



## OPEN ACCESS

## EDITED AND REVIEWED BY

Hervé Claustre,  
Centre National de la Recherche Scientifique  
(CNRS), France

## \*CORRESPONDENCE

Haiyong Zheng  
✉ zhenghaiyong@ouc.edu.cn

RECEIVED 24 September 2024

ACCEPTED 04 October 2024

PUBLISHED 22 October 2024

## CITATION

Zheng H, Nie J, Zhou H, Liu A-A and Zhang X  
(2024) Editorial: Deep learning for marine  
science, volume II.  
*Front. Mar. Sci.* 11:1501225.  
doi: 10.3389/fmars.2024.1501225

## COPYRIGHT

© 2024 Zheng, Nie, Zhou, Liu and Zhang. 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, volume II

Haiyong Zheng<sup>1\*</sup>, Jie Nie<sup>2</sup>, Huiyu Zhou<sup>3</sup>, An-An Liu<sup>4</sup>  
and Xiangrong Zhang<sup>5</sup>

<sup>1</sup>College of Electronic Engineering, Ocean University of China, Qingdao, China, <sup>2</sup>Faculty of Information Science and Engineering, Ocean University of China, Qingdao, China, <sup>3</sup>School of Computing and Mathematical Sciences, University of Leicester, Leicester, United Kingdom, <sup>4</sup>School of Electrical and Information Engineering, Tianjin University, Tianjin, China, <sup>5</sup>The Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, School of Artificial Intelligence, Xidian University, Xi'an, China

## KEYWORDS

Research survey, marine/underwater image enhancement/restoration/compression, marine/underwater visual recognition/detection/segmentation, marine process/phenomenon prediction/detection, marine physical/biogeochemical variable prediction/reconstruction, marine optics/acoustics, other marine/underwater applications

## Editorial on the Research Topic

### Deep learning for marine science, volume II

Deep learning (DL), a branch of artificial intelligence (AI), has become a pivotal technology across various scientific fields due to its ability to handle complex data and uncover patterns indiscernible to human analysts. In marine science, this technology has not only improved data processing capabilities but also provided novel insights into marine environments and phenomena. This editorial provides an overview of recent advancements in DL technology tailored for marine science, covering a range of research from image enhancement and visual understanding to predictive modeling of marine physical and biogeochemical processes.

This Research Topic, entitled “*Deep Learning for Marine Science, Volume II*”, is an extension of the previous topic of [Volume I](#), which continues to collate a significant collection of research works that leverage DL technologies to navigate and interpret the complexities of marine science. Featuring a total of 26 pioneering articles authored by 132 contributors, this Research Topic not only underscores the innovative applications of DL in marine and underwater image processing (spanning enhancement, restoration, and compression), but also in visual recognition and detection. Moreover, it extends to the predictive modeling of marine processes and phenomena, the reconstruction of biogeochemical variables, and advancements in marine optics and acoustics. Herein, we summarize the multifaceted contributions of these papers, analyzing and emphasizing their significance in the study of marine scientific research.

## 1 Research survey

While DL has shown significant potential in various scientific disciplines, its application in physical oceanography remains underexplored, particularly in areas such

as ocean circulation, ocean dynamics, ocean climate, ocean remote sensing, and ocean geophysics. To provide a comprehensive overview of recent advancements and guide future research, [Zhao Q. et al.](#) present a thorough review article that categorizes and analyzes the cutting-edge applications of DL in physical oceanography over the past three years. This review not only introduces the core concepts and methodologies related to DL models like CNNs, RNNs, and GANs but also highlights their applications across various oceanographic phenomena. Furthermore, it identifies the current bottlenecks and discusses innovative prospects, offering valuable insights for researchers aiming to leverage DL in their oceanographic studies.

## 2 Marine/underwater image enhancement/restoration/compression

To address the challenges of color attenuation and contrast reduction in underwater images caused by complex lighting conditions, [Liu T. et al.](#) propose a lightweight, zero-reference parameter estimation network (Zero-UAE) for adaptive enhancement of underwater images. The method introduces an underwater adaptive curve model based on light attenuation principles and a set of non-reference loss functions tailored for underwater scenarios. The experiments on three datasets demonstrate that Zero-UAE effectively enhances underwater images, achieving competitive state-of-the-art performance while maintaining minimal computational requirements and providing an application solution for extreme underwater conditions.

[Zhang H. et al.](#) propose an innovative framework for compressing underwater images aiming at improving machine vision applications in underwater environments. The framework includes two modules: the frequency-guided underwater image correction module (UICM) and the task-driven feature decomposition fusion module (DFDM). The UICM utilizes frequency priors to address noise issues and accurately identify redundant information, while the DFDM focuses on preserving machine-friendly information during compression by emphasizing task relevance. Extensive experiments on various downstream visual tasks, such as object detection, semantic segmentation, and saliency detection, show that the proposed framework significantly enhances performance at low bit rates, effectively mitigating the impact of compression on underwater visual tasks.

[Liu B. et al.](#) propose a novel lightweight DL model known as the Multi-Scale Dense Spatially-Adaptive Residual Distillation Network (MDSRDN) for underwater image super-resolution. The method aims to tackle the challenges of low image quality in underwater environments caused by scattering, absorption, and hardware limitations. MDSRDN utilizes a multi-scale dense spatially-adaptive residual distillation module and a spatial feature transformer layer to improve feature extraction, achieving high-quality image reconstruction with fewer parameters and computational cost. Experimental results on public datasets (USR-248 and UFO-120) demonstrate the superior performance

of the model in terms of PSNR, SSIM, and UIQM compared to state-of-the-art methods while maintaining efficient operation on edge devices.

## 3 Marine/underwater visual recognition/detection/segmentation

To address the challenges posed by quality degradation in underwater images, such as color casts, low contrast, and blurred details, [Wang Y. et al.](#) propose an innovative underwater superpixel segmentation network (USNet). The network incorporates a multi-scale water-net module (MWM) to enhance the quality of underwater images prior to segmentation. Additionally, it includes a degradation-aware attention mechanism (DA) that focuses on regions with significant quality degradation. By integrating deep spatial features with a dynamic spatial embedding module (DSEM), USNet effectively improves segmentation accuracy and robustness in complex underwater scenes. Extensive experiments demonstrate that USNet outperforms existing state-of-the-art methods in terms of segmentation accuracy, under-segmentation error, and boundary recall, thus providing a new benchmark for underwater image segmentation tasks.

To address the challenges posed by underwater image distortions due to turbulence, [Zhou et al.](#) propose a novel multi-scale aware turbulence network (MATNet) that incorporates a multi-scale feature extraction pyramid module with dense linking and position learning strategies. The network enhances object recognition by effectively extracting and correcting features from distorted underwater images. Experimental results show that MATNet outperforms state-of-the-art methods in both qualitative and quantitative assessments, providing a robust solution for underwater object recognition in complex environments.

To overcome the limitations of traditional zooplankton size measurement methods, such as labor-intensive manual sampling and analysis, [Zhang et al.](#) propose a novel DL-based approach utilizing a modified deep residual network (ResNet50) for accurate and efficient size measurement of zooplankton. The proposed method employs key point detection technology, replacing the fully connected layer with a convolutional layer to generate predictive heatmaps that enable precise size estimation, particularly for organisms with complex or curved shapes. The approach is validated against manual measurements from *in-situ* images of three zooplankton groups—copepods, appendicularians, and shrimps—collected by the PlanktonScope imaging system, demonstrating high consistency with a minimal average discrepancy of 1.84%. This automated method offers a rapid and reliable tool for large-scale zooplankton size measurement, facilitating improved ecosystem-based management decisions and advancing marine biological research.

To improve food safety, production efficiency, and economic benefits in the aquatic industry, [Kim et al.](#) develop a DL-based phenotype classification method for identifying and classifying three commercially important ark shell species: *Anadara*

*kagoshimensis*, *Tegillarca granosa*, and *Anadara broughtonii*. The study applies three convolutional neural network (CNN) models (VGGnet, Inception-ResNet, and SqueezeNet) to classify 1,400 images of ark shells and tested their performance on two different classification sets, as well as a combined classification set. The results show that SqueezeNet achieved the highest accuracy during the training phase, while Inception-ResNet performed best in the validation phase. This research provides a theoretical basis for image-based bivalve classification, offering a promising approach to enhance identification accuracy and efficiency in the aquatic products industry and supporting automation in production processes.

Aquatic biodiversity monitoring is crucial for conservation purposes, but identifying species in complex underwater environments can be challenging. To address these challenges, Ma D. et al. propose a novel semi-supervised learning (SSL) approach to improve species recognition by leveraging large amounts of unlabeled data. They also propose a wavelet fusion network (WFN) that can better capture high- and low-frequency features of underwater images, combined with a consistent balance loss (CEL) function to alleviate the long-tail class imbalance problem. The approach has significantly improved classification accuracy on the FishNet dataset, indicating its potential to advance automatic species identification in aquatic biodiversity monitoring and conservation.

## 4 Marine process/phenomenon prediction/detection

In response to the challenges of predicting fishing effort distribution due to the lack of integration between hydrological factors and fishing activity data, Shi et al. introduce HyFish, a novel DL model that combines Vessel Monitoring System (VMS) data with hydrological factor fields, such as Sea Surface Temperature (SST), Sea Surface Height (SSH), salinity, and ocean currents. The model utilizes residual networks and Long Short-Term Memory (LSTM) to capture both spatial and temporal dynamics, achieving highly accurate daily predictions of fishing effort distribution for the upcoming week with an average error ratio of 5.6%, as demonstrated on extensive datasets from the East China Sea.

Rui et al. present a two-stage spatiotemporal autoregressive model that improves ENSO prediction by integrating self-attention ConvLSTM networks and temporal embeddings of calendar month and seasonal information. The model consists of two phases: first, it employs a self-attention ConvLSTM to forecast meteorological time series, capturing both local and global spatiotemporal dependencies; then, it refines the predictions using a convolutional network to produce ENSO indicators. Their method demonstrates effective forecasting of ENSO up to 24 months in advance and successfully overcomes the spring predictability barrier. This approach outperforms existing models by leveraging short- and long-term spatiotemporal features as well as accounting for seasonal variations.

Xu et al. propose a machine learning-based approach to predict the features of convergence zones (CZ) in ocean front

environments, which are crucial for underwater acoustic propagation and target detection. After testing 24 different machine learning algorithms, they identify a hybrid model combining a multilayer perceptron (MLP) and random forest (RF) as the most effective for predicting the distance and width of CZs with high accuracy. The proposed model achieves 82.43% accuracy for CZ distance predictions within a 1 km error margin and demonstrated strong generalization capabilities across different datasets, proving its applicability in complex marine environments. The study also highlights the significance of turning depth and other environmental features in influencing CZ characteristics, suggesting that machine learning can effectively capture the nonlinear relationships between oceanographic features and CZ behavior.

Ding et al. propose a hybrid model named VMD-LSTM-rolling, to address the issue of non-stationarity in predicting significant wave height in the South Sea of China. This model combines Variational Mode Decomposition (VMD) with a Long Short-Term Memory (LSTM) neural network using a rolling decomposition method, which effectively avoids the information leakage problem present in traditional direct decomposition methods. By utilizing this approach, only known data is used in the prediction process, ensuring both prediction accuracy and practical applicability. Comparative experiments demonstrate that the VMD-LSTM-rolling model significantly improves both short-term and long-term prediction accuracy compared to the conventional LSTM model and the VMD-LSTM-direct model. These results highlight its effectiveness in handling non-stationary data for marine process prediction.

For the estimation of subsurface temperature anomalies (ESTA) associated with mesoscale eddies in the Northwest Pacific Ocean, Liu S. et al. propose a novel method that integrates multi-source satellite observations with Argo float data using a residual multi-channel attention convolution network (ERCACN). The ERCACN model effectively combines diverse remote sensing features, such as sea level anomaly (SLA), sea surface temperature anomaly (SSTA), and surface wind speed anomaly (SSWSA), with their components to accurately estimate the three-dimensional temperature structures at depths of up to 1000 m. The proposed approach significantly outperforms traditional methods, demonstrating a precision of 88.08% in predicting temperature anomalies, providing new insights into the spatial and temporal variability of oceanic processes driven by mesoscale eddies and contributing to an improved understanding of the global climate system.

To accurately reconstruct the acoustic fields of mesoscale eddies and improve the understanding of their impact on underwater sound propagation, Ma X. et al. develop a mesoscale eddy reconstruction method named EddyGAN based on generative adversarial networks (GAN). This method employs a hybrid algorithm for eddy identification using JCOPE2M high-resolution reanalysis data and AVISO satellite altimeter data to extract mesoscale eddy sound speed profile (SSP) samples. The EddyGAN model is then trained to reconstruct the mesoscale eddy acoustic field. The proposed model is evaluated against root mean square error (RMSE), structural similarity index (SSIM), and convergence zone (CZ) accuracy, achieving an RMSE of 1.7 m/s, an

SSIM of 0.77, and an average CZ accuracy exceeding 70%, thereby demonstrating superior performance over conventional GAN and other reconstruction methods.

## 5 Marine physical/biogeochemical variable prediction/reconstruction

In the area of Estimating Sea Surface Particulate Organic Carbon (POC) Based on Geodetector and Machine Learning,” [Wu et al.](#) conduct a systematic study that improved the estimation accuracy of global ocean POC concentrations through model evaluation and comparative analysis. The study primarily elucidates the core concepts and modeling processes of six machine learning-based POC estimation methods and provides an in-depth analysis of the current research status and main challenges of using geodetectors to identify key factors influencing POC concentration. It explores potential solutions to enhance the accuracy of POC estimation by integrating machine learning techniques and performs experimental validation and performance comparison. Finally, the study proposes several future directions for achieving higher accuracy in POC estimation in complex marine environments.

Based on the need for better understanding of the vertical distribution of chlorophyll-a (Chl-a) in the ocean, which is crucial for evaluating marine ecosystems under global climate change, [Zhao X. et al.](#) develop a Gaussian-activation deep neural network (Gaussian-DNN) model. This model reconstructs the three-dimensional structure of Chl-a in the northwestern Pacific Ocean using satellite-derived surface Chl-a data and *in-situ* vertical profiles of temperature and salinity. The results demonstrate that the Gaussian-DNN model accurately captures over 80% of Chl-a vertical profiles at a high spatial resolution of  $1^\circ \times 1^\circ$  and 1 m depth intervals. This approach provides a new method for long-term 3D Chl-a reconstruction, highlighting the key role of seawater temperature and salinity in controlling Chl-a distribution across different regions and offering insights into seasonal and interannual variability in marine biogeochemical processes.

Given the critical role of Marine Dissolved Oxygen Concentration (MDOC) in evaluating seawater conditions and its implications for global climate regulation, traditional approaches such as numerical computation and deep learning have faced challenges in terms of interpretability and computational transparency. To overcome these limitations, [Li et al.](#) propose a novel framework, CDRP, which integrates Causal Discovery, Drift Detection, RuleFit Model, and *Post Hoc* Analysis to achieve high-precision and interpretable MDOC inversion. This framework utilizes the PCMCI algorithm for causal discovery to elucidate the relationships between marine elements, and drift detection to optimize the selection of representative training data. The RuleFit model ensures both precision and transparency in inversion processes, while SHAP and LIME analyses provide comprehensive insights into operational mechanisms. This approach not only enhances the interpretability of the inversion process but also achieves optimal performance compared to existing methods, making a significant contribution to the field of marine biogeochemical variable reconstruction.

Based on a combination of oceanic and atmospheric data, [Aleshin et al.](#) develop a machine learning-based model to predict

chlorophyll-a (Chl-a) concentration in the northern marine regions, specifically focusing on the Barents Sea. Due to limitations of satellite observations during polar night, dense cloud cover, and sea ice, their approach leverages outputs from the Weather Research & Forecasting (WRF) model and the Nucleus for European Modelling of the Ocean (NEMO) to provide additional predictors when remote sensing data are unavailable. The study compares classical machine learning algorithms like LightGBM with DL approaches such as ResNet-18 for forecasting Chl-a concentration over an 8-day period. The models demonstrate different strengths: while LightGBM shows higher overall prediction accuracy ( $R^2 = 0.578$ ), ResNet-18 provides better performance in minimizing prediction errors across varying data points (MAPE = 0.528). This work contributes to advancing predictive capabilities for biogeochemical variables in challenging northern marine environments where limited data availability is due to adverse weather and lighting conditions.

## 6 Marine optics/acoustics

Based on the corrigendum of the article, [Huang et al.](#) introduced a meta-deep-learning framework designed for the spatio-temporal inversion of the Sound Speed Profile (SSP) in underwater environments. This framework utilizes artificial neural networks (ANN) and few-shot learning to enhance the model's generalization capability and minimize over-fitting, especially when training data is limited. The approach incorporates task-driven meta-learning (TDML) to adapt efficiently to various underwater environments by learning from a range of signal propagation simulations, thereby improving the model's capacity to accurately reconstruct SSP with fewer data samples. Despite a minor formula error in the original publication, it has been rectified by the authors without altering the scientific conclusions.

To enhance the automatic modulation classification (AMC) of underwater acoustic signals in complex environments, [Wang C. et al.](#) propose a novel spatial-temporal fusion neural network that combines Transformer and depth-wise convolution (DWC) networks. This method utilizes the attention mechanism of the Transformer to improve the recognition of key information and integrates DWC blocks to establish a spatial-temporal structure, thereby achieving superior classification performance at low signal-to-noise ratios (SNRs). Experimental results demonstrate that the proposed network outperforms state-of-the-art neural networks, achieving an average classification accuracy of 92.1% for SNRs ranging from -4 dB to 0 dB. These findings underscore its potential for robust underwater signal classification.

To address the challenges of recognizing ship-radiated noise in complex marine environments, where natural sound interference and signal distortion complicate the extraction of acoustic features, [Wang Y. et al.](#) propose a novel hybrid framework, DWSTr. This framework combines a depthwise separable convolutional neural network (DWSCNN) with a Transformer architecture to effectively isolate local acoustic features while capturing global dependencies. As a result, it enhances robustness against environmental interferences and signal variability. The experimental results on

the ShipsEar dataset demonstrate a notable 96.5% recognition accuracy, highlighting DWSTR's efficacy in accurate ship classification and its potential for real-time analysis in passive acoustic monitoring applications.

In their recent study, [Lyu et al.](#) have proposed a novel network architecture based on a space-time neural network to tackle the challenge of Automatic Modulation Identification (AMI) for underwater acoustic signals. The new method combines the strengths of Convolutional Neural Networks (CNNs) and Transformers, utilizing the attention mechanism to dynamically adjust feature aggregation weights based on the relationship between signal sequences and location information. Furthermore, the network features a hybrid routing structure to improve classification performance, especially in low signal-to-noise ratio (SNR) environments. Experimental findings indicate that the proposed method achieves an average recognition accuracy of 89.4% for SNRs ranging from -4 dB to 0 dB, surpassing other state-of-the-art neural network models.

[Zhu et al.](#) propose a novel algorithm-unrolled neural network model, named SSANet, to tackle the challenge of extracting the Normal-Mode Interference Spectrum (NMIS) from the received Sound Intensity Spectrum (SIS) in underwater environments with low signal-to-noise ratios. The SSANet model unrolls the traditional Singular Spectrum Analysis (SSA) algorithm into a DL framework, enhancing its noise robustness and reducing information loss during NMIS extraction. Simulation results in canonical ocean waveguide environments demonstrate that SSANet outperforms traditional methods such as Fourier Transform (FT), Multiple Signal Classification (MUSIC), and SSA, particularly under low SNR conditions. This provides a promising approach for applications such as underwater source ranging and waveguide-invariant estimation.

## 7 Other marine/underwater applications

To tackle the challenge of effectively identifying high-risk areas (accident black spots) in maritime search and rescue (MSAR) resource allocation, [Sun et al.](#) propose an optimization method for MSAR resource allocation based on accident black spot clustering. The method utilizes the Iterative Self-Organizing Data Analysis Technique (ISODATA) to cluster historical accident data and identify accident black spots, followed by the entropy weight method to assess the importance of each spot. Subsequently, a

multi-objective optimization is performed using a Non-Dominated Sorting Genetic Algorithm II combined with reinforcement learning (NSGAI-RL). Experimental results demonstrate that this method can save at least 7% of rescue time compared to traditional methods, significantly enhancing the efficiency and stability of maritime search and rescue operations.

Solving high-dimensional partial differential equations (PDEs) with complex structures presents a significant challenge for conventional numerical solvers and current DL methods. To address these challenges, [Chen et al.](#) propose DF-ParPINN, a parallel physics-informed neural network based on velocity potential field division and single time slice focus. This method divides the overall velocity potential field into multiple time slices and further into high-velocity and low-velocity fields, enabling parallel computation to handle large data efficiently. The experimental results demonstrate that DF-ParPINN achieves significantly higher accuracy and faster computation time than existing methods such as PINN, PIRNN, cPINN, and DeepONet, providing an effective solution for solving high-dimensional PDEs in ocean physics.

## Author contributions

HYZhe: Writing – original draft. JN: Writing – review & editing. HYZho: Writing – review & editing. A-AL: Writing – review & editing. XZ: 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.