



Linking fish activity and turbidity through visual and sensor data fusion and deep learning

Mohammad Jahanbakht^{a,b,*}, Andrea Tiernan^a, Alzayat Saleh^{a,b}, Nichola Stokes^c, Odette Langham^c, Mostafa Rahimi Azghadi^{a,b,d,*}, Nathan J. Waltham^{a,d}

^a College of Science and Engineering, James Cook University, Townsville, QLD, 4814, Australia

^b Centre for AI and Data Science Innovation, James Cook University, Townsville, QLD, Australia

^c North Queensland Bulk Ports, Level 1, Waterfront Place, Mulherin Drive, Mackay, QLD, 4740, Australia

^d Centre for Tropical Water and Aquatic Ecosystem Research (TropWATER), James Cook University, Townsville, QLD, 4814, Australia

ARTICLE INFO

Keywords:

Turbid water monitoring
Fish detection
Turbidity estimation
Deep learning
Image and sensor fusion

ABSTRACT

Monitoring underwater environments is crucial for industrial applications, providing data that can be used for reporting against corporate sustainability and environmental goals. This study presents a novel approach to integrating high-resolution underwater imaging and high-tech water quality sensing with deep learning models to detect fish, estimate turbidity in Nephelometric Turbidity Units (NTU), and analyze their interactions. An IP-based underwater camera and two advanced water quality sensors were deployed at the Port of Mackay (northern Queensland, Australia) to collect synchronized visual and water quality data. A significant portion of collected images lacked valid turbidity values due to camera and sensor synchronization issues. To address this, we developed a custom Convolutional Neural Network (CNN) model for image-based turbidity estimation. Additionally, YOLOWorld-based prompt-able object detectors were used and evaluated for fish detection, with YOLOWorld-v1 Large emerging as the best choice, achieving 89.7 % accuracy without any training. Our proposed CNN water turbidity estimation model gained root mean square error of 1.6 NTU. Using these deep learning models, we found a non-linear correlation between fish count and water turbidity with an R^2 of 0.93. This finding is aligned with previous research and highlights the complex interplay of environmental factors in marine ecosystems, while showcasing how technological advances can streamline ecological studies. Downstream applications of this technology could include permanently installed underwater cameras in port waters that record real-time data. Management responses could then be automatically triggered when water quality parameters exceed threshold levels, providing early warnings and enabling timely actions to protect marine ecosystems.

1. Introduction

Coastal port developments are among the most extensive anthropogenic modifications of marine ecosystems, fundamentally altering coastal environments on a global scale (Firth et al., 2024). These large-scale infrastructures are essential for international trade, economic growth, and urban expansion, but they introduce significant environmental challenges, particularly regarding the biodiversity and ecological function of marine communities (Verschuur et al., 2022). The proliferation of port and associated structures, including breakwaters, piers, and seawalls, has transformed natural coastlines, replacing soft-sediment and vegetated habitats with hard artificial substrates (Bishop

et al., 2017; Waltham and Sheaves, 2015). As a result, marine life around these engineered structures is often distinct from that of adjacent natural habitats, exhibiting altered community compositions, species interactions, and ecological processes (Airoldi et al., 2021).

Assessing the biodiversity of fish and other marine organisms around these structures is essential for understanding both the potential consequences and positive benefits of port expansions. This knowledge enables researchers and decision-makers to recognize ecological changes while also identifying opportunities to support marine life and enhance ecosystem services associated with these developments. However, quantifying fish communities in deep-water, highly engineered environments presents methodological challenges (Williams et al.,

* Corresponding authors.

E-mail addresses: mohammad.jahanbakht@jcu.edu.au (M. Jahanbakht), Mostafa.rahimiazghadi@jcu.edu.au (M. Rahimi Azghadi).

<https://doi.org/10.1016/j.marpolbul.2025.119070>

Received 3 August 2025; Received in revised form 24 November 2025; Accepted 25 November 2025

Available online 1 December 2025

0025-326X/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2021). Traditional survey methods employed to measure fish and aquatic species in coastal areas, such as netting, trapping, and diver-based observations (Whitfield and Elliott, 2002), are impractical due to depth, water clarity, and the physical/operational complexity of port structures (Malcolm et al., 2021).

In response to these limitations, underwater video methods have emerged as a non-invasive, and scalable approach, for assessing fish communities (Whitmarsh et al., 2017). This technique has proven especially valuable in marine environments where traditional sampling methods are constrained, offering an effective means of evaluating fish assemblages (Dunlop et al., 2015; Jahanbakh et al., 2023). Baited remote underwater video systems and remotely operated vehicles are increasingly being used to capture footage of fish communities, enabling species identification and abundance estimation (Langlois et al., 2020). These methods provide several advantages, including minimal disturbance to the environment, the ability to operate in deep and structurally complex areas, and the capacity for long-term, repeatable observations (Harvey et al., 2018).

Despite these advantages, underwater video monitoring is subject to several limitations, particularly in turbid environments where visibility is reduced (Cappo et al., 2004). Many port areas can experience high sediment resuspension due to vessel traffic, dredging activities, and hydrodynamic conditions, while the design of port infrastructure (e.g., lower seabed areas and the construction of breakwaters) further creates still-water environments that promote the accumulation and subsequent resuspension of fine sediments, collectively leading to decreased image clarity and difficulty in species identification (Seiler et al., 2020). Poor visibility impedes manual image analysis and reduces the effectiveness of traditional video-based monitoring approaches (Heagney et al., 2007). Addressing these challenges; for example, managing port related sediment relocation, requires developing new methods that improve the usefulness of underwater video in turbid coastal waters, where turbidity directly affects fish species by altering predator avoidance, foraging behavior, as well as community structure and distribution (Cyrus and Blaber, 1987; Lowe et al., 2015; Whitfield, 2021).

Recent advancements in artificial intelligence (AI) and computer vision have provided new opportunities for improving underwater video analysis (Saleh et al., 2024), particularly in challenging environmental conditions (Connolly et al., 2021; Goodwin et al., 2021). Machine learning algorithms have demonstrated significant potential for automating species identification (Saleh et al., 2022) from video footage, reducing the labor-intensive nature of manual analysis and increasing data processing efficiency (Rodriguez et al., 2022). These AI-driven approaches can enhance the accuracy of fish identification even in low-visibility conditions by leveraging spectral, shape, and movement-based recognition techniques (Spampinato et al., 2010).

The combination of AI with real-time image enhancement enables researchers to extract more accurate biological data from underwater footage, overcoming the challenges posed by suspended sediments and low-light conditions (Bryson et al., 2016). These integrated approaches have the potential to revolutionize the way fish communities are assessed around coastal port infrastructures, leading to more robust ecological assessments and informed conservation strategies. A potential downstream application involves the detection of water quality conditions exceeding species-specific thresholds, with this technology providing near real-time notifications to management to facilitate timely mitigation responses.

Overall, as coastal port developments continue to expand to meet global demand, understanding their ecological influences on marine life is critical for sustainable management and conservation efforts, necessitating the adoption of underwater video monitoring. However, the challenges posed by turbid water conditions require innovative solutions that integrate AI and advanced image processing technologies. This study deployed an underwater camera alongside a high frequency water quality turbidity logger to establish a workflow to examine fish communities occurring within port infrastructure and under variable

water clarity conditions – ranging from high to low turbidity visibility (Cartwright et al., n.d.).

The Port of Mackay (−21.112°; 149.227°), situated along the Queensland coast adjacent to the Great Barrier Reef World Heritage Area, is among several designated priority port developments in the region. The port's engineered seawalls and deep-water environments, which facilitate safe navigation, have been demonstrated to support a high diversity of fish species (Waltham et al., 2023; Bradley et al., 2023). Although the study was conducted within a port setting, our intention was not to assess pollution-related attributes of the Port of Mackay. Instead, the port provided a controlled and accessible environment in which to evaluate the performance of an advanced monitoring system. This setting allowed us to examine how turbidity influences fish behavior in built environments and to test the effectiveness of underwater video and high-frequency loggers in capturing these dynamics. The aim was to understand the relationship between water turbidity and fish activity through deep learning analyses. By combining these cutting-edge methodologies, researchers can enhance the accuracy and efficiency of fish community assessments in port environments, ultimately supporting more informed decision-making in coastal infrastructure planning and ecological management. In the below Sections, we describe the study development and our results.

2. Hardware description

To facilitate real-time underwater observation and to study water quality effects on fish species in turbid waters, an underwater camera system was deployed in the waters of North Queensland Bulk Ports Corporation Port of Mackay, alongside a water quality sensor. As shown in Fig. 1, the entire system is securely mounted on a metallic structure designed for stable underwater operation.

2.1. Underwater camera

The *View into the Blue*® (VITB) camera is a high-definition, IP-based, pan-tilt-zoom (PTZ) underwater camera, featuring *CleanSweep*™ technology for automated lens cleaning. This camera provides live 1080p video streaming at up to 30 frames per second with a resolution of 1920 × 1080 pixels. It features a 4.7–47 mm autofocus lens with an automatic day/night mode, offering a horizontal field of view ranging from 6.7 to 61.8°. The PTZ functionality allows for 360-degree pan, 90-degree tilt, and an 8× optical zoom.

For power and data transmission, the camera is connected via a single 48 V cable that delivers 13 W of power from the wharf and integrates it into the local area network (LAN) using a standard TCP/IP protocol. The camera is rated for depths of up to 50 m, and its *CleanSweep*™ system, equipped with a 9 N-m torque cleaning arm, ensures continuous optical clarity by preventing biofouling and sediment accumulation on the glass dome.

2.2. Water quality sensor

The system incorporates a high-frequency logging sensor from In-situ *Marine Optics*® (IMO): the Nephelometric Turbidity Unit (NTU) sensor. The sensor unit is compact, fully submersible up to 90 m, and equipped with an automated wiper to maintain clear measurement apertures. The unit features internal rechargeable lithium-ion batteries that must be charged before each deployment.

The NTU sensor measures particle concentration in water using an infrared LED emitter and a photodiode detector operating in synchronous scattered light detection mode. This configuration minimizes the interference from ambient light sources, such as sunlight. The NTU sensor must be deployed at a 0-degree angle, facing the horizon, with an integrated tilt sensor for deployment verification. The data collected includes LED temperature, detector temperature, NTU raw counts, and NTU calibrated values.

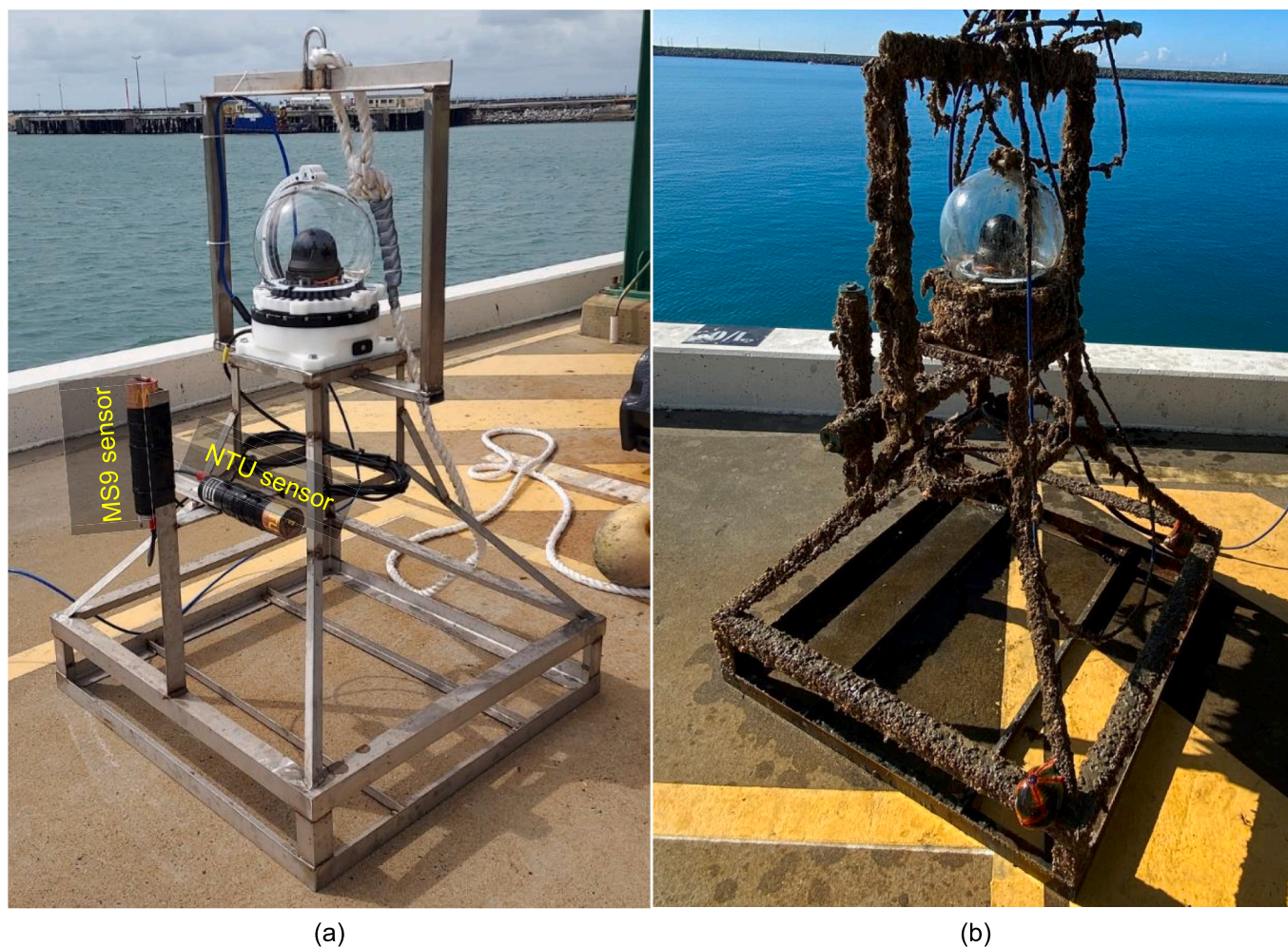


Fig. 1. The underwater camera mounted on a metallic frame with NTU sensor (a) before deployment in Mackay Port in March 2023 and (b) two months later. An MS9 sensor was also deployed; however, the data were excluded from this study because the analysis focused exclusively on turbidity. The biofilm growth in (b) occurs commonly in coastal waters and includes a mixture of encrusting bivalves, macroalgae and biofilm, soft sponges and possibly soft corals. Of note is the CleanSweep™ system that effectively cleans the Perspex dome around the camera unit. (c) Representative sample images taken by the underwater camera are provided to illustrate the fish diversity in the area. The corresponding NTU value for each image is displayed at the bottom of the image.

In addition to the abovementioned sensor-specific data, the following data is also collected similarly by both sensors: sensor's serial number, date and time, wiper position, input voltage, depth, water temperature, and tilt degree.

2.3. Customized Y-shaped cable

To integrate both sensors with the camera system, a custom-designed Y-shaped cable was fabricated. This cable allows both sensors to transmit their data via an RS232 serial protocol to the camera, which then makes the information accessible over the TCP/IP network.

The Y-shaped cable has an overall length of 6 m, with a 3-m segment before the split and two 3-m branches extending to each sensor. The

cable is designed to withstand depths exceeding 50 m. Each connector is equipped with female locking sleeves to ensure a secure and waterproof connection. Fig. 2 provides a schematic view of the cable design, illustrating the wiring diagram, connector heads, and pin configurations, which are numbered clockwise for male connectors and counterclockwise for female connectors.

This integrated custom-developed hardware setup enables continuous underwater video monitoring while simultaneously collecting crucial water quality parameters, contributing to efficient environmental observation and research in the Mackay Port area.

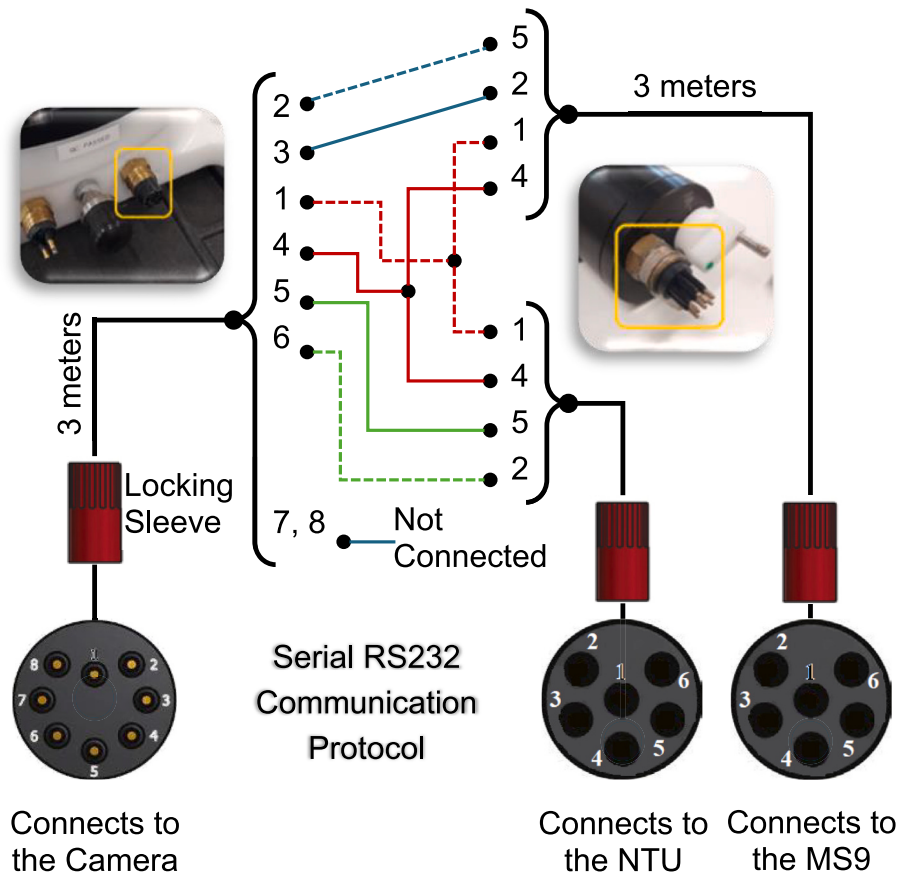


Fig. 2. The customized Y-shape cable to connect the two water quality sensors (NTU and MS9) into the camera, and from there to the worldwide web. Despite the installation of the MS9 sensor, its data were excluded from this study because the analysis focused exclusively on turbidity.

3. Software components

To effectively analyze underwater turbidity levels, a comprehensive dataset was collected using the underwater camera and water quality sensors. Given the inherent challenges of underwater monitoring, specialized data collection and processing techniques were developed to ensure high-quality and reliable operation. The following sections detail the dataset attributes, and two deep learning approaches employed to count fish species and to estimate turbidity levels from images, addressing the challenges posed by invalid or missing sensor readings.

3.1. Dataset attributes

The underwater monitoring system, consisting of the VITB camera and two IMO sensors, was installed at a depth of 11 m in the North Queensland Bulk Ports Corporation Port of Mackay (TropWATER, 2024). Since the camera and sensors operated as independent instruments, a custom Python-based data collection code was developed to ensure synchronized image acquisition and sensor measurements.

The data collection process begins with connecting to the camera system via its network APIs (VITB, 2025) to capture an underwater image and issuing a sensor data preparation command. Once the sensors complete their measurement process, they return the data to the code before the connection is terminated. Our code also verifies both camera and sensor data timestamps at the moment of data collection, ensuring their accurate timing. Due to the inherent operational delays and network latency, there is always a slight time difference between a captured image and its corresponding water quality data. To maintain measurement integrity, any sensor and camera data entry with a time difference exceeding 20 s was discarded. Additionally, the dataset was

restricted to daytime images to enhance consistency and usability.

Between March 21, 2023, and May 23, 2023, a total of 29,349 images were collected. Of these, 9331 images (32 %) were associated with valid NTU readings, while 20,018 images (68 %) were excluded due to excessive time differences with their corresponding NTU measurements, rendering them invalid.

To evaluate the performance of our fish-counting AI models, all images (including those with invalid NTU offsets) were manually reviewed by volunteer students to identify 1519 instances of fish presence. The review process was conducted on standalone images captured by the underwater camera; therefore, no video processing or frame extraction was required. Each image was binary labeled as either fish present or fish absent, depending on whether the human annotator could visually identify any part of a fish's body within the frame. To minimize annotation complexity, we did not consider factors such as fish size, species, or functional groups, primarily due to our limited access to volunteer annotators. Consequently, exploring the link between elevated turbidity and specific ecological mechanisms, rather than general fish count, falls beyond the scope of this study and is left for future research.

Additionally, images with valid NTU values were categorized into five turbidity levels (i.e., very low, low, medium, high, and very high) corresponding to those illustrated in Fig. 3. These levels represent five quantile ranges of turbidity, corresponding respectively to maximum values of 3, 6, 9, 12, and 15 NTU. The distribution of images across these categories is as follows:

- Very low: 1775 images
- Low: 4785 images
- Medium: 1944 images

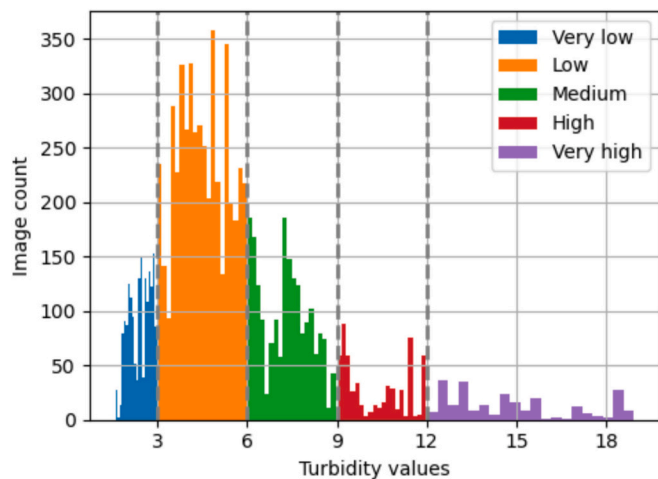


Fig. 3. The histogram of the valid/synchronous daytime image turbidities, categorized into five semantic groups.

- High: 575 images
- Very high: 252 images

A histogram of these categories is presented in Fig. 3, illustrating the distribution of turbidity levels within the collected dataset. This dataset provides valuable insights into underwater visibility conditions and water quality variations over the data collection period.

3.2. Deep learning models

3.2.1. Fish counting

To enable accurate fish counting in underwater environments, we employed two advanced deep learning models from the YOLOWorld framework: YOLOWorld-v1 Large and YOLOWorld-v2 XLarge. These models were selected after evaluating multiple versions and sizes of YOLOWorld for their suitability in detecting fish under varying turbidity conditions. Below, we detail the model structures, hyperparameters, and training methodologies utilized, focusing on their technical configurations tailored for this application.

The architecture of both YOLOWorld-v1 Large and YOLOWorld-v2 XLarge builds upon the YOLOv8 framework, enhanced for open-vocabulary object detection to accommodate the diverse fish species encountered in the Port of Mackay. For YOLOWorld-v1 Large, the backbone is based on YOLOv8-L, comprising approximately 43 million parameters dedicated to vision processing (Cheng et al., 2024). In contrast, YOLOWorld-v2 XLarge employs the larger YOLOv8-X backbone, with an estimated 68 million parameters (Cheng et al., 2024). Both backbones utilize a Darknet-based design, optimized for efficient feature extraction across complex underwater scenes. The neck of these models incorporates a *Re*-parameterizable Vision-Language Path Aggregation Network (RepVL-PAN), featuring Text-guided CSPLayer and Image-Pooling Attention modules to facilitate cross-modality fusion between visual and textual data (Cheng et al., 2024). This enhances the models' ability to associate fish detections with descriptive labels. The detection head, consistent with standard YOLO architecture, performs bounding box regression and generates object embeddings. Additionally, a frozen CLIP-base text encoder, with approximately 63 million parameters, processes textual inputs, contributing to a total parameter count of approximately 110 million when combined with the vision backbone (Cheng et al., 2024; Cheng, 2025).

The YOLOWorld models used in this study were employed in a zero-shot setting without any additional fine-tuning or re-training. According to the original YOLOWorld paper (Cheng et al., 2024), these models were trained, using the AdamW optimizer with an initial learning rate of 0.002 and a weight decay of 0.05 to prevent overfitting. The models

were trained for 100 epochs on 32 NVIDIA V100 GPUs, leveraging a batch size of 512 to ensure efficient parallel processing. To enhance robustness against underwater variability, data augmentation techniques were applied, including color adjustments (brightness and contrast), random affine transformations (scaling and rotation), random flips (horizontal and vertical), and mosaic augmentation combining four images per sample (Cheng et al., 2024). An online vocabulary of up to 80 nouns per mosaic sample was integrated during training. This vocabulary was dynamically sampled during training to expose the model to varied object-language associations. While the training vocabulary was broad to support open-vocabulary generalization, for the fish counting task, we constrained the inference prompt to a single class: "fish". This focused approach allowed the model to generalize across different species and shapes of fish without relying on species-level annotations, making it well-suited for detecting fish in diverse, turbid underwater conditions where species differentiation is less critical.

A key distinction between the models lies in their pre-training datasets. YOLOWorld-v1 Large was pre-trained on the Objects365 and Grounded Question Answering datasets, providing a robust foundation for general object detection (Cheng et al., 2024). YOLOWorld-v2 XLarge extends this pre-training by incorporating the Conceptual Captions 3 M dataset, which includes a large collection of image-text pairs, potentially enhancing its zero-shot detection capabilities across a broader range of underwater objects (Cheng et al., 2024). These pre-training differences, alongside the architectural variations, were considered during model selection to optimize performance for fish counting in turbid waters. The hyperparameters and structural configurations of YOLOWorld-v1 Large and YOLOWorld-v2 XLarge were meticulously designed to balance computational efficiency with detection accuracy, making them well-suited for real-time underwater monitoring.

In our work, we applied these pre-trained models as-is for inference on underwater datasets, constraining the inference prompt to a single class, "fish", to focus on general fish detection under diverse underwater conditions. This prompt-driven approach simplified the detection task while leveraging the model's open-vocabulary capabilities, making it adaptable to various ecological monitoring scenarios without needing extensive retraining or species-specific labelling. Further details on their comparative performance are presented in Section 4-A, where YOLOWorld-v1 Large was ultimately selected due to its optimal balance between detection accuracy and model efficiency.

3.2.2. Turbidity estimation

As discussed in Section 3-A, a total of 20,018 images were deemed invalid due to excessive time delays between image capture and their corresponding NTU measurements. Given the significant volume of discarded data, a deep learning-based approach was adopted to estimate water turbidity directly from the images.

To achieve this, we initially experimented with existing advanced Deep Neural Network architectures, including small, medium, large, and B3 variants of EfficientNet-v2 (Tan and Le, 2021). However, these models fall short due to their complex structures and huge number of training parameters. The high parameter count led to under-training and -fitting issues, making them inefficient for visual image turbidity estimation.

To address these limitations, we designed a simplified Convolutional Neural Network (CNN) tailored for turbidity estimation. The structural hyperparameters of this model were carefully optimized to maximize accuracy while maintaining computational efficiency. A schematic representation of the model is provided in Fig. 4.

The proposed CNN processes input images through a series of convolutional (Conv2D) and max-pooling layers, refining feature extraction at each stage. The extracted features are then batch normalized and flattened into a 1D array before being passed through a sequence of Multilayer Perceptron (MLP) layers. The number of convolutional layers and their filters, as well as the number of MLP layers and their neuron counts were optimized. It is worth noting that the final MLP layer

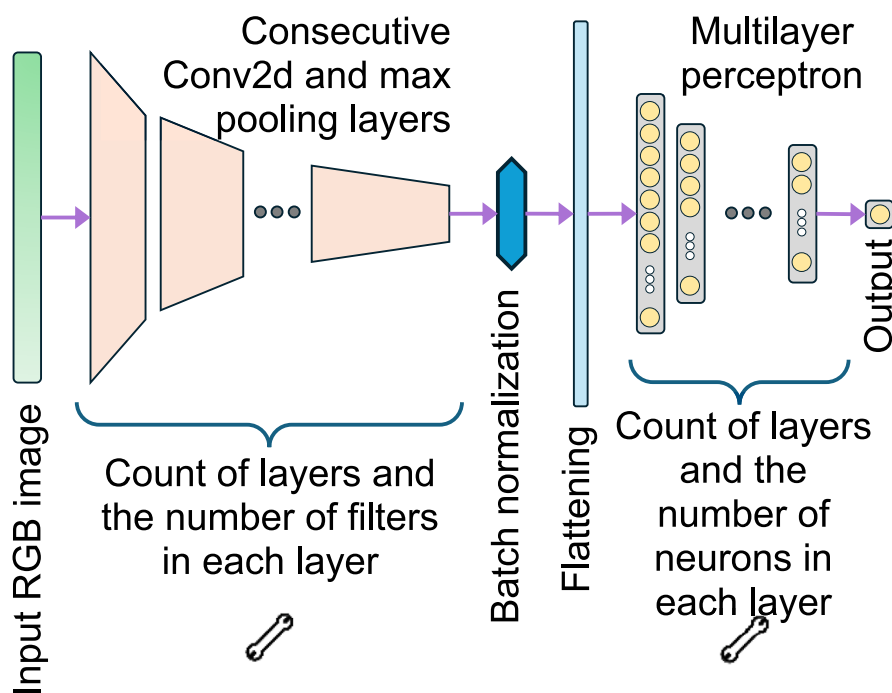


Fig. 4. The schematic view of the customized CNN model for water turbidity estimation, where the optimized variables are marked by a μ .

contains a single neuron representing the estimated NTU level.

Hyperparameter optimization results indicate that the optimal architecture consists of three consecutive Conv2D layers with 6, 13, and 21 filters, each followed by max pooling. These are then connected to three MLP layers with 1057, 84, and 21 neurons, respectively. This optimized model demonstrates superior accuracy in estimating turbidity levels while maintaining a lightweight structure suitable for real-world applications.

The optimized CNN model was trained on an NVIDIA GeForce RTX 4090 GPU using the Adam optimizer, with a batch size of 16 and an initial learning rate of 0.001. Training was carried out for 16 epochs, taking a total of 72.9 min. The model's performance was monitored on the validation dataset after each epoch, and the best-performing model was achieved during the 12th epoch. Early stopping based on validation performance was applied to prevent overfitting, and standard data augmentation and normalization techniques were used to improve generalization. The final selected model represents the optimal balance between training accuracy and validation performance.

4. Results and discussions

4.1. Fish counting performance

After evaluating various YOLOWorld-v1 and YOLOWorld-v2 model sizes (Small, Medium, Large, and XLarge), we identified that YOLOWorld-v1 Large achieved superior accuracy, whereas YOLOWorld-v2 XLarge demonstrated better precision. A comparative analysis of these two models is provided in Table 1.

YOLOWorld-v1 Large exhibited higher accuracy and recall, indicating its capability to correctly identify fish while minimizing False Negatives (FN). On the other hand, YOLOWorld-v2 XLarge had better precision and specificity, suggesting it excelled at reducing False Positives (FP). Additionally, YOLOWorld-v1 Large showed a slight advantage in F1-score, demonstrating a better balance between precision and recall. These differences are further reflected in their confusion matrices in Table 1, where YOLOWorld-v1 Large achieved comparable or superior True Positive (TP), True Negative (TN), and FN values. Based on these findings, YOLOWorld-v1 Large was selected as our deep learning

Table 1

Comparing the performance of two YOLOWorld variants in counting fish in daytime turbid waters.

Metrics	YOLOWorld-v1 Large	YOLOWorld-v2 XLarge																
Accuracy	0.897	0.844																
Precision	0.959	0.984																
Recall	0.813	0.673																
Specificity	0.970	0.990																
F1-score	0.880	0.800																
Confusion Matrix	<table border="1"> <tr> <td>TN</td> <td>FP</td> </tr> <tr> <td>52.2%</td> <td>1.6%</td> </tr> <tr> <td>FN</td> <td>TP</td> </tr> <tr> <td>8.6%</td> <td>37.6%</td> </tr> </table>	TN	FP	52.2%	1.6%	FN	TP	8.6%	37.6%	<table border="1"> <tr> <td>TN</td> <td>FP</td> </tr> <tr> <td>53.3%</td> <td>0.5%</td> </tr> <tr> <td>FN</td> <td>TP</td> </tr> <tr> <td>15.1%</td> <td>31.1%</td> </tr> </table>	TN	FP	53.3%	0.5%	FN	TP	15.1%	31.1%
TN	FP																	
52.2%	1.6%																	
FN	TP																	
8.6%	37.6%																	
TN	FP																	
53.3%	0.5%																	
FN	TP																	
15.1%	31.1%																	

model for fish counting throughout the study.

4.2. Turbidity estimation performance

Our optimized CNN model for turbidity estimation was benchmarked against EfficientNet-v2 Small, with results summarized in Table 2. To facilitate this comparison, all 9331 images with valid turbidity values were divided into a training set of 5598 images (60 %) and a validation set of 3733 images (40 %). The model was initially

Table 2

Comparison of the turbidity estimation performance between our optimized CNN model and an existing advanced deep learning model.

Metrics	EfficientNet-v2 Small	Optimized CNN
RMSE	5.078	1.685
Bias	-4.299	0.079
Abs. error	-4.320	0.057
	μ	
	σ	
Rel. error	2.710	1.680
	μ	
	σ	
Scatter index	-76.5 %	10.6 %
	11.1 %	31.5 %
	0.958	0.318

trained using only the training set, and its performance metrics were then evaluated on the validation set. As a deep learning regression model, our CNN significantly outperformed EfficientNet-v2 Small in terms of Root Mean Square Error (RMSE) of NTU, demonstrating a more precise turbidity estimation. Additionally, its bias (i.e., the average error) is calculated as,

$$Bias = \frac{1}{N} \sum_i (\hat{T}_i - T_i),$$

where N is the total number of samples and T_i and \hat{T}_i are the true and estimated turbidity values, respectively. The calculated bias is close to zero NTU, indicating minimal systematic errors in the predictions. This low bias is also reflected in the small mean (μ) and standard deviation (σ) of both the absolute and relative errors.

Dispersion, a measure of variability in a dataset, was also examined. Specifically, the scatter index (Bhattacharya and Sinha, 2022), an advanced and unitless dispersion metric based on relative observation frequency, is calculated as,

$$Scatter\ Index = \frac{RMSE}{\frac{1}{N} \sum_i T_i}$$

The calculated scatter index is noticeably lower in our optimized CNN model. This implies that our model achieved a more consistent and reliable turbidity estimation.

Using this optimized CNN model, turbidity levels were estimated for 20,018 images that previously lacked valid NTU values. The key findings from this regression process are:

- Min recorded image turbidity: 1.4
- Max recorded image turbidity: 20.0
- Min recorded turbidity in fish presence: 1.6
- Max recorded turbidity in fish presence: 14.8

This range in NTU aligns with the range experienced in the port waters, based on a long-term water quality monitoring program (Cartwright et al., 2024).

Fig. 5 presents a scatter plot comparing CNN-based regression

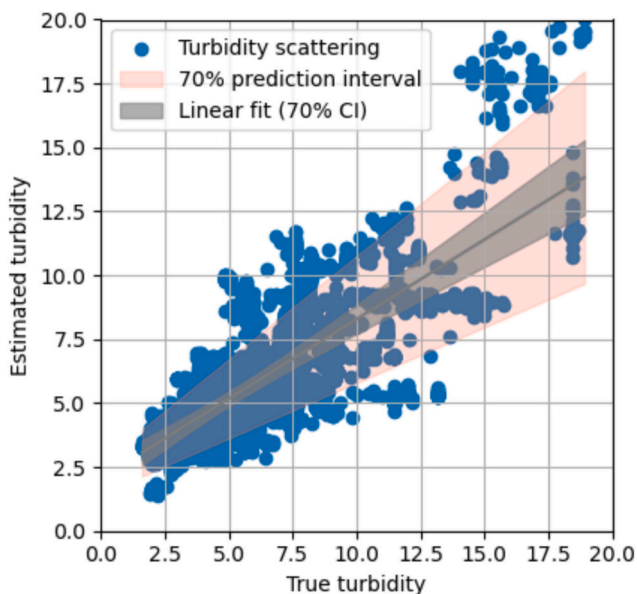


Fig. 5. Scatter plot of the optimized CNN model's output (i.e., the estimated turbidity on the y-axis) versus the true turbidity measured by NTU sensors. The plot includes the fitted linear regression, along with its 70 % prediction interval and 70 % confidence interval (CI).

estimates with true turbidity values. The 70 % prediction interval is measured by

$$RegrLine \pm (1 - 0.7) \times RegrLine,$$

where $RegrLine$ is the fitted linear regression line to the predicted values. Of the 9331 valid NTU samples, 8013 (~86 %) of the predicted values fall within the 70 % prediction interval, demonstrating the reliability of the model's predictions. This figure also contains the 70 % Confidence Interval (CI), calculated by

$$\hat{T} \pm t_{0.85} \times RMSE \times \sqrt{\frac{(T - \bar{T})^2}{\sum_i (T_i - \bar{T})^2}},$$

where $t_{0.85} \approx 1.04$ is the critical value from the t-distribution corresponding to a 70 % CI and \bar{T} is the average of the true NTU readings. The narrow width of the confidence interval (particularly in areas with a higher density of NTU measurements) indicates that the CNN model's turbidity predictions show a strong linear relationship and a good overall fit to the observed measurements.

4.3. Correlating fish count and turbidity

To analyze the relationship between fish count and water turbidity, we examined the correlation between the fish count in our dataset and the corresponding turbidity levels. This correlation is based on:

- 29,349 images (including 9331 valid and 20,018 estimated NTU samples) when using estimated turbidity values, and
- 9331 images (i.e., valid NTU samples) when using true turbidity values.

The results in Fig. 6 demonstrate a non-linear relationship. Fish count initially decreased as turbidity rose from very low to medium but then increased as turbidity reached very high levels. To further investigate this relationship, a detailed analysis is presented in Table 3. We began by calculating the Pearson and Poisson R^2 scores to assess the coefficients of determination for the linear and logarithmic regression models, respectively. In simple terms, the R^2 score quantifies how much of the variance in fish count (the dependent variable) can be explained by turbidity (the independent variable). The low R^2 values obtained for both Pearson and Poisson regressions indicate poor model fits, which was expected. Consequently, we explored polynomial regressions, where the cubic model achieved an R^2 score of 0.97, indicating an excellent fit.

It is important to note that this score was derived using estimated turbidity values, representing our engineered imputation approach to address missing data. To further evaluate the model's robustness, we also computed its R^2 score using true turbidity values, which yielded a value of 0.82 (see Table 3). This result suggests that 82 % of the variation in fish count can be explained by turbidity via the cubic model, highlighting a nonlinear relationship between turbidity and fish presence. The cubic regression fit, along with its 95 % confidence interval, is visualized in Fig. 6, as well.

4.4. Discussions and directions for future research

This study demonstrates the effectiveness of integrating underwater video imaging with deep learning models and high-frequency turbidity sensing to monitor fish communities in port environments. Traditional fish survey methods such as diver-based transects, nets, and traps are often impractical in ports due to restricted access, poor visibility, and physical hazards (Whitfield and Elliott, 2002; Malcolm et al., 2021). Our approach successfully addressed these limitations by using an IP-based underwater camera system combined with automated fish detection

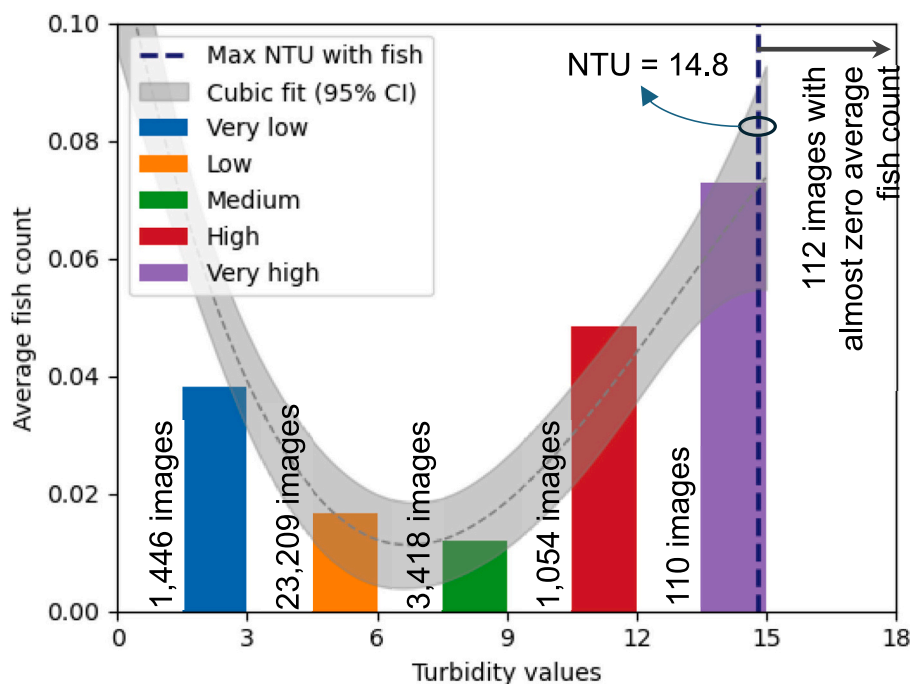


Fig. 6. Histogram of average fish counts across five turbidity levels, along with a cubic polynomial fit and its corresponding 95 % confidence interval (CI).

Table 3

Comparing multiple regression model performances (R^2 scores) when predicting estimated or ground-truth turbidity values.

Regression model	Estimated and true turbidities	True, turbidity values only
Linear – Pearson	0.41	0.05
Logarithmic – Poisson	0.29	0.03
Polynomial – Quadratic	0.93	0.19
Polynomial – Cubic	0.97	0.82

algorithms (YOLOWorld-v1 Large) and a custom CNN for turbidity estimation. The capacity to recover and utilize large volumes of image data, even where direct turbidity measurements were missing, marked a major strength of the approach. This advancement is consistent with global trends towards AI-enhanced marine monitoring (Marini et al., 2018; Monkman et al., 2022).

Nonetheless, the limitations of our approach warrant careful consideration. A significant challenge was the synchronization delay between video capture and sensor logging, which initially led to the loss of 68 % of image-turbidity pairs. Despite addressing this issue with a custom-build CNN model, future reliable systems must address it through improved network integration (Bouma et al., 2019). Additionally, although YOLOWorld-based models achieved high detection accuracy, visual fish detection accuracy for both human and machines inherently suffers from high turbidity conditions, where image degradation limits the ability to identify smaller or more cryptic species (Hannam et al., 2022; Ditría et al., 2020). Similar limitations have been reported in studies using baited remote underwater video systems, where visibility strongly affects detection rates (Harvey et al., 2012; Stobart et al., 2007). A solution is to focus the sampling approach to clear water periods (Schmid et al., 2017). Furthermore, our study focused only on daytime observations, thus missing nocturnal fish activity, which can be substantial in port ecosystems (Becker et al., 2018).

Both these limitations; namely, the constraints in technology to consider nocturnal activities and the reduction in visual fish detection in high turbidity environments, will reduce the ability to accurately assess fish communities and biodiversity. This may lead to the

underrepresentation of species adapted to turbid habitats and have implications on the effectiveness of ecosystem monitoring and management decisions. Also, while the water turbidity estimation CNN model performed well overall, any estimation approach introduces some uncertainty compared to direct instrument measurements. These issues highlight the need for continued methodological refinement, including improved sensor synchronization, night-vision capable cameras, and multi-modal sensor integration.

Comparison with previous studies highlights both consistency and advances. Underwater video studies in coastal and estuarine areas have shown declines in fish detection rates as turbidity increases (Whitmarsh et al., 2017; Cappo et al., 2003), and our results showing a non-linear fish-turbidity relationship are consistent with findings where fish use turbidity for predator avoidance (Lehtiniemi, 2005; Wenger et al., 2013). However, unlike previous studies, we successfully estimated turbidity directly from underwater images using deep learning, building on initial approaches explored by Spampinato et al. (Spampinato et al., 2010) and refined in more recent AI-aided ecological monitoring work (Ditría et al., 2020). The quadratic pattern we observed, where fish abundance initially decreased but increased again at very high turbidity, supports theories of trade-offs between foraging efficiency and predation risk under variable visibility (Abrahams and Kattenfeld, 1997). This insight suggests that management actions aimed at minimizing turbidity may not straightforwardly benefit all fish species but may shift community composition towards different functional groups depending on turbidity regimes.

An important ecological insight from this work is the complex, species-specific influence of turbidity on fish assemblages. While high turbidity often negatively affects visual predators and planktivores, it can simultaneously benefit species adapted to low-visibility environments (Wenger et al., 2013; Gray et al., 2012). The observed non-linear relationship, with fish abundance rising at very high turbidity levels, suggests that some species may actively exploit turbid conditions for refuge from predators or as improved foraging conditions (Lehtiniemi, 2005; Blaber and Blaber, 1980). Such findings are consistent with functional group shifts reported in estuarine and mangrove systems under changing turbidity (Breitburg, 1999). This nuance is critical for environmental management planning in coastal development, as the

observed non-linear relationship between fish abundance and turbidity highlights the importance of management strategies that recognize the ecological complexity of these environments. Rather than simply aiming to minimize turbidity in all situations, it is essential to support diverse functional groups and preserve ecosystem services within port and coastal systems.

Future research would benefit from trait-based analyses (e.g., feeding mode, sensory adaptations) which could be applied to a much larger dataset collected over a longer duration to include more fish with a wide range of traits. This more comprehensive data set could be used to improve predictions of fish assemblage responses to varying turbidity levels in coastal environments, including port facilities. Such information would enhance understanding of habitat selection patterns (such as the use of artificial infrastructure compared with natural vegetated habitats like mangroves or seagrass) and provide insights into food web dynamics and foraging strategies under different turbidity regimes. Additionally, it could inform new hypotheses regarding species-specific adaptations, including sensory compensation mechanisms employed by predators or prey to cope with varying turbidity conditions. Several studies have investigated fish responses and behavioral patterns under varying turbidity conditions using traditional sampling methods such as netting, visual surveys, and controlled tank experiments. In contrast, this study introduces an innovative approach that enables longer-term, minimally invasive deployments, reducing potential behavioral disturbance and experimental bias. This method provides enhanced insights for marine managers to support more effective management of coastal ecosystems.

Additionally, the current study was limited in its ability to resolve species-specific responses due to the constraints of automated fish detection models and the general challenge of species identification under poor visibility. Thus, developing AI models that better incorporate behavioral and morphological cues could significantly advance forward underwater ecological monitoring.

Considering these limitations, this study does ultimately demonstrate a scalable proof-of-concept for automated ecological monitoring in industrial port environments. Integrating underwater video, AI-based fish detection, and high-frequency water quality logging allows for near real-time assessment of environmental conditions and biological responses, supporting adaptive management approaches (Monkman et al., 2022; Hernandez et al., 2020).

A potential application of this study is the use of underwater cameras in port waters that record real-time data, not only on water quality parameters but also on fish assemblage behaviors as detected through advanced analytics. By integrating real-time monitoring of both environmental conditions and fish community responses, management systems could be programmed to automatically trigger responses when water quality parameters meet certain metrics. This offers the opportunity for these management responses and triggers to be calibrated based on real-time behavioral insights, ensuring that interventions are informed by the positive ecological roles observed at threshold turbidity levels, not just by environmental metrics alone. This could allow managers to proactively consider the broader ecological context and opportunities for ecosystem enhancement. By integrating sensor data with deep learning-based behavioral analysis, this adaptive framework supports more nuanced, ecosystem-based management strategies that can dynamically adjust to changing conditions, ultimately promoting the health and resilience of marine communities in areas adjacent to ports.

This study supports the recommendation that any development or review of threshold or region-specific water quality values should be directly informed by key ecological findings, particularly the non-linear responses of fish communities to turbidity. Fish in estuaries must adapt to highly variable turbidity by adjusting their foraging strategies, sensory reliance, and movement behavior. Some species shift from visual feeding to relying on lateral line or olfactory cues, while others modify swimming speeds or seek clearer microhabitats. Elevated or prolonged turbidity, however, can reduce feeding efficiency, impair predator

avoidance, disrupt migration cues, and smother critical habitats such as seagrass. As turbidity increases due to climate and catchment pressures, these effects may ultimately reduce growth, survival, and overall population resilience (Cyrus and Blaber, 1987; Lunt and Smeed, 2020; Johnston et al., 2007).

Management guidelines should acknowledge that some species may benefit from higher turbidity while others may be adversely affected. This would allow threshold values to better capture the unique characteristics of inshore environments and the diverse ways fish species respond to both natural and human-driven turbidity changes. During periods of human-induced turbidity, early detection of threshold breaches (e.g., those outlined in the Great Barrier Reef Water-Quality Guidelines (Environmental Protection (Water and Wetland Biodiversity) Policy, 2023)) using the system presented here are possible, helping to limit broader coastal ecosystem impacts.

Future research should include extended deployment of the camera rig over longer periods (6–12 months), providing not only a more comprehensive assessment of turbidity but also insights into seasonal patterns in fish community responses to changing turbidity conditions. Fully operationalizing this continuous logging at scale will require advancements such as improved biofouling mitigation, integration with acoustic monitoring for non-visual species detection (Williams, 2014), and extension into nocturnal monitoring. From a management perspective, the ability to detect biologically relevant turbidity thresholds and community shifts in near real-time could help port and coastal infrastructure managers to better align operations with environmental sustainability goals.

Future research should focus on extending this integrated monitoring framework across multiple sites, seasons, and habitat types to validate its broader applicability. Another extension to this application is to examine and establish effective detection distances under different turbidity conditions. This could be achieved by including similar design additions such as checkerboard visuals inside the field of view. Overall, this study advances the technical and conceptual foundations for smarter, more responsive marine biodiversity monitoring in heavily modified coastal systems.

5. Conclusion

This study successfully integrates deep learning models with underwater imaging and sensor data to achieve two core objectives: (1) detecting fish presence and (2) estimating water turbidity (in NTU). Key results demonstrate that:

- YOLOWorld-v1 Large delivers the optimal balance of accuracy (89.7 %) and recall for fish detection in turbid port waters, operating effectively in a zero-shot setting without additional fine-tuning.
- The optimized custom CNN model enables reliable image-based turbidity estimation, with a root mean square error (RMSE) of 1.6 NTU; outperforming benchmark models like EfficientNet-v2 Small.
- A strong nonlinear correlation ($R^2 = 0.93$ using estimated turbidity; $R^2 = 0.82$ using true turbidity) exists between fish count and turbidity, underscoring the complex interplay of environmental conditions and marine biotic responses.

Notwithstanding these contributions, the study faces two notable limitations: (1) synchronization delays between the underwater camera and NTU sensor resulted in 68 % of image-turbidity pairs being discarded (addressed via CNN imputation but requiring technical refinement); (2) the exclusion of nighttime data and reliance on daytime observations may underrepresent nocturnal fish activity.

To advance this work, future research should prioritize three actionable steps:

1. Implement network-synchronized timestamp protocols (e.g., Precision Time Protocol, PTP) to eliminate camera-sensor synchronization delays and reduce reliance on imputation.
2. Deploy low-light or infrared cameras, paired with deep learning models trained on augmented nighttime datasets, to capture nocturnal fish behavior and expand monitoring coverage.
3. Conduct extended deployments (6–12 months) to assess seasonal variability in turbidity and fish community responses, while integrating acoustic monitoring to detect non-visual species and improve biodiversity assessments.

Ultimately, this study provides a scalable proof-of-concept for AI-enhanced marine monitoring in industrial ports, and with continued refinement of the technical framework and resolution of current limitations, the approach can be operationalized to support real-time, ecosystem-based management. By enabling rapid detection of turbidity fluctuations and associated changes in fish presence, the system can serve as an early-warning tool for sediment resuspension, pollution events, and other anthropogenic disturbances, thereby facilitating automated conservation responses when water-quality parameters exceed ecologically relevant thresholds.

CRedit authorship contribution statement

Mohammad Jahanbakhht: Writing – original draft, Visualization, Validation, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation. **Andrea Tiernan:** Validation, Data curation. **Alzayat Saleh:** Software, Methodology. **Nichola Stokes:** Resources. **Odette Langham:** Resources. **Mostafa Rahimi Azghadi:** Writing – review & editing, Methodology. **Nathan J. Waltham:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We acknowledge the Traditional Owners of the land on which this research was completed, and pay respects to their Elders past, present and emerging. This study was funded by North Queensland Bulk Ports (NQBP), as part of the NQBP and James Cook University marine water quality and habitat monitoring partnership. The research project was completed in accordance with JCU animal ethics permit A2810. We also thank the staff at NQBP who assisted with the deployment of the camera and logger frame and assisted with maintenance of the set up over the study period. We also acknowledge assistance by J. Johns from JCU with logger maintenance and data processing.

Data availability

The authors do not have permission to share data.

References

Abrahams, M.V., Kattenfeld, M.G., 1997. The role of turbidity as a constraint on predator-prey interactions in aquatic environments. *Behav. Ecol. Sociobiol.* 40 (3), 169–174.

Airoldi, L., Turra, A., Chapman, M.G., 2021. The ecology of marine infrastructure: Insights and perspectives from the growing field of "urban marine ecology". *Oceanogr. Mar. Biol. Annu. Rev.* 59, 101–136.

Becker, A., et al., 2018. Managing the development of artificial reef systems: The need for quantitative goals. *Fish Fish.* 19 (4), 740–752.

Bhattacharya, D., Sinha, N., 2022. Scatter Index: An alternative measure of dispersion based on relative frequency of occurrence of observations. *Springer*, pp. 65–72.

Bishop, M.J., et al., 2017. Effects of ocean sprawl on ecological connectivity: impacts and solutions. *J. Exp. Mar. Biol. Ecol.* 492, 7–30.

Blaber, S.J.M., Blaber, T.G., 1980. Factors affecting the distribution of juvenile estuarine and inshore fish. *J. Fish Biol.* 17 (2), 143–162.

Bouma, J.A., et al., 2019. Assessing the value of information for monitoring and evaluating ecosystem restoration projects. *Ecol. Econ.* 157, 28–38.

Bradley, M., Sheaves, M., Waltham, N.J., 2023. Urban-industrial seascapes can be abundant and dynamic fish habitat. *Front. Mar. Sci.* 9, 1034039.

Breitburg, D.L., 1999. *Are three-dimensional structure and healthy oyster populations the keys to an ecologically interesting and important fish community.* Oyster reef habitat restoration: a synopsis and synthesis of approaches. Virginia Institute of Marine Science Press, Gloucester Point, Virginia, pp. 239–250.

Bryson, M., et al., 2016. True color correction of autonomous underwater vehicle imagery. *Journal of Field Robotics* 33 (6), 853–874.

Cappo, M., Speare, P., De'ath, G., 2004. Comparison of baited remote underwater video stations (BRUVS) and prawn (shrimp) trawls for assessments of fish biodiversity in inter-reefal areas of the Great Barrier Reef Marine Park. *J. Exp. Mar. Biol. Ecol.* 302 (2), 123–152.

Cappo, M.A., et al., 2003. Potential of video techniques to monitor diversity, abundance and size of fish in studies of marine protected areas. In: *Aquatic Protected Areas—What Works Best And How Do We Know*, 1, pp. 455–464.

Cartwright, P., Johns, J., Waltham, N.J., 2024. Port of Mackay and Hay Point Ambient Marine Water Quality Monitoring Program: Annual Report. In: James Cook University: Centre for Tropical Water & Aquatic Ecosystem Research (TropWATER), p. 91.

Cheng, T., 2025. Yolo-World. Available from: <https://github.com/AI-Lab-CVC/YOLO-World>.

Cheng, T., et al., 2024. Yolo-world: Real-time open-vocabulary object detection. In: *Computer Vision And Pattern Recognition. IEEE/CVF*.

Connolly, R.M., et al., 2021. Improved accuracy for automated counting of a fish in baited underwater videos for stock assessment. *Front. Mar. Sci.* 8, 658135.

Cyrus, D., Blaber, S., 1987. The influence of turbidity on juvenile marine fishes in estuaries. Part I. Field studies at Lake St. Lucia on the southeastern coast of Africa. *J. Exp. Mar. Biol. Ecol.* 109 (1), 53–70.

Ditria, E.M., et al., 2020. *Automating the Analysis of Fish Abundance Using Object Detection: Optimizing Animal Ecology With Deep Learning.* Frontiers in Marine Science 7.

Dunlop, K.M., Marian Scott, E., Parsons, D.R., 2015. Baited cameras: A potential tool for investigating the deep-water fish assemblages of the UK continental shelf. *Deep-Sea Research Part I: Oceanographic Research Papers* 106, 123–134.

Environmental Protection (Water and Wetland Biodiversity) Policy, 2023. Parliamentary Counsel. Government, Queensland.

Firth, L.B., et al., 2024. Coastal greening of grey infrastructure: an update on the state of the art. In: *Proceedings Of The Institution Of Civil Engineers-Maritime Engineering.* Emerald Publishing Limited.

Goodwin, S.D., Nitschke, G., Taylor, M.D., 2021. AI and machine learning applications for automated fish species identification and classification. *Ecological Informatics* 61, 101225.

Gray, S.M., McKinnon, J.S., Bolnick, D.I., 2012. Colour pattern divergence in freshwater sticklebacks: A role for predation and foraging environment. *Evol. Ecol. Res.* 14 (7), 769–783.

Hannam, M., et al., 2022. A review of challenges and solutions for underwater visual fish survey methods. *Rev. Fish Biol. Fish.* 32, 831–857.

Harvey, E.S., et al., 2012. Bait attraction in remote underwater video: An experiment in different habitats. *J. Exp. Mar. Biol. Ecol.* 416, 190–195.

Harvey, E.S., et al., 2018. The influence of bait quantity on a temperate fish assemblage sampled using stereo-BRUVS. *Mar. Ecol. Prog. Ser.* 592, 213–222.

Heagney, E.C., et al., 2007. Pelagic fish assemblages assessed using mid-water baited video: standardising fish counts using bait plume size. *Mar. Ecol. Prog. Ser.* 350, 255–266.

Hernandez, F., et al., 2020. Environmental monitoring and decision making in ports and harbors: opportunities for new technologies. *Front. Mar. Sci.* 7, 571.

Jahanbakhht, M., Azghadi, M.R., Waltham, N.J., 2023. Semi-supervised and weakly-supervised deep neural networks and dataset for fish detection in turbid underwater videos. *Ecological Informatics* 78, 102303.

Johnston, R., Sheaves, M., Molony, B., 2007. Are distributions of fishes in tropical estuaries influenced by turbidity over small spatial scales? *J. Fish Biol.* 71 (3), 657–671.

Langlois, T.J., et al., 2020. The effects of depth and a marine reserve on fish assemblage composition assessed using baited remote underwater video systems. *J. Environ. Manag.* 261, 110209.

Lehtiniemi, M., 2005. Swim or hide: predator cues cause species specific reactions in young fish larvae. *J. Fish Biol.* 66 (5), 1285–1299.

Lowe, M., Morrison, M., Taylor, R., 2015. Harmful effects of sediment-induced turbidity on juvenile fish in estuaries. *Mar. Ecol. Prog. Ser.* 539, 241–254.

Lunt, J., Smeed, D.L., 2020. Turbidity alters estuarine biodiversity and species composition. *ICES J. Mar. Sci.* 77 (1), 379–387.

Malcolm, H.A., et al., 2021. Assessing the effectiveness of underwater video to monitor fish in estuarine habitats. *Estuar. Coast. Shelf Sci.* 249, 107089.

Marini, S., et al., 2018. Tracking Fish Abundance by Underwater Image Recognition. *Sci. Rep.* 8 (1), 13748.

Monkman, G.G., Hyder, K., Garcia, L., 2022. Automated methods for monitoring fish populations in aquatic ecosystems. *Front. Ecol. Environ.* 20 (8), 454–461.

Rodriguez, J.P., Cramer, L., Gomez, J., 2022. Automated species identification in marine environments: Machine learning approaches and applications. *Mar. Biol. Res.* 18 (4), 332–348.

- Saleh, A., Sheaves, M., Rahimi Azghadi, M., 2022. Computer vision and deep learning for fish classification in underwater habitats: A survey. *Fish Fish.* 23 (4), 977–999.
- Saleh, A., et al., 2024. Applications of deep learning in fish habitat monitoring: A tutorial and survey. *Expert Syst. Appl.* 238, 121841.
- Schmid, K., et al., 2017. Baited remote underwater video as a promising nondestructive tool to assess fish assemblages in clearwater Amazonian rivers: testing the effect of bait and habitat type. *Hydrobiologia* 784 (1), 93–109.
- Seiler, L.M.N., Fernandes, E.H.L., Siegle, E., 2020. Effect of wind and river discharge on water quality indicators of a coastal lagoon. *Reg. Stud. Mar. Sci.* 40, 101513.
- Spampinato, C., et al., 2010. Automatic fish classification for underwater species behavior understanding. In: *Proceedings Of The First Acm International Workshop On Analysis And Retrieval Of Tracked Events And Motion In Imagery Streams*.
- Stobart, B., et al., 2007. A baited underwater video technique to assess shallow-water Mediterranean fish assemblages: Methodological evaluation. *J. Exp. Mar. Biol. Ecol.* 345 (2), 158–174.
- Tan, M., Le, Q., 2021. Efficientnetv2: Smaller models and faster training. In: *Proceedings Of The International Conference On Machine Learning*.
- TropWATER, 2024. **Marine Management In Industries With Artificial Intelligence.** Available from. https://www.linkedin.com/posts/tropwater_biodiversity-data-quality-activity-7047437054045011968-yRPr/.
- Verschuur, J., Koks, E., Hall, J., 2022. *Ports' criticality in international trade and global supply-chains.* *Nature. Communications* 13 (1).
- VITB, 2025. **Octopus Flagship Self-Cleaning PTZ Vision System.** Available from. <https://www.viewintothelblue.com/products/octopus>.
- Waltham, N.J., Sheaves, M., 2015. Expanding coastal urban and industrial seascape in the Great Barrier Reef World Heritage Area: Critical need for coordinated planning and policy. *Mar. Policy* 57, 78–84.
- Waltham, N.J., et al., 2023. Remotely operated vehicle reveals fish orientate to the substrate underneath marina floating pontoons. *Estuar. Coast. Shelf Sci.* 280, 108184.
- Wenger, A.S., Harvey, E., Wilson, S., 2013. Suspended sediment impacts on fish: A global review. *Aquat. Sci.* 75 (4), 477–480.
- Whitfield, A., Elliott, M., 2002. Fishes as indicators of environmental and ecological changes within estuaries: a review of progress and some suggestions for the future. *J. Fish Biol.* 61, 229–250.
- Whitfield, A.K., 2021. Estuaries—how challenging are these constantly changing aquatic environments for associated fish species? *Environ. Biol. Fish* 104 (4), 517–528.
- Whitmarsh, S.K., Fairweather, P.G., Huvneers, C., 2017. What is Big BRUVver up to? Methods and uses of baited underwater video. *Rev. Fish Biol. Fish.* 27, 53–73.
- Williams, K., 2014. Acoustic tracking of aquatic animals: scale, design and deployment of listening arrays. *Mar. Ecol. Prog. Ser.* 492, 1–17.
- Williams, M.J.P., et al., 2021. The heat is on: Gulf of Alaska Pacific cod and climate-ready fisheries. *ICES J. Mar. Sci.* 79 (2), 573–583.