaneous Restoration and Super-Resolution GAN (SRSRGAN). The model

uses GANs and consists of two stages of a cascading architecture to restore and super-resolve damaged underwater images coarse to fine. The proposed method is experimentally validated and demonstrates its superiority in underwater image restoration, super-resolution, and simultaneous restoration and super-resolution.

3 Marine/underwater visual recognition/detection

In order to realize the fast navigation of Unmanned Surface Vehicle (USV) in complex marine environments, a target detection algorithm with high detection speed and accuracy is essential. To address this Research Topic, [Zhang et al.](#) propose a YOLOv5 lightweight object detection algorithm that leverages the Ghost module and Transformer, resulting in high-efficiency and high-precision object detection. The proposed algorithm is tested on ship videos collected by the “JiuHang 750” USV in different marine environments and demonstrates promising results.

To address the problem of ship instance segmentation in Synthetic Aperture Radar (SAR) images with high resolution and complex backgrounds, [Yasir et al.](#) propose a unique YOLOv7 improved high-resolution remote sensing (HR-RS) image segmentation single-stage detection method. The method enhances the accuracy, efficiency, and model robustness of ship instance segmentation through improvements made to the single-stage detector, backbone network, and network feature fusion part, and promising results have been achieved.

To enhance the economic and environmental performance of the fishery, [Avsar et al.](#) utilize underwater images captured by an in-trawl video recording system to obtain quantitative information on the capture rate of *Nephrops norvegicus*, a target species. The study employs real-time detection, tracking, and counting techniques to monitor the entry of the target species into the trawl. The detection is done using the YOLOv4 algorithm, which has a proven track record in real-time processing underwater images to determine the target species' capture rate. Additionally, the algorithm has the potential to process multiple species simultaneously.

[Saito et al.](#) utilize DL to investigate the suspended particles in the depths of the sea. To analyze the variability of suspended particle abundance in the images taken by the standard fixed camera “Edokko Mark 1”, they implement object detection technology through the YOLOv5 algorithm to create a suspended particle detection model. They conduct the first excavation test of cobalt-rich ferromanganese crust in the world. The ability of the model to measure changes in the concentration of deep-sea suspended particles is assessed, and the effectiveness of the proposed method in detecting temporal changes of suspended particles and detecting significant abrupt changes, such as mining effects, is validated.

Collecting data on marine fish can be a challenging task due to the nature of their environment, often resulting in poor-quality data. Moreover, identifying various fish categories from small sample images can be difficult, especially regarding fine-grained classification. [Zhai et al.](#) propose a new attention network called the

Sandwich Attention Covariance Metric Network (SACovaMNet), which applies metric learning and incorporates attention modules to comprehensively improve the feature extraction capability from global and local perspectives. The result is an excellent performance in the task of fine-grained fish classification.

Prior et al. develop automated video post-processing models to implement automated image analysis of commercially important Gulf of Mexico fish species and habitats. In addition to traditional metrics used to measure the performance of Artificial Intelligence and Machine Learning (AI/ML) models, such as mean Average Precision (mAP), the automated counts are compared to validated set counts to ensure accuracy. The adapting comparative otolith aging methods and metrics are used to measure the model performance, which helps researchers analyze and make management decisions. This approach provides a valuable tool for analyzing Gulf of Mexico fish species and habitats.

Han et al. propose a few-shot domain adaptive underwater object detection framework to address the issues of expensive establishment of marine species database and unstable domain shifting of underwater objects caused by the complex marine environment. The framework includes a novel two-stage training method and a lightweight feature correction module that can adapt to image-level and instance-level domain shifting on multiple datasets. The method quickly demonstrates its knowledge transfer capability in detecting two similar marine species.

Through the sea trial experimental data, Guo et al. propose to automatically identify inbound and outbound ships by utilizing the phenomenon that the sound field interference structures of inbound and outbound ships are different due to the variation of the topography of the shallow continental shelf. The approach utilizes only a single scalar hydrophone to collect data and employs four convolutional neural networks to classify inbound and outbound ships. And this research method can be applied to the intelligent monitoring of ships entering and leaving ports.

To address the challenge of applying DL algorithms to underwater target detection tasks due to the complex underwater environment and low image quality, Zhang et al. propose an underwater target detection algorithm based on an improved version of YOLOv4. This proposed method achieves superior detection performance and efficiency in experiments by incorporating a newly designed convolutional network module, loss function, and detector strategy.

Large-scale research on plankton classification, which uses machine learning techniques, requires powerful computing resources. The exponential computing power of quantum computers makes quantum machine learning a potential solution for large-scale data processing. Therefore, Shi et al. propose a hybrid quantum-classical convolutional neural network (CNN) for the identification task of phytoplankton. The model demonstrates the feasibility of using quantum deep neural networks for phytoplankton classification for the first time. The proposed model exhibits a faster convergence rate, higher classification accuracy, and lower accuracy fluctuation compared to classic CNN-based models.

Commercial fishing vessels face difficulties in collecting acoustic data required for species classification and population evaluation

due to the limited calibration capability and frequent data loss of current commercial echo sounders. To address this issue, Tong et al. develop an automatic detection and classification model for Pacific saury (*Cololabis saira*) echo trace using the YOLOv5m algorithm. This model enables the measurement of *in-situ* values of Pacific saury using a single fish echo trace. Furthermore, the living fish calibration method is utilized to facilitate rapid calibration of commercial echo sounders.

To measure the fish without disturbing their natural habitat and overcome the limitation of manual measurement with potentially harmful intervention, Marrable et al. propose a generalized, semi-automatic method that combines the DL method with the high-precision stereo-BRUVS calibration method. The calibration cube is used to ensure that the accuracy of the calculated length is within a few millimeters and that the measurement accuracy is close to the accuracy of human measurements.

In order to distinguish the subtle changes of marine organisms and achieve accurate fine-grained classification, Si et al. propose a new transformer-based framework, token-selective vision transformer, and also propose a token-selective self-attention to select important tokens with discrimination for attention calculation, so as to limit attention to more accurate local areas. Experiments on three marine biological datasets verify that the proposed method can achieve state-of-the-art performance.

Current DL methods face challenges in processing *in-situ* plankton images due to large computation and long consumption time. To address this issue, Yue et al. propose an inter-class similarity distillation algorithm. This method enables the student network (small scale) to acquire excellent plankton recognition ability under the guidance of the teacher network (large scale). The experiment proves helpful in improving the accuracy and speed of plankton recognition, establishing effective DL models, and facilitating the deployment of underwater plankton imaging systems.

To address the ever-changing marine environments and diverse marine life, Schmid et al. implement edge computing technology by integrating the latest *In-situ* Ichthyoplankton Imaging System-3 (ISIIS-3) in the Northern California Current. The edge server utilizes DL techniques to achieve high-throughput *in-situ* plankton classification technology for real-time data adaptive sampling.

In order to develop and evaluate a subtidal seagrass detector method, Langlois et al. adopt a DL model to detect most forms of seagrass appearing in various habitats in the seascape of northeast Australia from underwater images, and classify them according to the coverage degree of seagrass to obtain high accuracy, and better application value and prospects.

To create a non-invasive method to recognize leopard coral grouper (*Plectropomus leopardus*), Wang et al. develop a multiscale image processing method based on matched filters with Gaussian kernels and partial differential equation (PDE) multiscale hierarchical decomposition with the deep convolutional neural network models VGG19 and ResNet50 to extract shape and texture image features of individuals. They then use these features to identify individual *Plectropomus leopardus* in sequence images captured over 50 days. To achieve this, they employ random forest, support vector machine, and multi-layer perceptron methods for individual recognition. The experimental results demonstrate that

the CNN based on PDE decomposition can identify *Plectropomus leopardus* effectively and with great accuracy.

4 Dataset and labeling

Catalán et al. create a new labeling dataset with the aim to further study and improve the application of DL techniques in identifying and classifying fish in underwater images. The dataset consists of more than 18,400 recorded Mediterranean fish from 20 different species, which are obtained through various operations such as different backgrounds, sample size, labeling quality, etc. These fish were extracted from underwater images captured from over 1,600 diverse backgrounds, which will assist in improving the use of DL in studying underwater life.

To achieve efficient data labeling and reduce the cost of manual labeling, Zhang et al. propose a weakly supervised learning framework for labeling marine biological data. This method utilizes crowdsourcing interfaces to converge to a labeled image dataset through multiple training and production loops. Experimental results demonstrate that training with a small subset and iterating over the results can converge to a large, highly annotated dataset with a small number of iterations.

Remote sensing technology can potentially capture aerial images of cetaceans across a vast observation area. However, current limitations in automated analysis techniques require biologists to manually analyze all images, leading to exorbitant tagging costs. Boulent et al. propose a human-in-the-loop approach that merges the proficiency of biologists with DL-based automation capabilities to create a reliable AI-assisted annotation tool for large-scale cetacean monitoring.

DL has been applied to the image classification of marine echinoderms in response to the need for automatic classification in marine biology research worldwide. Zhou et al. collect image data of marine echinoderms and classify them according to systematic taxonomy. Based on the DL model EfficientNetV2, an automatic classification tool (EchoAI) is developed. The EchoAI tool, along with methods and strategies, can classify images of other categories of marine organisms, thus helping researchers investigate the diversity, abundance, and distribution of marine species.

5 Marine process/phenomenon prediction/detection

Song et al. propose a new method called Time-Sequence-Involved Space Discretization neural network (TSI-SD) to solve the problem of large computation amount and high complexity of the fluid numerical model. This method extracts grid correlations from both spatial and temporal views simultaneously and combines TSI-SD with finite volume format as an advection solver for passive scalar advection in a two-dimensional unsteady flow field. Compared to the previous method that only considers spatial context, TSI-SD achieves higher simulation accuracy and reduces the calculation amount. Comprehensive experiments have verified the superior computational efficiency and accuracy of this method.

Song et al. propose a spatio-temporal transformer network that overcomes the defects of existing methods in network structure design and prediction errors to accurately, quickly and effectively predict ENSO events. This network simulates the inherent characteristics of spatio-temporal variations of sea surface temperature anomaly maps and heat content anomaly maps and takes into account the influence of seasonal variations on the prediction of ENSO phenomena. Additionally, an effective recurrent forecasting strategy is proposed, which takes previous predictions as prior knowledge to improve the reliability of long-term forecasting.

Aiming at addressing the problem that the current method only uses single-modal Sea Surface Height (SSH) data to detect mesoscale eddy, which often leads to inaccurate results, Zhao et al. propose an end-to-end mesoscale eddy detection method based on multi-modal data fusion, and add the data of the Sea Surface Temperature (SST) and the velocity of flow. The superior performance of the proposed method is demonstrated on various multi-modal mesoscale eddy datasets.

In view of the problem that the ocean front detection method in the Southwestern Atlantic Front (SAF) mainly adopts the thermal gradient method while ignoring dynamic features, which leads to inaccurate manifestation of SAF. Wang et al. develop a DL model, SAFNet, to detect the SAF through the synergistic effect of satellite SST and SSH observation data in 10 years (2010-2019), to achieve high-precision SAF detection with the fusion of thermal and dynamic features.

6 Marine physical/biogeochemical variable prediction/reconstruction

Based on satellite observations, machine learning has successfully reconstructed the high-resolution ocean subsurface thermohaline structure. However, due to the macro-tidal environment and limited *in-situ* observations, the offshore subsurface parameter estimation accuracy will be affected. Yu et al. propose a new approach by coupling the TPXO tidal model and light gradient boosting machine algorithm to develop an inversion model of offshore subsurface thermal structure for the South Yellow Sea (SYS) using sea surface data and *in-situ* observations. The experimental results show that the reconstruction is reliable in the SYS area, and the proposed method also provides a new exploration direction for reconstructing offshore ocean thermal structures.

For the reconstruction of satellite-derived chlorophyll-a concentration in a global scale, Roussillon et al. propose a method based on physical predictors, and uses a multi-mode convolutional neural network to globally account for interregional variabilities via learning and combining different modes spatially. The different modes show regional consistency with ocean dynamics, and the work contributes to new insights into the physical-biogeochemical processes that control temporal and spatial variability in phytoplankton on a global scale.

The current status of the sea surface carbon dioxide partial pressure (pCO₂) in the Yellow Sea is unclear due to limited availability of *in-situ* spatial and temporal distribution data. To

address this problem, [Li et al.](#) develop a pCO₂ model using a random forest algorithm. The model uses 14 cruise datasets from 2011 to 2019, as well as input variables such as remote sensing satellite sea surface temperature, chlorophyll concentration, diffuse attenuation of downwelling irradiance, and *in-situ* salinity. The model is trained and tested, yielding excellent prediction and evaluation results.

[Cutolo et al.](#) develop a CLuster Optimal Interpolation Neural Network (CLOINet) to combine remote-sensing data with *in-situ* observation and create a comprehensive 3D reconstruction of the ocean state. CLOINet combines the robust mathematical framework of the optimal interpolation scheme with a self-supervised clustering method and also effectively segments remote sensing images into clusters to reveal non-local correlations and enhance fine-scale ocean reconstruction. The network is trained using the output of the Ocean General Circulation Model and shows good reconstruction results in various testing scenarios.

7 Marine optics/acoustics

[Huang et al.](#) propose a Task-driven Meta-Deep-Learning (TDML) framework to solve the problem that the nonuniform distribution of sound speed will bring difficulties to underwater accurate positioning. It learns the common features of the Sound Speed Profile (SSP) through multiple base learners, accelerates the model convergence on new tasks, and enhances the model's sensitivity to changes in sound field data through metatraining. Thus, the over-fitting effect is weakened, and the inversion accuracy is improved. Experimental results show that the proposed TDML method can achieve fast and accurate spatio-temporal SSP inversion.

To fully consider how water environment and communication equipment affect signal transmission and accurately simulate the complex characteristics of the Underwater Wireless Optical Communication (UWOC) systems, [Huo et al.](#) develop a UWOC

channel emulator based on deep convolutional conditional generative adversarial networks, which are tested in experiments to verify their excellent performance in the time domain, frequency domain, and universality under different water turbidity levels.

To achieve full acoustic tracking of whales with reverberation interference, [Jin et al.](#) propose an intelligent acoustic tracking model that enables horizontal direction discrimination and distance/depth perception by mining unpredictable features of position information directly from signals received from two hydrophones. The proposed method not only achieves satisfactory prediction performance, but also effectively avoids the reverberation effect of signal propagation over long distances.

Author contributions

HZ: Writing – original draft. HB: Writing – review & editing. XC: Writing – review & editing. MB: Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.