Automated object detection based on for monitoring benthic population dynamics: A new approach combining photogrammetry and opensource tools applied to sea cucumbers

Remote Sensing in Ecology and Conservation
Open Access

RESEARCH ARTICLE

Automated object detection based on YOLOv11 for monitoring benthic population dynamics: A new approach combining photogrammetry and open-source GIS tools applied to sea cucumbers

Gian Mario Sangiovanni[1], Giovanna Jona Lasinio[1], Daniele Poggio[2], Gianluca Mastrantonio[2], Alessio Pollice[3], Arnold Rakaj[4], Midhun Mohan[5], Stefano Moro[6], Edoardo Casoli[7] & Daniele Ventura[7]

[1]Department of Statistical Sciences, University of Rome 'La Sapienza', Rome, Italy
[2]Department of Mathematical Sciences 'G. L. Lagrange', Politecnico di Torino, Torino, Italy
[3]Department of Economics and Finance, Università degli Studi di Bari Aldo Moro, Bari, Italy
[4]Experimental Ecology and Aquaculture Laboratory, Department of Biology, University of Rome Tor Vergata, Rome, Italy
[5]Ecoresolve, San Francisco, CA, USA
[6]Department of Integrative Marine Ecology (EMI), Stazione Zoologica Anton Dohrn, Rome, Italy
[7]Department of Environmental Biology, University of Rome 'La Sapienza', Rome, Italy

Keywords
image augmentation, machine learning, Mediterranean Sea, structure from motion (SfM)

Correspondence
Daniele Ventura, Department of Environmental Biology, University of Rome 'La Sapienza', V. le dell'Università 32, 19 00185 Rome, Italy. E-mail: daniele.ventura@uniroma1.it

Funding Information
This work was partially funded by the Pure Ocean Fund (3DR-4-Seac research grant) and PRIN 2002 (Project #20224LYJM5). The involvement of Daniele Poggio in formal data analysis and presentation was made possible through the PNRR-NGEU project, which received financial support from the MUR following DM 630/2024. We extend our gratitude to Dr. Przemysław Aszkowski for his contributions in implementing significant features within the final workflow. Additionally, we are grateful to the anonymous reviewers and associate editor whose constructive feedback played a crucial role in refining and enhancing the final version of this manuscript.

Editor: Dr. Kylie Scales
Associate Editor: Dr. Jacquomo Monk

Abstract

Benthic organisms, such as sea cucumbers, play a crucial role in delivering essential ecosystem services, including nutrient recycling through feeding, excretion and bioturbation processes. Considering recent harvest activities in subtidal zones worldwide, it is imperative to define rapid and cost-effective solutions to improve monitoring tools. Large-scale Structure from Motion (SfM) photogrammetric mapping was conducted along Mediterranean infralittoral sea bottoms to provide training and testing datasets for developing a new object detection model. Several high-resolution (0.5 cm/pixel) orthophoto mosaics were used to train a new object detection model based on the YOLOv11 architecture. Subsequently, georeferenced imagery was used to assess the accuracy of manual counts compared to the outcomes produced by the model. The model exhibited strong performance on the validation set, achieving a precision of 83% and a recall of 88%. The implementation of the open-source GIS model through the Deepness plugin in QGIS software enabled the rapid deployment of the model (15.9 ± 1.35 min) over large ($5500\ m^2$) orthophoto mosaics, revealing the effects of habitat and seasonality on the detection results. The length of the diagonal of the bounding boxes used as a proxy for the body length of sea cucumbers, combined with a Gaussian kernel density estimator analysis, revealed comparable results between the estimated and observed values of sea cucumbers, confirming the validity of the proposed method for extracting biologically relevant information for natural population monitoring. This research presented a new and scalable workflow for identifying small benthic organisms through underwater photogrammetric-based imagery. The findings underscore the promise of using cost-effective and open-source object detection tools to provide a valuable tool for ecological assessments of natural populations. Data from automatic detection systems represent a promising non-invasive approach for directly analysing species distribution patterns based solely on image data, marking a significant improvement in ecological monitoring techniques.

Received: 26 May 2025; Revised: 19 November 2025; Accepted: 2 December 2025

doi: 10.1002/rse2.70051

Introduction

Sea cucumbers (phylum Echinodermata, class Holothuroidea) are marine benthic invertebrates essential in temperate and tropical waters (Purcell et al., 2016). They are extensively distributed across the Mediterranean Sea (Rakaj & Fianchini, 2024), the Atlantic Ocean (Félix et al., 2024), and various reef ecosystems (Purcell et al., 2013). As epibenthic deposit feeders, these organisms can selectively consume surface sediments, feeding on non-living fine particles and their microorganisms, including bacteria, protozoans, and diatoms (Hammond et al., 2020). This feeding behaviour plays a crucial role in the recycling of detritus on the seabed, thereby influencing biotic interactions within soft-bottom communities and improving oxygen levels in the sediment through the process of bioturbation (Schneider et al., 2013). The digestion of organic matter, which includes ingested coral sand and rubble, results in the dissolution of acidic calcium carbonate particles, thereby increasing local alkalinity in reef habitats (Schneider et al., 2013). Besides providing essential ecosystem services, sea cucumbers are a significant source of food and medicinal products in Asian and Middle Eastern nations (Wolfe & Byrne, 2017). These organisms are harvested primarily to produce processed forms (e.g., salted or dried) known as beach-de-mer (Hamel et al., 2022). Consequently, approximately 70% of the tropical sea cucumber species are classified as exploited or overexploited, with numerous local populations being entirely depleted (Purcell et al., 2014). In recent years, sea cucumber fishing has expanded rapidly in the Atlantic and Mediterranean regions, where they are used both as bait for fishing and as a food source (Félix et al., 2021; Azevedo e Silva et al., 2023). Furthermore, the implementation of integrated multi-trophic aquaculture (IMTA) systems, which incorporate deposit feeder organisms to mitigate organic matter accumulation from fish farming, has sparked increased commercial interest in temperate sea cucumber species (Grosso et al., 2021). Regrettably, adequate management measures for the sustainable fishing of these target species remain limited or entirely lacking along the Mediterranean coasts (Pasquini et al., 2022).

To establish sustainable harvest levels, stakeholders and decision-makers must conduct ecological assessments examining the spatio-temporal dynamics of natural populations. This evaluation should integrate data from various sources, including fisheries-independent surveys (Kilfoil et al., 2020). To date, most fisheries-independent surveys have relied on underwater visual censuses (UVCs), which yield detailed estimates based on counts obtained from transects conducted by SCUBA divers or snorkelers (Woo et al., 2013). However, these methods are often time-consuming and susceptible to observer bias and are affected by limited spatial coverage and significant operational and logistical costs (Prescott et al., 2013).

In recent years, considerable emphasis has been placed on using remotely operated vehicles (ROVs) and uncrewed aerial vehicles (UAVs) as substitutes for traditional sampling methods. Although using ROVs (Baharin & Kamarudin, 2025; Guo et al., 2019) allows the exploration of deep waters, their usage is restricted to relatively small areas due to limitations in deployment and operational costs. UAV-based mapping methods (Kilfoil et al., 2020; Li et al., 2021; Williamson et al., 2021) demonstrate considerable effectiveness primarily in open, sandy, shallow-water reef environments, becoming less appropriate for deeper infralittoral (1–50 m) communities present in the Mediterranean, typically marked by heterogeneous sea bottoms characterised by diverse substrates that include a mix of rocky, sandy and seagrass habitats.

To address this aspect and provide ecological information on population dynamics and habitat preferences for Mediterranean sea cucumbers, a Structure from Motion (SfM) photogrammetric application based on direct acquisition of underwater imagery using Diver Propulsion Vehicles (DPV) has been recently proposed (Ventura et al., 2025). However, the labor-intensive and time-consuming nature of manually detecting sea cucumbers in Geographic Information System (GIS) environments, which requires skilled operators for imagery analysis, renders such applications impractical for monitoring programs across extensive regions.

Although recent advances in deep learning (DL) have offered innovative technical solutions that aim to improve automation levels and reduce production costs within the aquaculture sector (Ge et al., 2024), applications for processing large datasets from natural environments are still limited (Chu et al., 2025; Wang et al., 2023). The use of imagery for detecting sea cucumbers and, in general, tiny benthic organisms in their natural environment presents significant challenges, primarily due to complex lighting conditions, water turbidity, diverse body shapes and the numerous confounding features present in the underwater environment, such as aquatic vegetation, other marine organisms and anthropogenic waste (Zhai et al., 2022). DL methods, such as Convolutional Neural Network (CNN)-based identification technology, represent a recent approach to image processing and data analysis, characterised by its impressive accuracy and significant potential, which is increasingly being adopted in various fields of ecology (Zhai et al., 2022). The conventional object detection technique known as Faster R-CNN divides the detection process into two separate stages: classification and regression. On the contrary, You Only Look Once (YOLO) (Redmon et al., 2016) adopts a more efficient methodology by tackling the regression challenge in a unified step, allowing image analysis to be significantly faster. The YOLO models are specifically designed for encoding image data, producing positional and categorical information alongside simultaneous bounding boxes. This aspect makes YOLO particularly effective for a wide range of detection tasks, including sea cucumber identification in natural marine habitats (Chu et al., 2025; Ge et al., 2024). Object detection underwent a paradigm shift with the introduction of YOLO, which pioneered a unified approach that combined object localisation and classification into a single neural network (Khanam & Hussain, 2024). This innovation streamlined the detection process compared to traditional multistage DL methods.

This research investigated the potential of advanced yet user-friendly DL methods to improve the effectiveness of image-based population assessments for benthic marine species. During several campaigns conducted around the coasts of Giglio Island, which featured complex seabed communities typical of Mediterranean coastal waters, an automated detector based on the YOLOv11 architecture was used to identify and count sea cucumbers in high-resolution imagery captured by underwater photogrammetry. To the best of our knowledge, this represents the first application for sea cucumber detection based on large georeferenced orthophoto mosaics used for training and validation. This approach can demonstrate the broad applicability of photogrammetrically derived data collection in conjunction with GIS and YOLO technologies for monitoring marine benthic populations. In addition, three key questions were formulated to gain a greater insight into the model's performance: (i) Did the object detector accurately predict the presence of sea cucumbers on the seafloor? (ii) Can ecological traits such as body length and densities of the observed (i.e. derived from manual count) and estimated (i.e., derived from object detection) processes be compared? (iii) What factors, including habitat type, might influence any discrepancies between them?

By responding to these inquiries, we provide an efficient method to enhance the detection of subtidal sea cucumbers, addressing the challenges associated with time-consuming manual counts in high-resolution imagery. By integrating photogrammetric acquisition and YOLO in GIS environments, the model achieves high accuracy, enabling its validation and comparison with biologically meaningful information derived by expert observers. This workflow is flexible enough to be applied to other subtidal species and ecosystems, presenting a scalable solution for enhancing ecological surveys in complex environments and ultimately delivering significant assistance for studies on population dynamics and conservation efforts.

Materials and Methods

Study area

The research site was located along the northeast coast of Giglio Island (Tyrrhenian Sea, Italy) close to Giglio harbour, specifically at the affected location known as Punta Lazzaretto (N:42.364867°; E:10.920210°), which was affected by the Costa Concordia shipwreck (Fig. 1). This area, covering approximately 5500 m^2, featured complex infralittoral seabeds with depths ranging from 8 to 27 m. The region has experienced multiple disturbances due to the presence of the wreck and subsequent removal operations (Mancini et al., 2019). Specifically, the shadowing effect produced by different structures (such as wrecks and operational vessels) along with the sedimentation resulting from scraping activities for platform installation has significantly harmed the natural *P. oceanica* meadow, resulting in its decline (Mancini et al., 2021). After the restoration of the natural environmental conditions post the completion of the wreck removal operation in July 2014 and the seabed cleaning phase in April 2018, the transplanting of *P. oceanica* fragments commenced in 2019 at the site. Currently, the area exhibits a mosaic of habitats mainly composed of sandy regions interspersed with rocky substrates (granitic boulders and pebbles covered by photophilic algae). Additionally, both natural (high-density) and transplanted (low-density) *P. oceanica* beds, characterised by discontinuous patches of dead *P. oceanica* rhizomes with trapped sediments, referred to as 'matte', are present. These habitats were widely spread across the entire Mediterranean basin and are indicative of the presence of sea cucumbers, thereby facilitating a broad applicability of the testing phase for the detection model we have proposed.

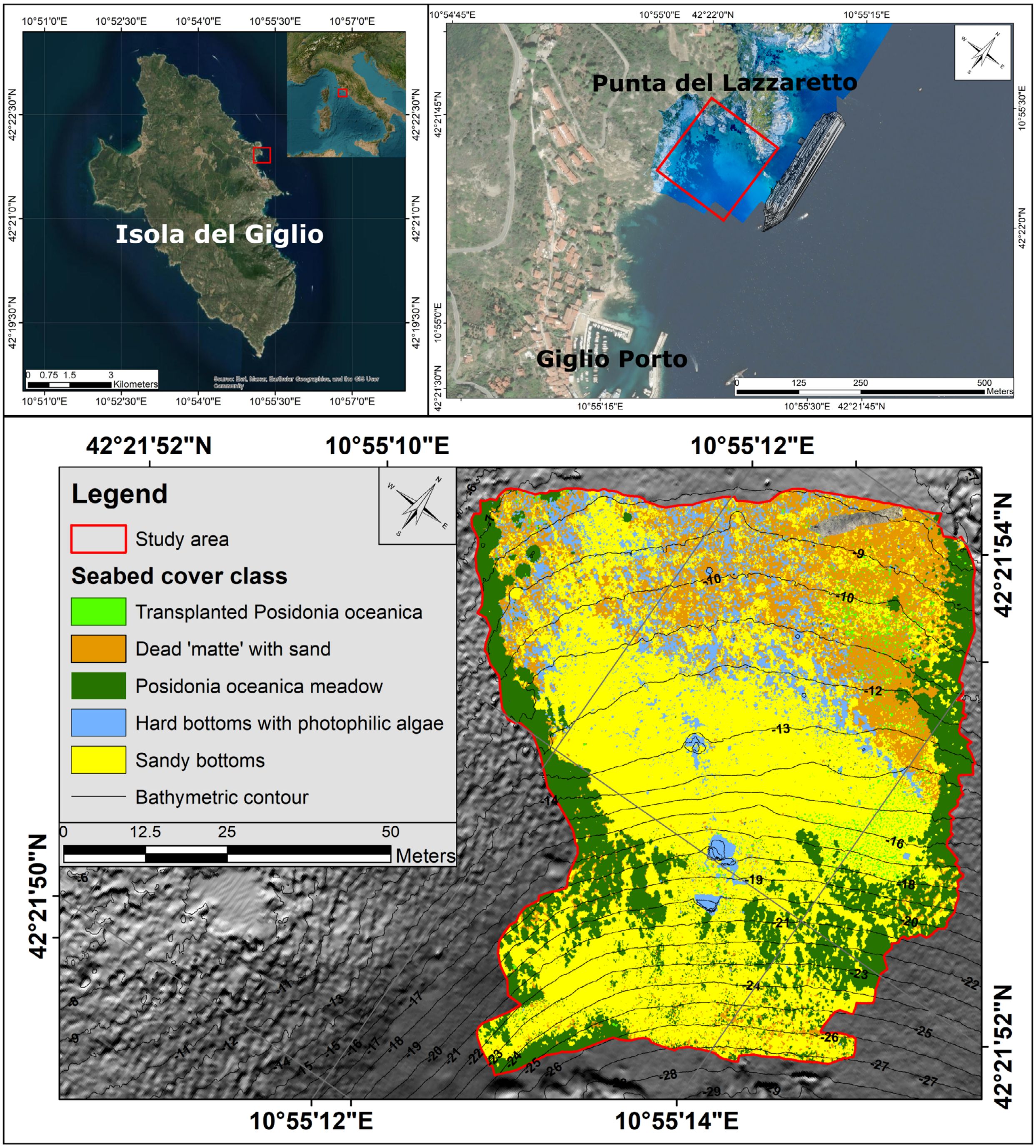

Figure 1. Map of the study area along the NE coast of Giglio Island (Tuscany archipelago, Italy). The red polygon indicates the mapped area where the automated sea cucumber detection workflow was tested. A detailed orthophoto of an uncrewed aircraft vehicle (UAV) is draped over satellite base imagery (*Source*: Esri, DigitalGlobe) and a hill shade relief map derived from multibeam echosounder data is displayed to improve the visualisation of coastal and seabed features, respectively. 1-m bathymetric contours and main seabed cover classes (adopted from Ventura et al., 2025) were shown.

Underwater imagery collection and cartographic output generation

Five underwater photogrammetric surveys were conducted in the region between January and November 2022, following the established protocol for seagrass transplantation mapping (Ventura et al., 2022) and to determine the ecological preferences of sea cucumbers within the study area (Ventura et al., 2025). The methods involved defining an underwater regular mapping grid, which enabled the acquisition of HD (5568×4176 pixels) images using a consumer-grade action camera (GoPro Hero 10). To cover the entire area while keeping the permanence of the diver limited as much as possible, the camera was mounted on a diver propulsion vehicle (DPV). The diver operated the DPV at an approximate distance of 5 m from the seabed to ensure an overlap of 75% between consecutive and adjacent images, thereby achieving a ground sample distance of 3 mm per pixel, which was sufficient for the accurate identification of sea cucumbers and other benthic features. Six ground control points (GCPs), with coordinates determined from the surface using the Global Positioning System (GPS), were placed throughout the area to enhance the accuracy of georeferencing and scaling of the photogrammetric results (Ventura et al., 2025). Imagery processing was achieved in Agisoft Metashape v. 1.8.2 (Agisoft LLC, Russia), a low-cost platform for 3D reconstructions using SfM algorithms. Sparse and dense point clouds, digital surface models (DSMs) and red-green-blue (RGB) orthophoto mosaics were obtained by aligning digital photos derived from each mapping mission. On these high spatial resolution ($\sim$0.5 cm/pixel) cartographic products, an object-based image analysis (OBIA) routine was applied to determine the main seabed cover classes (Ventura et al., 2025). The RGB orthophoto mosaics and OBIA classification were stored as raster layers as 32-bit GeoTIFF (*.tif) files for further GIS implementation.

Dataset generation and manual labelling of sea cucumbers

The creation of the data set with the locations of labelled sea cucumbers was based on twelve orthophoto mosaics obtained in 2021 from sites outside the study area, specifically from other transplantation sites or natural environments, located along the northeast coast of the island. These areas encompass a plethora of habitats, including bare and vegetated rocks with photophilic algal cover, sandy, muddy and pebbly bottoms, as well as seagrass meadows formed by different species (*Posidonia oceanica* and *Cymodocea nodosa*) with variable cover and shoot density (from dense to low-density meadows). All these habitat types are commonly found along the shallow waters (1–20 m in depth) of the Mediterranean coasts, ensuring the scalability of the trained model along different geographic regions of the basin. To enhance the generalisability of the training dataset, we used photogrammetric orthophoto mosaics generated using different equipment (i.e., images acquired by consumer-grade action cameras and professional full-frame cameras) and environmental conditions (i.e., summer-winter acquisition with variable lighting and turbidity). Naturally, to ensure good results of the photogrammetric acquisition, surveys were not performed under conditions of poor visibility (water visibility <5 meters). The original orthophotos acquired in 2021 were cropped to a standardised size of 640×640 pixels, with a resolution of 0.5 cm per pixel, using the plugin Deepness (Aszkowski et al., 2023) for QGIS software v3.34 (Fig. 2). This pre-processing tiling procedure was motivated by three key considerations: (i) processing large, high-resolution images was computationally intensive, (ii) sea cucumbers were relatively small objects compared to the seafloor background, making their detection challenging and (iii) the use of orthophoto mosaics guaranteed a more uniform resolution of the raster images due to the photogrammetric processing compared to original photos. Research has shown that image tiling can significantly improve the detection of small objects (Ozge Unel et al., 2019). Dividing the images into even smaller tiles that match the size of individual sea cucumbers was tested, but this approach did not produce satisfactory results. A total of 3609 tiles were initially considered, of which 2292 tiles containing at least one sea cucumber were selected for the labelling process.

Furthermore, 200 tiles that showed only a seabed without sea cucumbers (negative examples) were included. This inclusion aimed to improve the model's generalisation capabilities and ensure that the object detection model could effectively manage various underwater scenarios, such as open sandy areas, patchy seagrass habitats and continuous meadows. Tiles that were blurred and of low quality were excluded. Subsequently, two independent reviewers labelled the 2.492 tiles using the Roboflow annotation toolkit from Roboflow (Alexandrova et al., 2015) to conform to the Ultralytics image annotation format. Following standard machine learning practices, the dataset was partitioned using the 70–20–10 rule into training sets (1880 images), validation sets (368 images) and test sets (244 images).

YOLOv11 general outlook

YOLO, a widely recognised single-shot detector, has seen advancements in its recent iteration, YOLOv11, which

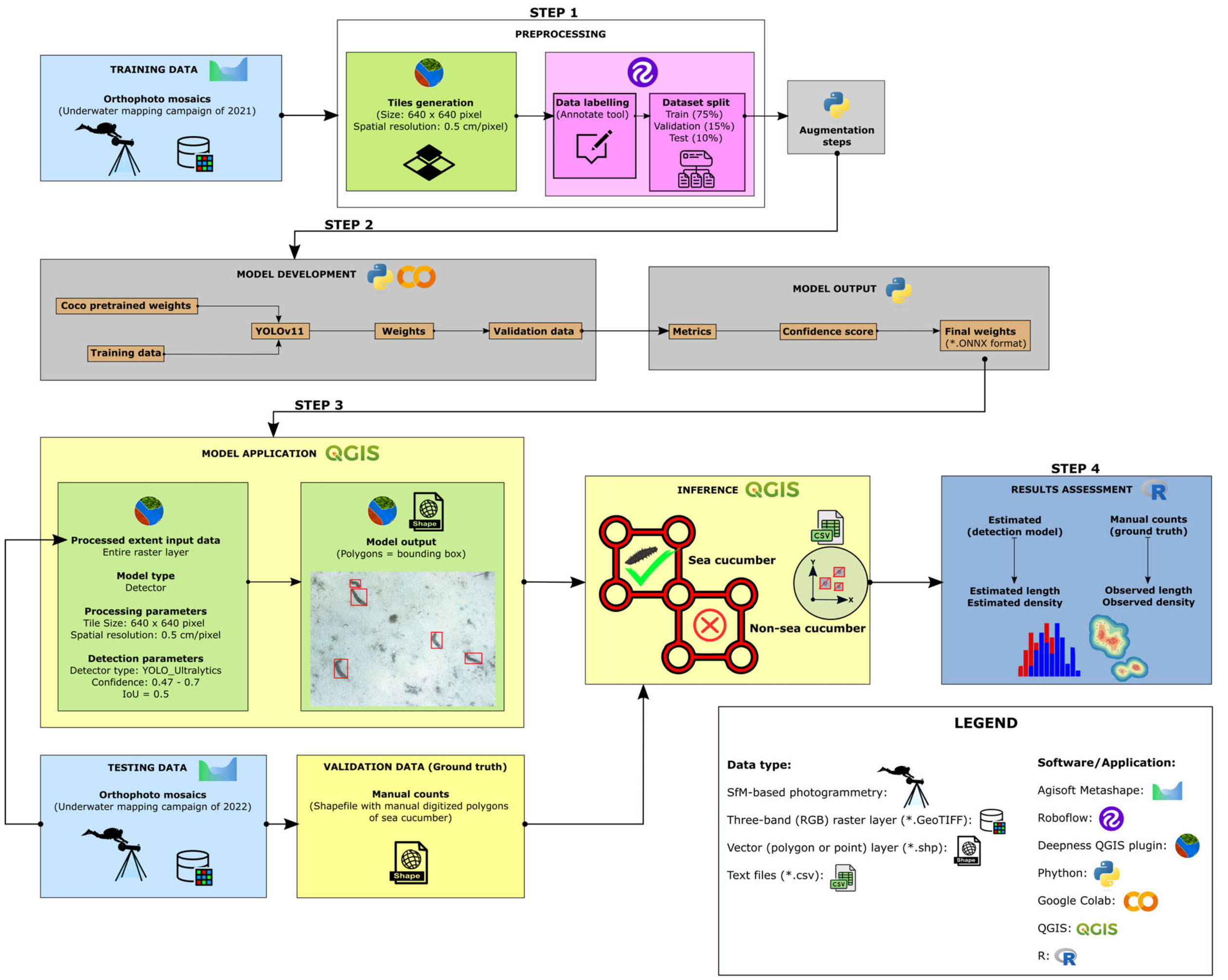

Figure 2. Workflow of the proposed method: The icons and colours refer to the software used.

improves computational efficiency and accuracy across various object sizes. Specifically, enhancements in architecture and optimisation of parameters have led to reports of improved detection performance (Khanam & Hussain, 2024). YOLOv11 processed an image, which was represented as a tensor with dimensions $h \times w \times c$, along with the coordinates of the vertices that outline the bounding box surrounding the object of the specified class. The image in the given scenario was the three (RGB) channels tile (640×640 pixels), with one class constituted by labelled sea cucumbers. Different versions of the YOLOv11 model existed, distinguished by the number of parameters involved (i.e. the number of layers). Consequently, weights were learned by minimising a loss function specific to the architecture. Without loss of generality, the loss in every YOLOv11 model consisted of three main components. First, the bounding box regression loss ($\mathfrak{L}_{box}$) measured the accuracy of the predicted bounding boxes compared to the ground truth, improving the model's precision. Second, the class probability loss ($\mathfrak{L}_{cls}$), evaluated the difference between the predicted class probabilities and the ground truth, ensuring accurate classification. Third, the distributed focal loss ($\mathfrak{L}_{dfl}$), extended the original focal loss by prioritising detecting hard-to-classify objects and assigning lower weights to those that are easier to classify. Total loss (Equation 1)

$$\mathfrak{L}_{tot} = \mathfrak{L}_{\mathfrak{box}} + \mathfrak{L}_{\mathfrak{cls}} + \mathfrak{L}_{\mathfrak{dfl}} \tag{1}$$

can be minimised during training using gradient-based optimisation techniques over multiple epochs to refine the model's weights and improve detection accuracy progressively.

Model training and validation

For the implementation, the medium version of YOLOv11 (YOLOv11m) was chosen as the object detection model due to its balance between the number of parameters and the accuracy in general detection tasks. During the training phase, computations were performed on an NVIDIA Tesla T4 GPU (15 GB VRAM) with driver version 550.54.15 and CUDA 12.4 (utilising the Google Colab free plan). During execution, the GPU operated in default compute mode with a maximum power draw of 70 W. The weights have been initialised considering the nano version of YOLOv11 (YOLOv11n), pre-trained on the Coco (Common Objects in Context) dataset. This approach was chosen based on established research showing that transfer learning techniques significantly improve both generalisation capabilities and convergence rates of model weights (Öztürk et al., 2023).

Moreover, to enhance the generalisation and robustness of the model, a range of data augmentation methods, adhering to the standard configurations specified in the YOLO documentation on the Ulralytics web page (https://docs.ultralytics.com/guides/yolo-data-augmentation/) were implemented on the cropped dataset (Table 1 and Fig. 3). This approach ensured replicability and effectively tackled issues related to diverse seabed textures, lighting variations and the orientations of sea cucumbers (Zoph et al., 2020). The augmentation techniques have been applied online, ensuring that the original dataset remained unaltered and each training batch received a dynamically augmented data version, thereby enhancing model adaptability without increasing the dataset size. The model was trained for 90 epochs, with early stopping implemented to prevent overfitting, and all other hyperparameters were set to their default values (Table 2). The final epochs of the training process were conducted without enhancing mosaic data. This phase allowed the model to fine-tune its weights on actual data, reinforcing its ability to generalise to real-world conditions without relying on artificially modified inputs.

Establishing a suitable decision threshold during the validation phase is crucial for accurate prediction. To define this threshold, called confidence score or box confidence in object detection, which indicated the likelihood that a predicted bounding box encompasses an object and the intersection over union (IoU) between the detected box and the actual ground truth, predictions were produced on a range of confidence scores, from 0 to 1, to facilitate the creation of a confusion matrix which offered a comprehensive analysis of the predictions and allowed the derivation of several related metrics based on established confidence and IoU thresholds.

Therefore, the precision (P), which was the ratio of true positives (correctly detected objects, TP) to the total number of predicted positives (TP + FP) and the recall (R), namely the proportion of correctly identified objects (TP) out of all actual objects (TP + FN), were computed. A high recall value indicated a lower number of false negatives, whereas a high precision indicated a lower number of false positives.

The precision-recall curve illustrates the trade-off between precision and recall, which is similar to a receiver operating characteristic (ROC) curve. A high area under the curve (AUC) indicated that a suitable confidence threshold was found. Furthermore, the model's average precision (AP) has been computed as the area under the precision-recall curve. AP summarised the model performance by averaging the precision values at different recall levels. The mean average precision (mAP) is typically reported as an overall performance measure, averaging the AP across the various classes. However, since the data set contained only one class, AP was equivalent to mAP.

Table 1. Main augmentation steps applied during model development.

Augmentation methods	Type	Settings	Aim
Scaling	Geometric	Scaling factor = 0.5	To diversify object sizes
Flipping	Geometric	Vertical flip probability = 0.1, horizontal flip probability = 0.5	To simulate different sea cucumber orientations
Translation	Geometric	Random translation probability = 0.2	To introduce spatial variability
Mosaic	Composite image creation	Probability = 0.5 of combining four images into a single composite image. For the last 10 epochs this option was disabled.	To enhance contextual diversity
HSV	Colour adjustments	Hue = 0.015, saturation = 0.7 and value (brightness) = 0.3	To account for underwater colour variation
Gaussian Noise	Blur and noise jitter	Introduction of noise and blur	To enhance model robustness against possible natural distortions

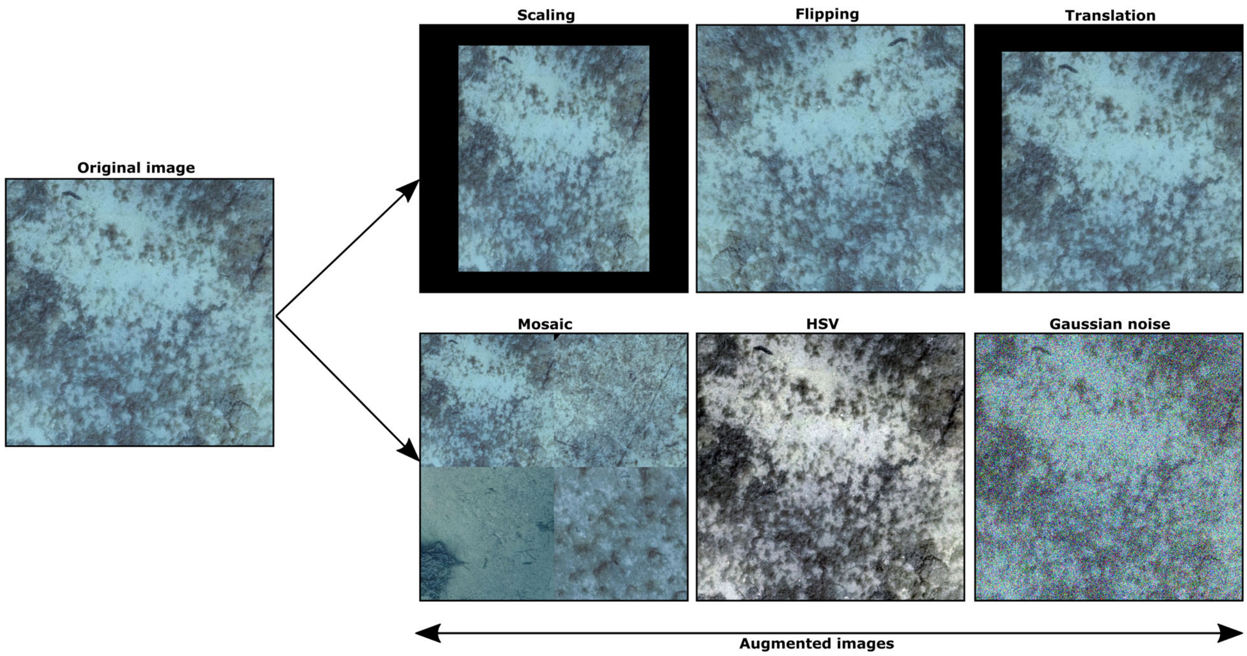

Figure 3. Example of the augmentation steps applied to original images during model development.

Table 2. Hyperparameters used during the training phase.

Hyperparameter	Value	Hyperparameter	Value
imgsz	640	Momentum	0.937
lr0	0.1	lrf	0.01
Single_cls	True	Batch	16
optimiser	ADAM	Epochs	90

Lastly, to determine the optimal value for the confidence threshold, the F_β-Score is computed, which generalises the standard F_1-Score. Formally, it is defined as:

$$F_\beta = \frac{\left(1+\beta^2\right)(P+R)}{\left(\beta^2 * P\right)+R} \qquad (2)$$

The parameter β determines the relative importance of recall compared to precision, and therefore, different values are used. The value that maximises the F_β-Score (the argmax of the function) represents the optimal confidence threshold that achieves the most balanced trade-off between recall and precision. To facilitate scientific reproducibility and collaborative development, we provide a GitHub repository containing the code used to train the model, along with the final trained weights (https://github.com/justgimmi/Sea-Cucumbers-Object-Detection). The trained weights are included in both ONNX format and YOLO-compatible format, allowing them to serve as a starting point for fine-tuning. Moreover, the dataset used for model training is freely available at https://universe.roboflow.com/olo-dan/sea-cucumbers-new-tiles.

GIS model implementation and comparison between model detection results and manual counts

Manual counts of sea cucumbers were performed to evaluate the model's performance using the five orthophoto mosaics obtained in 2022 for the study area (Fig. 1). These mosaics were used solely to assess the final object detection model and were not involved in the tiling process necessary for training the model. Two independent users carried out a visual inspection and manual polygonisation in QGIS using the 'Add Polygon Feature' in the editing toolbar. The inter-observer agreement was evaluated using the intra-class correlation coefficient (ICC) (Shrout & Fleiss, 1979), implemented in the 'IRR' R package. It provided a single value estimate that reflected the agreement and consistency within counts. The body length of the sea cucumbers was measured and added to the attribute table as a new field. The results obtained from each observer, stored as geospatial vector shapefiles (*.shp), were compared to ensure that only the polygons representing the sea cucumbers were retained. The Deepness plugin was used to implement automated object detection in QGIS by running the model in ONNX format (*.onnx). The entire raster dataset (i.e., the orthophoto mosaics of 2022) was used as the input file. The processing parameters, specifically resolution (cm/pix) and tile size (pixels), should correspond to the values used during the training phase, which were 0.5 cm/pix and 640×640 pixels, respectively. The detection parameters encompassed the specification of the detector type as 'YOLO_Ultralytics' to align with the model's format. At the same time, the confidence score and the IoU threshold were set at 0.47 and 0.5, respectively. Multiple confidence scores were evaluated, starting with a balanced approach giving equal weight to recall and precision, and then considering scenarios where recall was weighted less than precision. The tested values ranged from 0.47 to 0.79, corresponding to those that maximised the F_β-Score for different β values. A standard threshold used is 0.5, meaning that a predicted box must have an IoU of at least 0.5 with a ground truth box to be considered a true positive detection.

The Deepness plugin provided the outcomes of object detection in the form of a polygon shapefile, reporting the confidence level for each bounding box (Fig. S1), which was helpful for a first assessment of accuracy. To precisely identify the correct (matching) detection (i.e., the model predicted a bounding box that corresponded to a sea cucumber identified from the manual count), false positive (FP) (i.e., the model predicted a bounding box but no correspondence is found within the manual detection) and false negative (FN) or missing predictions (i.e., a sea cucumber was present but was not recognised by the model), a polygon overlay analysis through spatial join based on centroids location was performed. Centroids were used instead of the original polygons to prevent multiple matching due to the clustering of sea cucumbers. True negatives were excluded from this analysis, which is a standard practice in object detection tasks that focus on identifying the presence of objects rather than their absence. Buffer polygons encompassing an area of $0.25\,m^2$ were established around each centroid of the bounding box produced in Deepness to assess the habitat's effects on detection outcomes. Considering the limited movement speed of sea cucumbers and the duration of photogrammetric acquisition (approximately 70 min to map the area), the chosen buffer area ensured a satisfactory representation of the habitat association. Within each buffer, the 'zonal histogram' tool was used to quantify the number of pixels associated with each cover class, converting these data into cover percentages (Fig. S2).

Given that the detections were represented as rectangular shapes surrounding each sea cucumber, the length in cm of the diagonal, estimated using the formula $d = \sqrt{l^2 + w^2}$, where the length (l) and width (w) of each shape were evaluated by the oriented minimum bounding box tool available in the vector geometry toolbox of QGIS and used as a proxy of the actual length manually measured for each specimen detected. Only matching detection was used, and the variability in body length between manual and automatic detection processes was assessed by comparing linear regression models and histograms.

A Gaussian kernel density estimator (KDE) was implemented to analyse the spatial density patterns of both processes. The bandwidth parameter, which controls the smoothness of the resulting density estimate, was selected through cross-validation, and Jones–Diggle improved edge correction, available in the Spatstat package (Baddeley & Turner, 2014), was applied. This edge correction method reduces estimation bias near the study area boundaries by adjusting the kernel weights to account for the proportion of the kernel that extends beyond the study region. The Gaussian KDE with edge correction provides a robust estimate of the expected number of points per unit area at any location within the study region, a concept similar to the intensity of a point pattern.

Results

Model performances

The model's performance was summarised by its comprehensive loss values during the validation phases, as well as several validation metrics previously described (Fig. 4).

A threshold of 0.47 was selected to maintain equal importance between precision and recall, as it demonstrated superior performance during the validation phase. The performance of the model was revealed through an analysis of the F_β-Score and precision-recall curve (Fig. 5), other than visualising its predictive capabilities on raw imagery (Fig. S3).

Confusion matrices showed that the model exhibited strong performance on the validation set, achieving a precision of 83% and a recall of 88% (Fig. 6). The F1 score stabilised around 86%. At the same time, the mean mAP at the IoU threshold of 0.5 reached 0.91. Initial testing on a small preliminary test set yielded promising results, with the model achieving 75% precision, 92% recall and an F1 score of 82%. These results suggested that even a relatively small number of images was sufficient to develop a robust and generalised model.

Comparison and effectiveness of model estimations versus manual counts

The model required 23 milliseconds to process each image during the inference. In the study area, which comprises approximately $5500\,m^2$, the cumulative processing duration across the five campaigns was 15.9 ± 1.35 min,

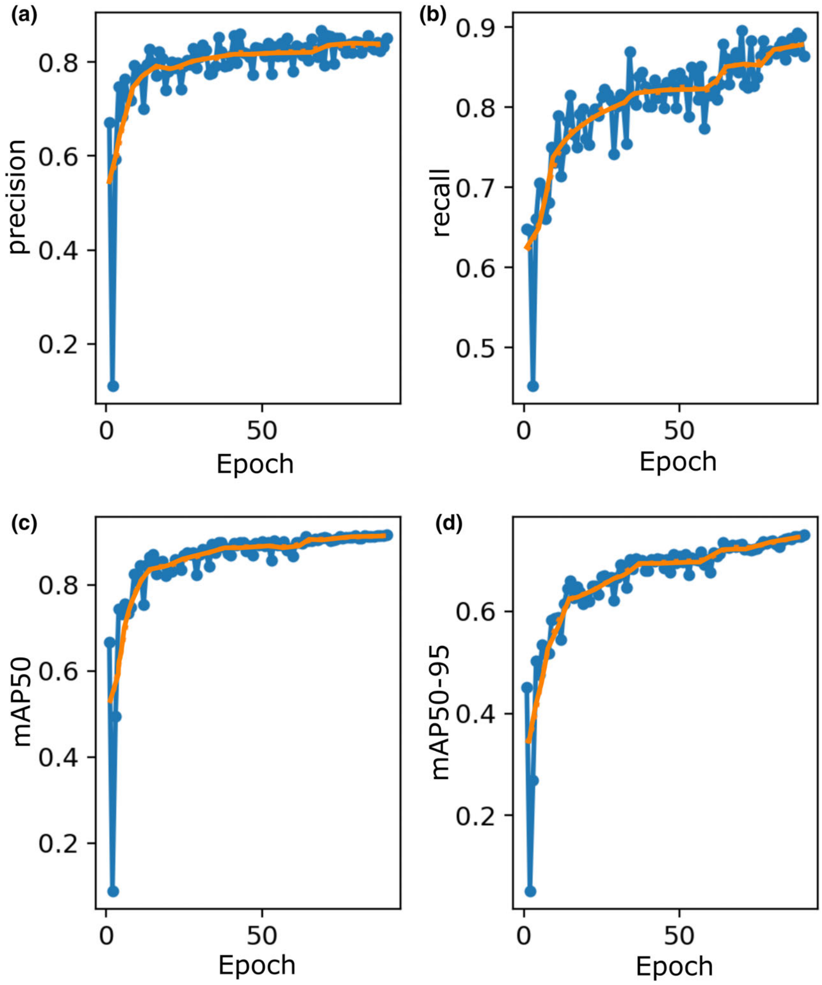

Figure 4. Validation performance of the YOLOv11m model over 90 epochs. Validation metrics including precision (a), recall (b) and mAP values (c, d) at different IoU thresholds for the sea cucumber class. The blue line represents the raw results, and the orange dotted line represents the smoothed trend.

using an ASUS ROG G703 workstation equipped with an Intel Core i7-7700HQ CPU, 48 GB of RAM and an Nvidia GeForce GTX 1080 graphics card. This processing time was significantly lower than the manual count phase, which extended for approximately two weeks. The evaluation of accuracy (Table 3), which relied on manual counts indicating complete concordance among the observers (ICC = 1; $P < 0.001$), served as the ground truth, while the automated counts were treated as estimates. This assessment revealed consistently high precision across all campaigns, despite the recall values being relatively low (however, obtaining good F_1 values). The analysis revealed a seasonal pattern in the detection performance, with the model achieving its best results during the spring and summer months while showing reduced accuracy during fall.

The variability in recall rates across five distinct benthic habitat types during the sampling campaigns indicated significant seasonal fluctuations in model performance, with recall values ranging from 0.13 to 0.74 (Fig. 7). Each habitat type exhibited unique structural and spectral characteristics that affected the detectability and classification performance of sea cucumbers (Fig. S4). Generally, open areas clear of vegetation, such as sandy patches and dead 'matte', consistently demonstrated high classification accuracy, achieving peak recall during the winter and spring months, maintaining values above 0.59, except for the survey conducted in October. In contrast, natural *Posidonia oceanica* and transplanted *Posidonia* displayed notable seasonal variation due to the behavior of specimens that were frequently observed in close proximity to the plant's rhizomes, which contributed to increased

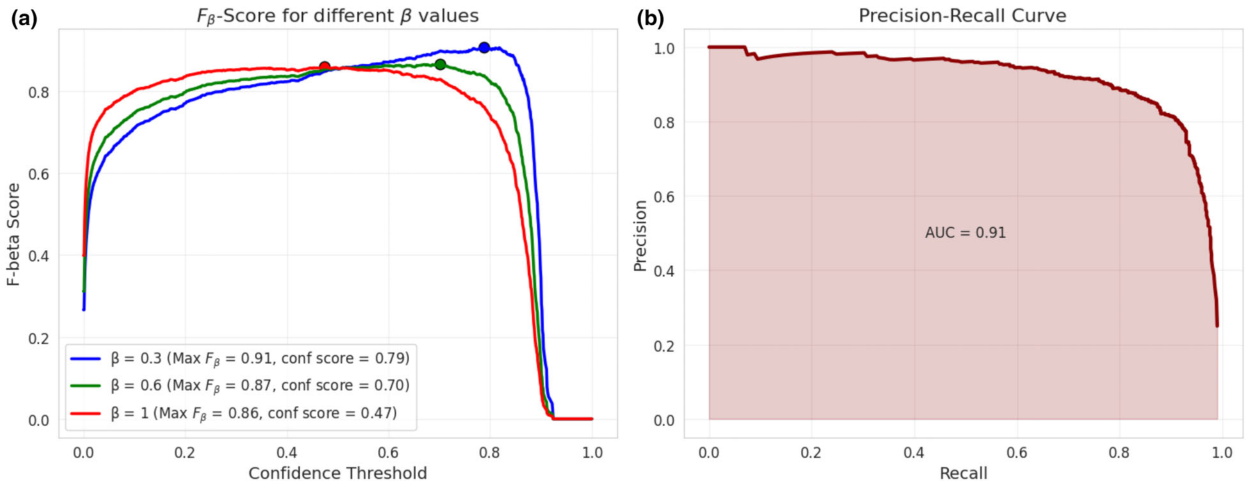

Figure 5. Analysis of validation metrics used to evaluate the model performance. F_β-Score for different values of β, with corresponding argmax (a) and the precision-recall curve with the corresponding mAP for β equals to 1 (b). The AUC is shown in red.

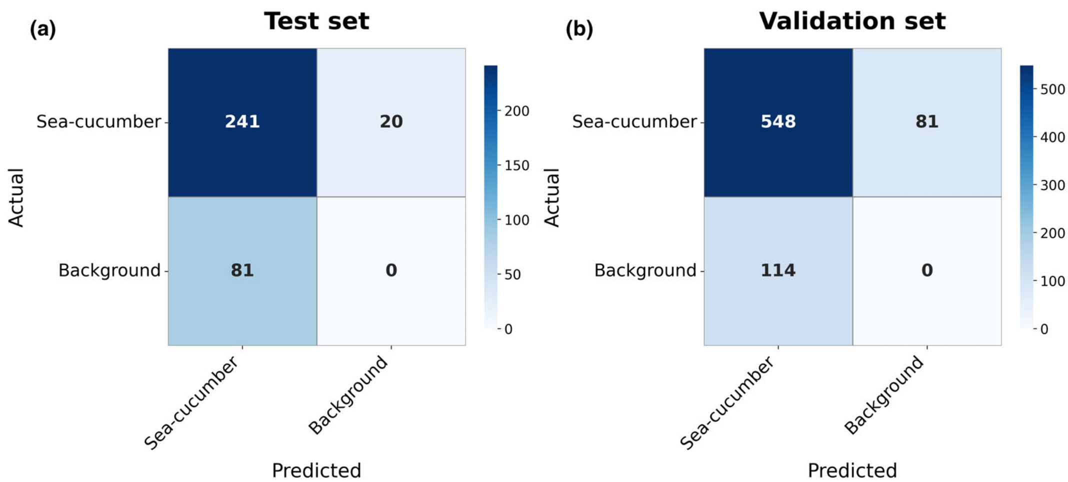

Figure 6. Confusion matrices obtained in the reduced test set (a) and the full validation set (b). The top-left cell corresponds to TP (true positive = correctly detected objects), the top-right to FN (false negative = model failed to detect an object that actually exists in the ground truth label), the bottom-left to FP (false positive = a detection that does not correspond to any ground truth object), and the bottom-right to TN (true negative = the model correctly predicts an instance as negative, not available).

Table 3. Count-based metrics (Manual, Auto, Match, False Positives, Misses and Difference) and performance-based metrics (precision, recall and F_1 − score) derived after model testing on the five raster datasets acquired during the campaigns in 2022.

	Count metrics						Performance metrics		
Sampling occasion (orthophoto mosaic)	Manual	Auto	Match	FP	FN	Diff	Precision	Recall	F_1
20/02/2022	985	721	649	72	336	264	0.9	0.659	0.761
13/04/2022	1479	1137	1041	96	438	342	0.916	0.704	0.796
20/06/2022	938	672	598	74	340	266	0.88	0.638	0.743
24/08/2022	770	472	390	82	380	298	0.826	0.506	0.628
20/10/2022	668	284	258	26	410	384	0.908	0.386	0.542

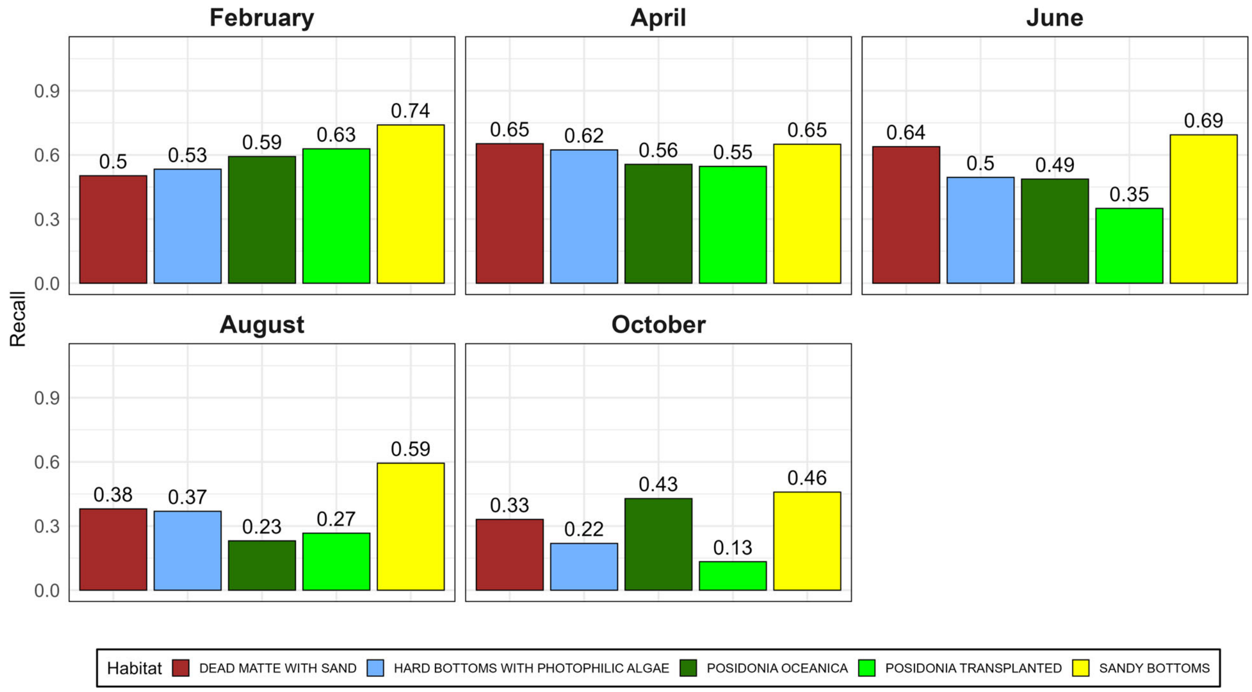

Figure 7. Bar plots reporting the Recall (i.e. TP/(TP + FN)) according to habitat type cover in each sampling occasion (month). The colour of the seabed cover classes was consistent with those reported in Fig. 1.

visual complexity and potential occlusion in these areas. Precision exhibited a more uniform distribution across habitat types (Fig. S5), with the exception of June and August, when it was more significantly influenced by *P. oceanica*. However, this trend is primarily driven by the limited number of predictions made for the *P. oceanica* habitat during these 2 months (Table S1). Overall, it can be stated without loss of generality that the underlying habitat structure had only a limited influence on false detections.

The results regarding object size revealed a significant correlation between the dimensions estimated from the diagonal measurements of bounding boxes in detection shapefiles produced by the Deepness QGIS plugin and the actual body length obtained through manual measurements of the individuals. The diagonal measurements showed a strong positive correlation with the exact body lengths, indicating a notably precise approximation with minimal measurement error. The frequency distribution analysis (Fig. 8) revealed a slight tendency toward overestimation for medium- to small-sized sea cucumbers and a correspondingly minimal underestimation in other size ranges.

The correlation analysis yielded *P*-values consistently below 0.01, which provided strong statistical evidence against the null hypothesis of no meaningful relationship. A regression analysis (Fig. S6) further contextualised the relationship between estimated and observed lengths.

The kernel density estimation analysis highlighted a strong relationship between the detector's overall performance metrics and the similarity between the estimated and observed density patterns (Fig. 9). Specifically, during periods of optimal detector performance, the estimated density distribution was closely aligned with the observed one, reinforcing the reliability of the detection process. Moreover, the estimated density patterns consistently preserved the spatial clustering characteristics of the observed distribution, albeit with reduced intensity, highlighting the aggregation areas (hotspots) of sea cucumbers during the summer and fall months. However, the KDE fitted on the estimated sites showed a potential over-smoothing effect for some campaigns (June and October) due to the reduced number of points.

Discussion

Although we utilised a case study involving sea cucumbers as an illustrative example, high-performing and non-destructive image-based techniques, including photogrammetry combined with open-source DL and GIS applications, provide substantial opportunities for customisation and scalability. In fact, our approach can be

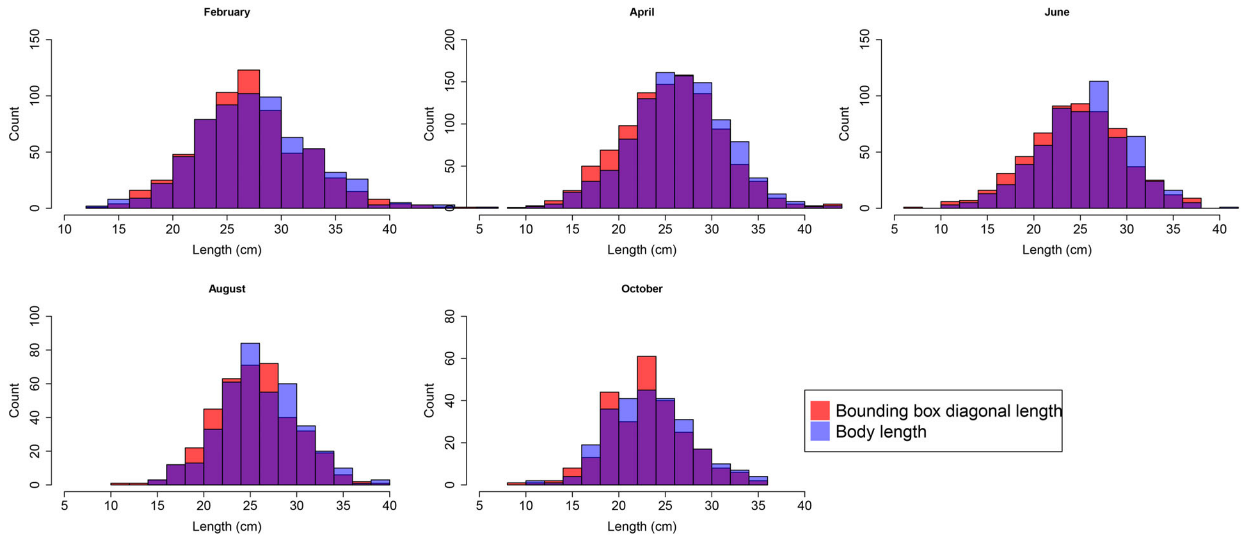

Figure 8. Frequency length distribution histograms comparing the size of sea cucumbers estimated from the diagonal of the bounding boxes in detection shapefiles generated by the Deepness QGIS plugin and the observed body length derived from manual measurements of individuals.

adapted, after specific generation of new training dataset, to a plethora of small benthic organisms, which often pose difficulties for automatic detection and therefore are helpful for ecological applications aimed at characterising ecological traits of other natural benthic assemblages such as those formed by sea urchins (Piazza et al., 2019a), sea stars (Montalbetti et al., 2022) and gorgonians (Palma et al., 2018). Specific conservation and monitoring approaches will be essential for the scientific sampling of vulnerable marine species, particularly when avoiding fishing-related methods, to protect the local population, especially in evaluations carried out within marine protected or sensitive areas. The widespread adoption of imagery-based protocols, even in remote areas (Peirano et al., 2023; Piazza et al., 2019b), will provide large datasets suitable for generating an effective trained model for detection in the near future.

This study represents the first application of the YOLOv11 object detection model, providing a practical approach for the low-cost and time-effective extraction of biologically relevant information to understand the dynamics of the sea cucumber population. Our methodology offers several benefits over the techniques established by the underwater visual census by significantly improving the efficiency of data collection and information retrieval. Specifically, the extraction of spatial locations of individuals is a fundamental requirement for ecological applications that seek to delineate species preferences and habitat associations, regardless of the type of remotely sensed data used, encompassing both aerial and underwater georeferenced imagery. Mapping solutions based on underwater photogrammetry (Ventura et al., 2025) offer a valuable approach for studying sea cucumber populations in coastal environments, where turbidity in the water often hinders the use of drone-based cartographic products (Kilfoil et al., 2020; Li et al., 2021). These efforts frequently generate vast amounts of data that require manual analysis, resulting in an impractically large workload, particularly when applied to multitemporal monitoring plans. Density-derived maps can be easily integrated into GIS environments, along with benthic habitat and bathymetric layers, to enhance the understanding of the factors influencing the distribution of sea cucumbers among complex infralittoral bottoms. The current results highlight that sea cucumbers exhibited a non-uniform and non-random distribution within their habitats, instead forming distinct clusters with seasonal patterns. Consequently, integrating mapping with automated object detection methods will be particularly advantageous for investigating natural populations, as they can be surveyed at a considerably larger scale in a shorter duration than traditional methods. By using our approach as a preliminary investigation of the study area before deploying divers or snorkelers, researchers can effectively assess the overall spatial distribution of the target species and potentially discern the environmental factors, such as proximity to seagrass meadows or artificial structures, that influence the observed spatial distribution patterns. However, the current model cannot identify sea cucumbers at the species level due to the constraints imposed by their similar body morphology, which necessitates direct examination and handling of specimens to ensure correct identification at the species level. For some cryptic species pairs (e.g., *H. tubulosa* vs. *H. mammata*),

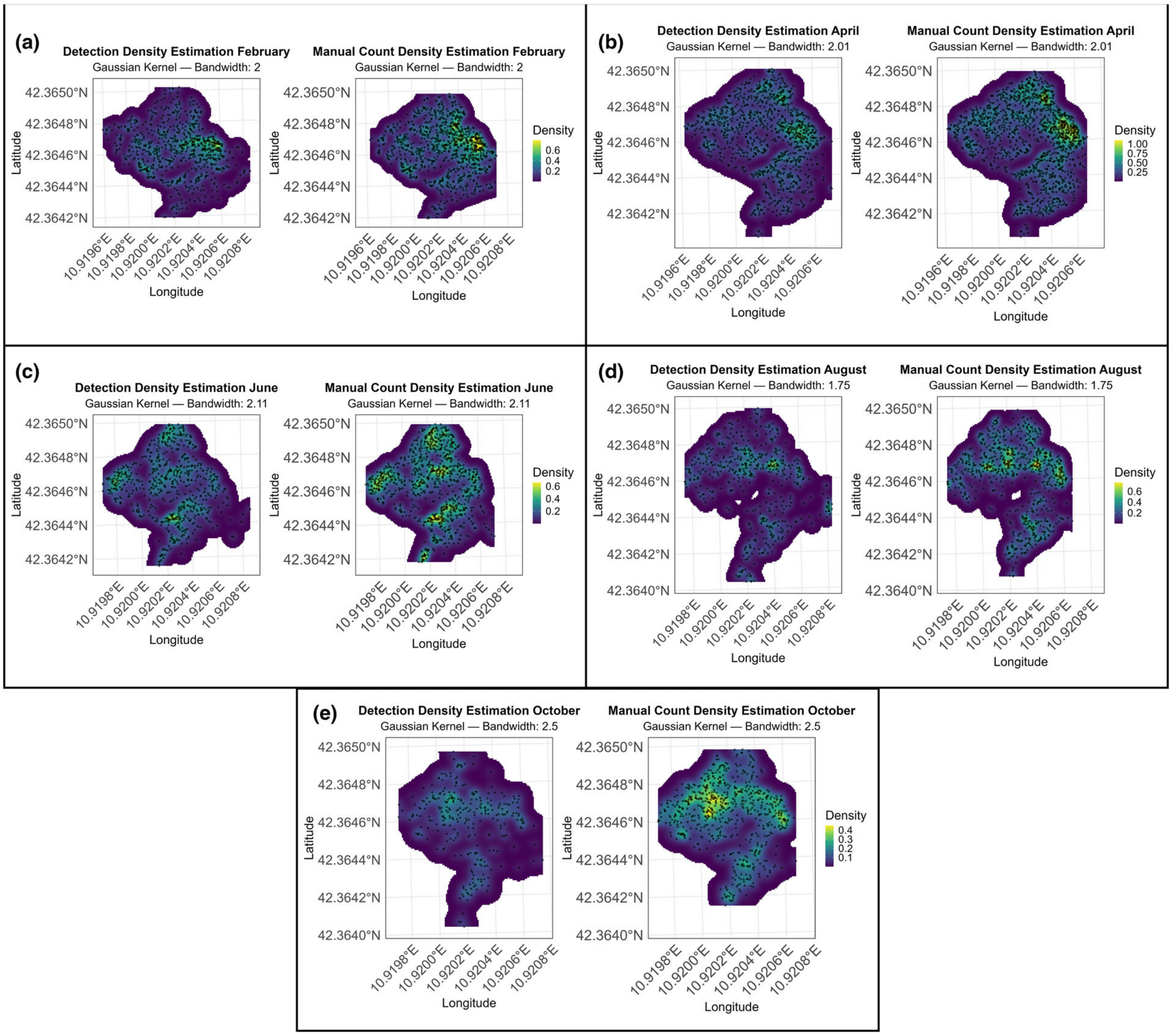

Figure 9. Gaussian kernel density estimation (KDE) heat maps comparing the spatial distribution patterns of sea cucumbers, as determined by automated counts from the object detector, against manually recorded counts of individuals. The results of each sampling campaign (months) were reported from panel (a) to (e).

accurate identification requires microscopic examination of ossicles or even dissection to observe internal organs (Rakaj & Fianchini, 2024). Even if we can improve the spatial resolution of the acquired imagery, we will likely remain unable to distinguish between various species. Consequently, *in situ* surveys will continue to serve as an essential complement to remote sensing surveys, enhancing our understanding of ecological interactions and the roles of specific species. Moreover, it is worth noting that our training dataset relies on Mediterranean species; therefore, for its application to reef habitats and tropical species, we recommend substantial retraining for accurate detection results.

Large training datasets are commonly believed to be essential to achieve high prediction accuracy (Juba & Le, 2019). However, the defined model for sea cucumber detection demonstrated an average precision of 92% despite being trained on a relatively small dataset ($\sim$2500 tiles). This finding is particularly significant, as it indicates that effective models can be trained with limited data resources, potentially reducing the time and effort required for data collection and annotation in similar applications. Furthermore, the approach does not require significant computational resources, making it accessible and efficient. Additionally, a wide range of models from the literature can be applied to this specific scenario, offering flexibility in methodological choices. Lastly, no modifications were made to the original infrastructure, yet the results remain highly satisfactory. Comparable findings regarding precision were observed in YOLOv8 models trained on a restricted dataset of images obtained from underwater video recordings captured by a consumer-grade camera (GoPro Hero 9 Black) in turbid water conditions (Chu et al., 2025). When comparing performance metrics across validation and test sets, the results align well with those reported in recent literature (Ge et al., 2024; Li et al., 2021). More specifically, precision and recall tend to be higher on average than those of other methods, despite being trained on a smaller dataset.

The inference speed was approximately 23 ms per image, comparable to other state-of-the-art approaches. When applied to the original study area using the Deepness plugin, the process takes approximately 25 min, demonstrating the model's stability and efficiency. This consistency with published benchmark models suggested the robustness of the newly proposed approach. However, direct comparisons should be made cautiously due to differences in experimental conditions, data set characteristics and evaluation protocols. Establishing a standardised benchmark dataset would be valuable for allowing more meaningful comparisons across different methodologies. Despite some studies claiming superior results, the reliability of such statements remains uncertain due to the lack of standardised evaluation criteria. Similarly, it is crucial to thoroughly evaluate the direct comparison of the proposed methodology with established underwater visual census protocols. However, this is not a trivial task because even under comparable environmental conditions, direct estimates in simple scenarios, such as open sandy habitats, may yield varying results due to the observers' experience. Consequently, in this context, a significant strength of the proposed approach is that a permanent record provided by the captured imagery can facilitate a more objective assessment, allowing for a lasting annotation of presence/absence data that can be revisited in light of new advancements in models and training datasets. The results obtained after QGIS deployment of the model showed slight discrepancies from those obtained from the small test set, denoting a notable shift in the balance between precision and recall, while the F_1-score remains relatively stable. The primary reason for this difference was the larger number of objects defined in the extended test area. However, this bias toward precision over recall aligned with the research objectives, as having highly reliable detections is more crucial for population reconstruction than achieving a high detection rate. Notably, these strong results were achieved over a substantially larger study area compared to previous studies in the literature.

The seasonal variation, implying a differential presence of sea cucumber-like shapes, can affect false positive rates and therefore the performance in detection efficiency. In fact, like terrestrial plants, *P. oceanica* loses its leaves mainly in autumn (Mutlu et al., 2022); thus, the presence of detached dead leaves in the areas at the end of august and in october might affect the detection of sea cucumbers by adding more confounding objects to the scene. Moreover, other factors, such as lighting conditions (i.e., the formation of caustics) and fish behaviour, such as the damselfish *Chromis chromis*, which may form schools during spawning periods, should be considered during the comparison and evaluation of the results, especially when different periods of the year have been surveyed. Consequently, the standardisation of the underwater sampling procedure is a crucial factor to consider in reducing variations in habitat-related performance resulting from temporal fluctuations induced by seasonal changes.

Regarding the ecological information derived from model implementation in GIS, we confirmed strong agreement with the manually derived information from our previous work. In fact, the spatial pattern clearly indicates a high density of sea cucumbers near seagrass habitats, especially within the *Posidonia* transplantation area, confirming the essential role of this restored habitat in retaining fine organic matter, which serves as a food source for deposit-feeding holothurians (Ventura et al., 2025). Similarly, estimating length using the diagonal of the bounding box can serve as a reliable approximation for the actual length of an individual. Notably, significant discrepancies in measurements were observed in individuals displaying bending and curvature, particularly those with U-shaped forms.

The density estimation reveals a distinctive spatial pattern where the automated detection process appeared to generate a (spatially) thinned representation of the original point process, which can be seen as a "realised" intensity, a degradation of the "potential" one (Chakraborty et al., 2011). Although the degradation of the original process may be primarily related to the detector's capabilities (specifically affecting only where detection occurs), the thinning mechanism likely depends on additional environmental factors, including habitat characteristics. This distinction is important because, although they may have the same impact on the original process, there is a slight but significant difference between independent thinning and intensity degradation. However, further research is necessary to understand the underlying mechanisms fully.

The developed approach can be regarded as substantial proof of concept, demonstrating the technical feasibility of employing low-cost and open-source applications for swift automated population assessments in Mediterranean waters. However, further validation will be necessary to comprehensively evaluate habitat limitations before its practical implementation in conservation decision-making for sustainable fisheries management. This is particularly important given the substantial variations in the scale of fishing activities, the status of potentially harvestable stocks and the capabilities of management agencies worldwide (Purcell et al., 2010).

Conclusions

As awareness of the ecological significance and economic potential of sea cucumbers grows, also in Mediterranean waters, researchers will continue to develop effective monitoring techniques for holothurians. Additionally, there is a growing trend in applying advanced deep learning technologies to ecological research. Our study introduced an automated model to detect sea cucumbers through underwater photogrammetric mapping. It was the first to validate the model's results on a real dataset, assessing sea cucumber ecological traits, population and density across a sizable underwater area. In this paper and the existing literature, it has been shown that object detection models perform exceptionally well for sea cucumber detection, yielding impressive results even during generalised testing on orthophotos. From a statistical perspective, a promising direction for future research would be to investigate whether it is possible to reconstruct the original process and observed population based on the point pattern estimated by the object detector. This reverse engineering approach could provide valuable insights into sea cucumber population dynamics and spatial distribution patterns, potentially enhancing our understanding of marine benthic communities.

Furthermore, this approach could enable direct inference of species distribution patterns using only image data, representing a significant advance in ecological monitoring techniques. Such a capability would streamline field research methodologies, potentially reducing the need for labor-intensive manual surveys. The results collectively highlighted the potential of employing affordable and open-source object detection applications for sea cucumbers and other benthic species to improve efficiency, replicability and coverage in surveying and monitoring practices.

Acknowledgements

This work was partially funded by the Pure Ocean Fund (3DR-4-Seac research grant) and PRIN 2002 (Project #20224LYJM5). The involvement of Daniele Poggio in formal data analysis and presentation was made possible through the PNRR-NGEU project, which received financial support from the MUR following DM 630/2024. We extend our gratitude to Dr. Przemysław Aszkowski for his contributions in implementing significant features within the final workflow. Additionally, we are grateful to the anonymous reviewers and associate editor whose constructive feedback played a crucial role in refining and enhancing the final version of this manuscript.

Author Contributions

Gian Mario Sangiovanni: Writing – original draft; methodology; validation; visualization; writing – review and editing; software; formal analysis. **Giovanna Jona Lasinio:** Data curation; supervision; resources; project administration; formal analysis; writing – original draft; writing – review and editing; methodology. **Daniele Poggio:** Writing – original draft; software; validation; visualization. **Gianluca Mastrantonio:** Writing – original draft; methodology; validation; supervision; data curation. **Alessio Pollice:** Formal analysis; supervision; data curation; validation; visualization; writing – original draft. **Arnold Rakaj:** Writing – original draft; visualization; writing – review and editing; data curation; investigation; validation. **Midhun Mohan:** Writing – original draft; data curation; investigation; validation. **Stefano Moro:** Writing – original draft; software; methodology; validation. **Edoardo Casoli:** Writing – original draft; investigation; data curation; visualization. **Daniele Ventura:** Conceptualization; investigation; funding acquisition; writing – original draft; methodology; validation; visualization; writing – review and editing; software; formal analysis; supervision; data curation.

Conflict of Interest Statement

The authors acknowledge any possible sources of conflicts of interest and affirm that they possess no recognised financial interests or personal relationships that might be perceived as influencing the research presented in this paper.

Statement on Inclusion

This research involves contributors from various nations, including researchers located in the country where the study was conducted. All contributors were involved from the beginning in the research and study design to guarantee that the wide range of perspectives they embody was taken into account from the start. Relevant literature authored by local scientists was cited whenever applicable and attempts were made to include significant works published in the local language.

Data Availability Statement

The data that support the findings of this study are openly available in GitHub at https://github.com/justgimmi/Sea-Cucumbers-Object-Detection.

References

Alexandrova, S., Tatlock, Z. & Cakmak, M. (2015) RoboFlow: A flow-based visual programming language for mobile manipulation tasks. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, WA, USA, pp. 5537–5544. https://doi.org/10.1109/ICRA.2015.7139973.

Aszkowski, P., Ptak, B., Kraft, M., Pieczyński, D. & Drapikowski, P. (2023) Deepness: deep neural remote sensing plugin for QGIS. *SoftwareX*, **23**, 101495.

Azevedo e Silva, F., Brito, A.C., Pombo, A., Simões, T., Marques, T.A., Rocha, C. et al. (2023) Spatiotemporal distribution patterns of the sea cucumber *Holothuria arguinensis* on a rocky-reef coast (Northeast Atlantic). *Estuaries and Coasts*, **46**, 1035–1045.

Baddeley, A. & Turner, R. (2014) Package 'spatstat.' Compr. R Arch. Netw. 146

Baharin, U.N.A. & Kamarudin, K.R. (2025) Using remote-controlled vehicles (ROV) as tools for sea cucumber conservation: A review. *Maritime Technology and Research*, **7** (2), 274346. https://doi.org/10.33175/mtr.2025.274346

Chakraborty, A., Gelfand, A.E., Wilson, A.M., Latimer, A.M. & Silander, J.A. (2011) Point pattern modelling for degraded presence-only data over large regions. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, **60**, 757–776.

Chu, C., Gu, Y.F., Wong, A. & Russell, B.D. (2025) Automated detection of sea cucumbers in turbid subtidal marine habitats: An explainable approach. *Ecological Informatics*, **90**, 103342. https://doi.org/10.1016/j.ecoinf.2025.103342

Félix, P.M., Azevedo e Silva, F., Simões, T., Pombo, A., Marques, T.A., Rocha, C. et al. (2024) Modelling the habitat preferences of the NE-Atlantic Sea cucumber *Holothuria forskali*: Demographics and abundance. *Ecological Informatics*, **80**, 102476. https://doi.org/10.1016/j.ecoinf.2024.102476

Félix, P.M., Pombo, A., Azevedo e Silva, F., Simões, T., Marques, T.A., Melo, R. et al. (2021) Modelling the distribution of a commercial NE-Atlantic Sea cucumber, *Holothuria mammata*: demographic and abundance Spatio-temporal patterns. *Frontiers in Marine Science*, **8**, 1–17.

Ge, F., Xuan, K., Lou, P., Li, J., Jiang, L., Wang, J. et al. (2024) Multi-object detection and behavior tracking of sea cucumbers with skin ulceration syndrome based on deep learning. *Frontiers in Marine Science*, **11**, 1–16.

Grosso, L., Rakaj, A., Fianchini, A., Morroni, L., Cataudella, S. & Scardi, M. (2021) Integrated multi-trophic aquaculture (IMTA) system combining the sea urchin *Paracentrotus lividus*, as primary species, and the sea cucumber *Holothuria tubulosa* as extractive species. *Aquaculture*, **534**, 736268.

Guo, X., Zhao, X., Liu, Y. & Li, D. (2019) Underwater sea cucumber identification via deep residual networks. *Information Processing in Agriculture*, **6**, 307–315.

Hamel, J.-F., Eeckhaut, I., Conand, C., Sun, J., Caulier, G. & Mercier, A. (2022) Global knowledge on the commercial sea cucumber *Holothuria scabra*. *Advances in Marine Biology*, **91**, 1–286.

Hammond, A.R., Meyers, L. & Purcell, S.W. (2020) Not so sluggish: movement and sediment turnover of the world's heaviest holothuroid, *Thelenota anax*. *Marine Biology*, **167**, 1–9.

Juba, B. & Le, H. S. (2019) Precision-recall versus accuracy and the role of large data sets. In Proc. AAAI Conf. Artif. Intell. Vol. 33, pp. 4039–4048.

Khanam, R. & Hussain, M. (2024). Yolov11: An overview of the key architectural enhancements. arXiv Prepr. arXiv2410.17725.

Kilfoil, R.-P.I., Kiszka, J.J., Heithaus, M.R., Zhang, Y., Roa, C.C., Ailloud, L.E. et al. (2020) Using unmanned aerial vehicles and machine learning to improve sea cucumber density estimation in shallow habitats. *ICES Journal of Marine Science*, **77**, 2882–2889.

Li, D.S., Joyce, K.E. & Xiang, W. (2021) SeeCucumbers: using deep learning and drone imagery to Detect Sea cucumbers on coral reef flats. *Drones*, **5**, 1–19.

Mancini, G., Casoli, E., Ventura, D., Jona Lasinio, G., Belluscio, A. & Ardizzone, G.D. (2021) An experimental investigation aimed at validating a seagrass restoration protocol based on transplantation. *Biological Conservation*, **264**, 109397. https://doi.org/10.1016/j.biocon.2021.109397

Mancini, G., Casoli, E., Ventura, D., Jona-Lasinio, G., Criscoli, A., Belluscio, A. et al. (2019) Impact of the Costa Concordia shipwreck on a *Posidonia oceanica* meadow: A multi-scale assessment from a population to a landscape level. *Marine Pollution Bulletin*, **148**, 168–181.

Montalbetti, E., Fallati, L., Casartelli, M., Maggioni, D., Montano, S., Galli, P. et al. (2022) Reef complexity influences distribution and habitat choice of the corallivorous seastar *Culcita schmideliana* in the Maldives. *Coral Reefs*, **41**, 253–264.

Mutlu, E., Olguner, C., Gökoğlu, M. & Özvarol, Y. (2022) Seasonal growth dynamics of *Posidonia oceanica* in a pristine Mediterranean gulf. *Ocean Science Journal*, **57**, 381–397.

Ozge Unel, F., Ozkalayci, B. O. & Cigla, C. (2019) The power of tiling for small object detection. In Proc. IEEE/CVF Conf. Comput. Vis. pattern Recognit. Work.

Öztürk, C., Taşyürek, M. & Türkdamar, M.U. (2023) Transfer learning and fine-tuned transfer learning methods' effectiveness analyse in the CNN-based deep learning models. *Concurrency and Computation: Practice and Experience*, **35**, e7542.

Palma, M., Rivas Casado, M., Pantaleo, U., Pavoni, G., Pica, D. & Cerrano, C. (2018) SfM-based method to assess gorgonian forests (*Paramuricea clavata* (Cnidaria, Octocorallia)). *Remote Sensing*, **10**, 1154.

Pasquini, V., Porcu, C., Marongiu, M.F., Follesa, M.C., Giglioli, A.A. & Addis, P. (2022) New insights upon the reproductive biology of the sea cucumber *Holothuria tubulosa* (Echinodermata, Holothuroidea) in the Mediterranean: implications for management and domestication. *Frontiers in Marine Science*, **9**, 1–16.

Peirano, A., Bordone, A., Corgnati, L.P. & Marini, S. (2023) Time-lapse recording of yearly activity of the sea star *Odontaster validus* and the sea urchin *Sterechinus neumayeri* in Tethys bay (Ross Sea, Antarctica). *Antarctic Science*, **35**, 4–14.

Piazza, P., Cummings, V., Guzzi, A., Hawes, I., Lohrer, A., Marini, S. et al. (2019) Underwater photogrammetry in Antarctica: long-term observations in benthic ecosystems and legacy data rescue. *Polar Biology*, **42**, 1061–1079.

Prescott, J., Vogel, C., Pollock, K., Hyson, S., Oktaviani, D. & Panggabean, A.S. (2013) Estimating sea cucumber abundance and exploitation rates using removal methods. *Marine and Freshwater Research*, **64**, 599–608.

Purcell, S.W., Conand, C., Uthicke, S. & Byrne, M. (2016) Ecological roles of exploited sea cucumbers. *Oceanography and Marine Biology: An Annual Review*, **54**, 367–386.

Purcell, S.W., Lovatelli, A. & Pakoa, K. (2014) Constraints and solutions for managing Pacific Island sea cucumber fisheries with an ecosystem approach. *Marine Policy*, **45**, 240–250.

Purcell, S. W. Lovatelli, A. Vasconcellos, M. & Ye, Y. (2010) Managing sea cucumber fisheries with an ecosystem approach.

Purcell, S.W., Mercier, A., Conand, C., Hamel, J.F., Toral-Granda, M.V., Lovatelli, A. et al. (2013) Sea cucumber fisheries: global analysis of stocks, management measures and drivers of overfishing. *Fish and Fisheries*, **14**, 34–59.

Rakaj, A. & Fianchini, A. (2024) Mediterranean sea cucumbers—Biology, ecology, and exploitation. In: *The world of sea cucumbers*. Academic Press, New York, USA, pp. 753–773.

Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016) You only look once: Unified, real-time object detection. In Proc. IEEE Conf. Comput. Vis. pattern Recognit. pp. 779–788.

Schneider, K., Silverman, J., Kravitz, B., Rivlin, T., Schneider-Mor, A., Barbosa, S. et al. (2013) Inorganic carbon turnover caused by digestion of carbonate sands and metabolic activity of holothurians. *Estuarine, Coastal and Shelf Science*, **133**, 217–223.

Shrout, P.E. & Fleiss, J.L. (1979) Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, **86**(2), 420.

Ventura, D., Mancini, G., Casoli, E., Pace, D.S., Lasinio, G.J., Belluscio, A. et al. (2022) Seagrass restoration monitoring and shallow-water benthic habitat mapping through a photogrammetry-based protocol. *Journal of Environmental Management*, **304**, 114262.

Ventura, D., Rakaj, A., Lasinio, G.J., Sangiovanni, G.M., Poggio, D., Mastrantonio, G. et al. (2025) Detecting habitat preferences and monitor population dynamics of sea cucumbers in coastal ecosystems through underwater photogrammetry. *Journal of Environmental Management*, **377**, 124589.

Wang, Y., Fu, B., Fu, L. & Xia, C. (2023) In situ sea cucumber detection across multiple underwater scenes based on convolutional neural networks and image enhancements. *Sensors*, **23**, 2037.

Williamson, J.E., Duce, S., Joyce, K.E. & Raoult, V. (2021) Putting sea cucumbers on the map: projected holothurian bioturbation rates on a coral reef scale. *Coral Reefs*, **40**, 559–569.

Wolfe, K. & Byrne, M. (2017) Biology and ecology of the vulnerable holothuroid, Stichopus herrmanni, on a high-latitude coral reef on the Great Barrier Reef. *Coral Reefs*, **36**, 1143–1156.

Woo, S.P., Yasin, Z., Ismail, S.H. & Tan, S.H. (2013) The distribution and diversity of sea cucumbers in the coral reefs of the South China Sea, Sulu Sea and Sulawesi Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, **96**, 13–18.

Zhai, X., Wei, H., He, Y., Shang, Y. & Liu, C. (2022) Underwater Sea cucumber identification based on improved YOLOv5. *Applied Sciences*, **12**, 9105.

Zoph, B., Cubuk, E. D., Ghiasi, G., Lin, T.-Y., Shlens, J. & Le, Q. V. (2020) Learning data augmentation strategies for object detection. In Comput. Vision–ECCV 2020 16th Eur. Conf. Glas. UK, August 23–28, 2020, Proceedings, Part XXVII 16, pp. 566–583. Springer.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1 Supplementary Information.

