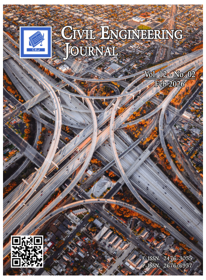
Available online at www.CivileJournal.org

Civil Engineering Journal

(E-ISSN: 2476-3055; ISSN: 2676-6957)

Vol. 12, No. 02, February, 2026



An Automated Framework for Benthic Habitat Classification and Segmentation Based on Deep Learning Algorithms

Hassan Mohamed [1*], Kazuo Nadaoka [2, 3]

[1] Department of Geomatics Engineering, Shoubra Faculty of Engineering, Benha University, Cairo 11672, Egypt.

[2] Graduate School of Information Science and Engineering, Tokyo Institute of Technology, Tokyo 152-8552, Japan.

[3] Kajima Technical Research Institute (KaTRI), Chofu 182-0036, Japan.

Received 15 July 2025; Revised 21 January 2026; Accepted 26 January 2026; Published 01 February 2026

Abstract

Although benthic habitats represent some of the largest, most diverse, and productive ecosystems on Earth with great environmental, and economical value, they are increasingly threatened and declining in many locations worldwide. Every year, numerous underwater images are collected for monitoring these habitats. Still, the manual labelling process remains tedious and time-consuming, creating a huge gap between data collection and extraction of meaningful information. In this study, an automated framework is proposed for single-label classification and semantic segmentation of benthic habitats using convolutional neural networks (CNNs). The framework integrates and evaluates various pre-trained CNNs, bagging of features (BOF), color spaces, and texture descriptors for benthic habitat classification. Furthermore, the classified images served as training and validation samples to assess the semantic segmentation performance of pre-trained CNNs with different architectures (e.g., ResNet-50, AlexNet, Xception, etc.). Both high- and low-quality underwater images of benthic habitats collected from six diverse study areas located off Australia and Japan were used to evaluate the proposed framework. The analysis revealed that the ResNet-50 FC1000 combined with BOF, color space, and texture attributes yielded the highest automatic classification accuracy. Moreover, the ResNet-50 network outperformed all the tested networks for automatic semantic segmentation of benthic habitats. Overall, the presented framework enhanced the automation of benthic habitat classification and semantic segmentation processes.

Keywords: Underwater Images; Benthic Habitats Classification; Benthic Habitats Semantic Segmentation; Convolutional Neural Networks.

1. Introduction

Benthic habitats and seagrass species are complex and biodiverse marine ecosystems with immense environmental and economic value [1]. These irreplaceable ecosystems provide habitats for numerous marine species, protect shorelines, sequester blue carbon, facilitate nutrient cycling, and generate revenue through tourism [2-4]. However, they are highly fragile and increasingly vulnerable to human stressors such as climate change, pollution, and overfishing, leading to their rapid decline in recent decades [5-7]. Recent studies report terrible statistics, 25 % of coral reefs were bleached, more than 50% are currently under threat, and up to 70% may be severely damaged if degradation continues [8, 9]. Protecting these important ecosystems requires accurate classification, mapping, and long-term monitoring programs for benthic habitats and seagrass meadows [10]. Still, there is a huge global lack of benthic habitats accurate maps limiting the conservation efforts for these high value ecosystems [11, 12].

* Corresponding author: hassan.mohamed@ejust.edu.eg; hassan.hussein@feng.bu.edu.eg

https://doi.org/10.28991/CEJ-2026-012-02-022

Remote sensing imagery provides valuable data for mapping large-scale benthic habitat areas [13, 14]. Free satellite images from Landsat 8 and Sentinel-2 with 10 m spatial resolution offer multi-temporal datasets for broad benthic habitat classification. However, these pixel sizes lack the detail necessary to accurately map small-scale features and substrate complexity [15]. High-resolution commercial satellites, including Planet Dove, WorldView-2, and Spot 7, provide improved spatial detail, enabling more precise mapping of benthic habitats. Nevertheless, the high cost of these high-resolution datasets remains a significant challenge [16]. Despite their advantages, these technologies have limitations, including insufficient spatial resolution for detailed benthic studies, environmental challenges such as cloud cover, shadows, water clarity, surface roughness, and sun glint, which can obscure reef features [17].

Numerous remote sensing techniques including scuba diving, towed cameras, and Autonomous Underwater Vehicles (AUVs) have been employed to collect high spatial and temporal resolution underwater images for monitoring changes in benthic habitats [18-20]. Millions of underwater images of the seafloor are collected using these platforms every year [21]. For instance, in Western Australia an AUV collects around 200,000 images annually [22]. However, the manual annotation of this unprecedented volume of underwater imagery by marine experts is both laborious and time-consuming, as each image may require up to 30 minutes for complete annotation [23]. According to the National Oceanic and Atmosphere Administration (NOAA) less than 2% of the underwater images acquired annually are adequately analyzed by marine experts, resulting in a substantial gap between data collection and extrapolation [24].

Automated classification of benthic habitats and seagrass images remains a challenging task for two primary reasons: the inherent nature of these ecosystems and the difficulties of underwater imaging environments [25, 26]. First, the dynamic nature and continuous movement of benthic habitats and seagrasses, caused by currents and waves, alter their visual appearance in images. Additionally, their irregular 3D structure and plasticity forms of benthic habitats added difficulties to define their shapes and edges [27]. Also, many organisms have visual similarities (e.g., algae and corals) and often coexist in the same spatial locations [24]. Second, underwater images usually had major drawbacks, including inadequate lighting conditions, water turbidity, and optical distortions decreasing image quality and clarity. Furthermore, the variations in the water column result in low contrast, color distortion, and blurred underwater images [28]. However, recent progress in computer vision and deep learning techniques offers promising potential to automate the annotation of underwater images and help bridge this critical bottleneck [29].

Generally, previous studies on processing underwater images can be summarized into image classification, object detection, region proposal, and semantic segmentation. The present study focuses on two major techniques: single-label classification and semantic segmentation. First, several studies have explored single-label classification, where each underwater image in a dataset is classified as a single class as coral, algae, or seagrass [30, 31]. Numerous studies combined hand-crafted features (i.e., shape, color, and texture) to discriminate between underwater images. For example, Jamil et al. [32] proved the superiority of Bagging of Features (BOF) extracted features compared to other descriptors such as histogram of oriented gradients (HOG), texture, and local binary patterns (LBP) features for classifying bleached and unbleached coral underwater images. Using a dataset of 230 underwater images from the Great Barrier Reef of Australia, they achieved about 99% overall accuracy. Furthermore, the effectiveness of BOF has been proved in other marine applications, such as bleached coral detection [33] and monitoring benthic assemblages [34]. Similarly, Awalludin et al. [35] combined Hue Saturation Value (HSV) color attributes and Local Binary Pattern (LBP) texture descriptors to classify four classes of coral reef images. Their proposed approach resulted in 95% overall accuracy using an 800-image dataset. Additionally, Srividhya & Ramya [36] evaluated 14 texture features for classifying three classes sea cucumber, fish, and corals using a dataset of 200 images and achieved 97% overall accuracy.

Moreover, recent studies developed novel features descriptors enhanced the discrimination power of classification models [37, 38]. Ganesan and Santhanam [39] developed a unique feature descriptor by detecting directional edge details in each coral reefs image and then contacting the local neighborhood pixels. They compared the proposed descriptor with traditional descriptors using five coral reefs datasets: EILAT (1123 images & 8 classes), EILAT2 (303 images & 5 classes), MLC (776 images & 9 classes), Brodartz (2800 images & 112 classes), and RSMAS (24159 images & 14 classes). The proposed descriptor outperformed the baseline methods, achieving overall accuracies ranging from 96% to 98%. Nazmi et al. [40] combined the symmetric binary method and the median enhanced binary method to reduce the number of extracted features. This integration merged selected features from the evaluated coral reefs images and enabled the extraction of more detailed information. Using the same abovementioned datasets, the proposed method improved the overall accuracy by approximately 10% compared with traditional descriptors. Sotoodeh et al. [41] presented two local color texture descriptors that were robust to illumination variations and noise in coral reefs images, effectively capturing microstructural texture relations. These descriptors were evaluated using the EILAT, EILAT2, RSMAS, and MLC coral reefs datasets and they provided higher accuracies than conventional descriptors.

Recently, convolutional neural networks (CNNs) have provided robust frameworks for coastal applications including the extraction of offshore raft aquaculture areas [42], prediction of nearshore waves and hydrodynamics [43], and identification of coastlines [44] and sea fog [45] from satellite imagery. Furthermore, CNNs have shown significant promise in efficiently processing large-scale image datasets, which has accelerated research in benthic habitats images classification [46-48]. For instance, Fawad et al. [49] integrated BOF with hybrid color texture and AlexNet feature descriptors to classify bleached and unbleached corals. Their approach was evaluated using 342 coral images from the Great Barrier Reef, achieving an overall accuracy of 96.2%. Furthermore, Game et al. [47] evaluated the VGG16 [50] CNN for feature extraction and subsequently fed the extracted features into a Support Vector Machine (SVM) classifier. Using between 574 and 8,353 images from three datasets to classify three classes, their approach achieved overall accuracies ranging from 87% to 95%. In another study, Yasir et al. [51] proposed two image enhancement techniques combined with CNN-extracted feature descriptors to classify coral reefs images. The MLC dataset including 2055 images was used to classify nine classes. They achieved 87.40% overall accuracy using DenseNet-169 [52] feature descriptors.

The second technique is semantic segmentation, which classifies each underwater image into various classes using either machine learning algorithms [53-55] or, more recently, deep learning algorithms [56-58]. Jagadish et al. [59] proposed a machine learning framework for monitoring coral reefs' health in the Great Barrier Reef, Australia. After preprocessing images and extracting Gray-Level Co-occurrence Matrix (GLCM) and spectral index features, both pixel- and object-based segmentation were applied. The Random Forest classifier was then employed to model the nonlinear relationships among these features, achieving a high segmentation accuracy of 99.59%. In de Oliveira et al. [60], six machine learning algorithms were evaluated to classify live corals, dead corals, coral rubble, and sediments in the Piddington Mound area, Ireland, using underwater imagery collected by a remotely operated vehicle. Among the tested algorithms, gradient boosting trees achieved the highest overall accuracy of 95%. Nieuwenhuis et al. [61] integrated unmanned aerial vehicle images with digital elevation models to classify corals, microalgae, rubble, sand, and seagrass along the Saudi Arabian Red Sea coast using object-based image analysis, obtaining an overall accuracy of 84.4%. Driven by recent advances in deep learning algorithms, many studies have developed novel CNN architectures to enhance the automated semantic segmentation of benthic habitats. For example, Zhong et al. [62] designed a new deep neural network that integrates image mosaics and digital surface models to segment coral reefs into three classes. They yielded a superior mean Intersection over Union (mIoU) of 85%. Similarly, Zuo et al. [63] proposed a transformer-based model, the underwater segmentation transformer (UWSegFormer), to improve semantic segmentation of low-quality underwater images. UWSegFormer integrates the Underwater Image Quality Attention (UIQA) and Multi-scale Aggregation Attention (MAA) modules to enhance semantic feature representation and recover fine structural details, along with an Edge Learning Loss (ELL) to refine object boundary detection. Experiments conducted on two datasets demonstrated that UWSegFormer outperformed existing methods in segmentation accuracy and boundary sharpness, achieving a peak mIoU of 82.12%.

Although numerous studies have contributed to the advancement of underwater image classification and segmentation, several important limitations persist in the existing literature. First, most existing studies focus on either hand-crafted descriptors or CNN-extracted features alone, with limited systematic evaluation of how shape, color, texture, and deep features compare across diverse benthic environments. Second, the majority of prior work relies on a single dataset, making it difficult to assess the generalizability of proposed methods to images with different water qualities, illumination conditions, and habitat compositions. Third, very few studies have explored the complementary strengths of traditional and deep-learning-based descriptors through an integrated feature representation, even though benthic habitats exhibit complex visual patterns that often cannot be captured by a single feature type alone. Finally, while both single-label classification and semantic segmentation have been widely investigated, a unified framework capable of handling the millions of underwater images collected annually across heterogeneous underwater datasets remains largely absent in the literature. To address these gaps, the present study systematically investigates multiple descriptors, proposes an integrated feature configuration that enhances discrimination across habitat types, and rigorously evaluates both classification and segmentation models on six distinct underwater datasets. These methodological advances collectively enhance robustness, improve cross-dataset generalization, and provide a more comprehensive framework for automated benthic habitat mapping.

This study presents an automated framework for the single-label classification and semantic segmentation of benthic habitats. The principal contributions of this research are summarized as follows: (i) several combinations of shape-, color-, texture-, and CNN-based feature descriptors were systematically investigated for underwater image classification; (ii) the ResNet-50 network descriptors demonstrated superior performance compared with other state-of-the-art CNN architectures; (iii) integrating all descriptor types yielded higher classification accuracies than using CNN-based features alone across six distinct underwater image datasets; (iv) the classified images were subsequently employed to assess multiple CNN architectures for automated semantic segmentation of benthic habitats; and (v) the experimental findings confirmed that the ResNet-50 network achieved the highest segmentation accuracy among all evaluated models across diverse underwater environments.

2. Material and Methods

2.1. Datasets

For the benthic habitat classification and segmentation experiments, six different datasets were employed. To ensure the generalizability of the proposed framework, these datasets encompassed a broad spectrum of characteristics, including differences in geographic locations, water turbidity, illumination conditions, the number and nature of benthic habitat species, and underwater imagery quantity and quality. Four datasets originated from the Australian benthic datasets (Benthoz15) [64] as samples of georeferenced high-quality benthic images. These images were obtained using the AUV Sirius as part of Australia's Integrated Marine Observation System (IMOS) benthic monitoring initiative across various locations in Australia. The AUV Sirius was equipped with an illumination system and GPS receiver as detailed in Johnson-Roberson et al. [65]. Underwater imagery was captured using a high-resolution stereo camera with an approximate field of view of 42 × 34 degrees, producing 1,360 × 1,024 pixel RGB images at a standard altitude of 2 meters. Marine experts manually annotated the imagery following the Collaborative and Automation Tools for Analysis of Marine Imagery and Video (CATAMI) classification scheme [66, 67].

This scheme has been widely adopted in numerous studies [68-70] due to its ability to facilitate consistent processing and interpretation of benthic habitat imagery. To ensure data reliability, all images within each class were carefully reviewed prior to classification. Two additional low-quality, georeferenced datasets were derived from field surveys in the Shiraho and Fukido regions of Ishigaki Island, Japan. Both survey areas had an approximate water depth of three meters, and data collection took place during the typhoon season in August, which increased water turbidity. Underwater images were captured using a GoPro HERO3 Black Edition towed video camera with 12-megapixel photo resolution [71], affixed on a wooden frame positioned alongside a motorboat and submerged just below the water surface. Moreover, the coordinates of the surveyed underwater images were simultaneously recorded using a differential global positioning system (DGPS) mounted directly above the camera. Subsequently, 3000 benthic habitat images from Shiraho were manually classified into seven benthic classes, while 1500 seagrass images from Fukido were categorized into four seagrass species. The spatial distribution of all study areas and representative examples of the corresponding benthic classes are illustrated in Figures 1 and 2, respectively. Moreover, a summary of the evaluated datasets is provided in Table 1.

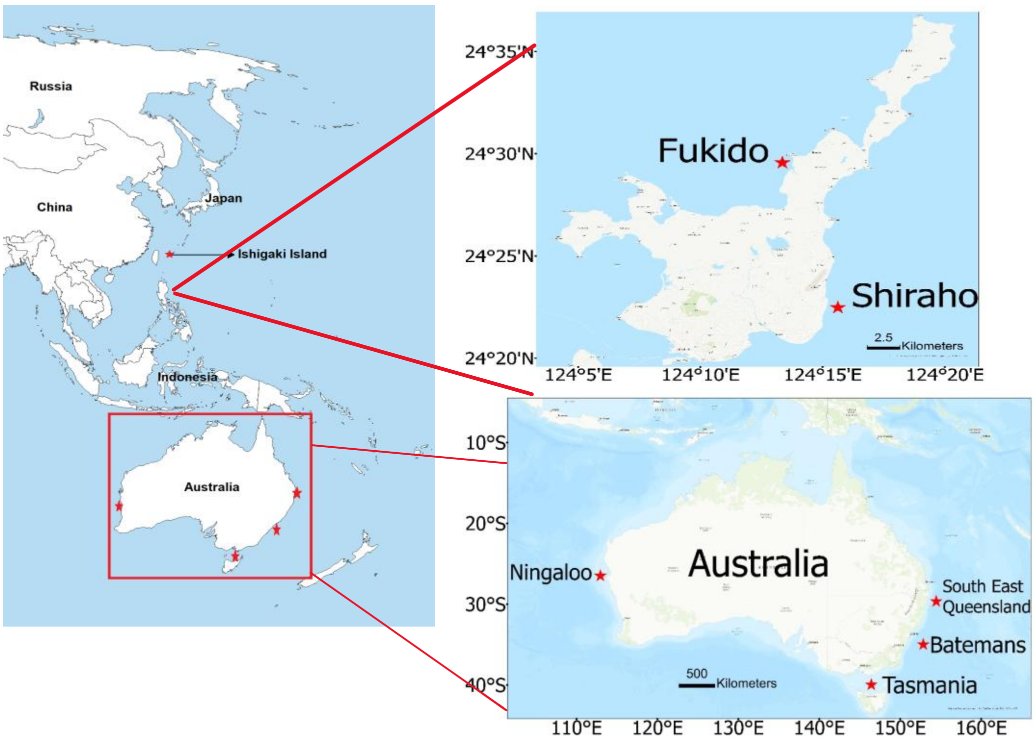


Figure 1. The datasets locations shown with red stars; Australia and Japan

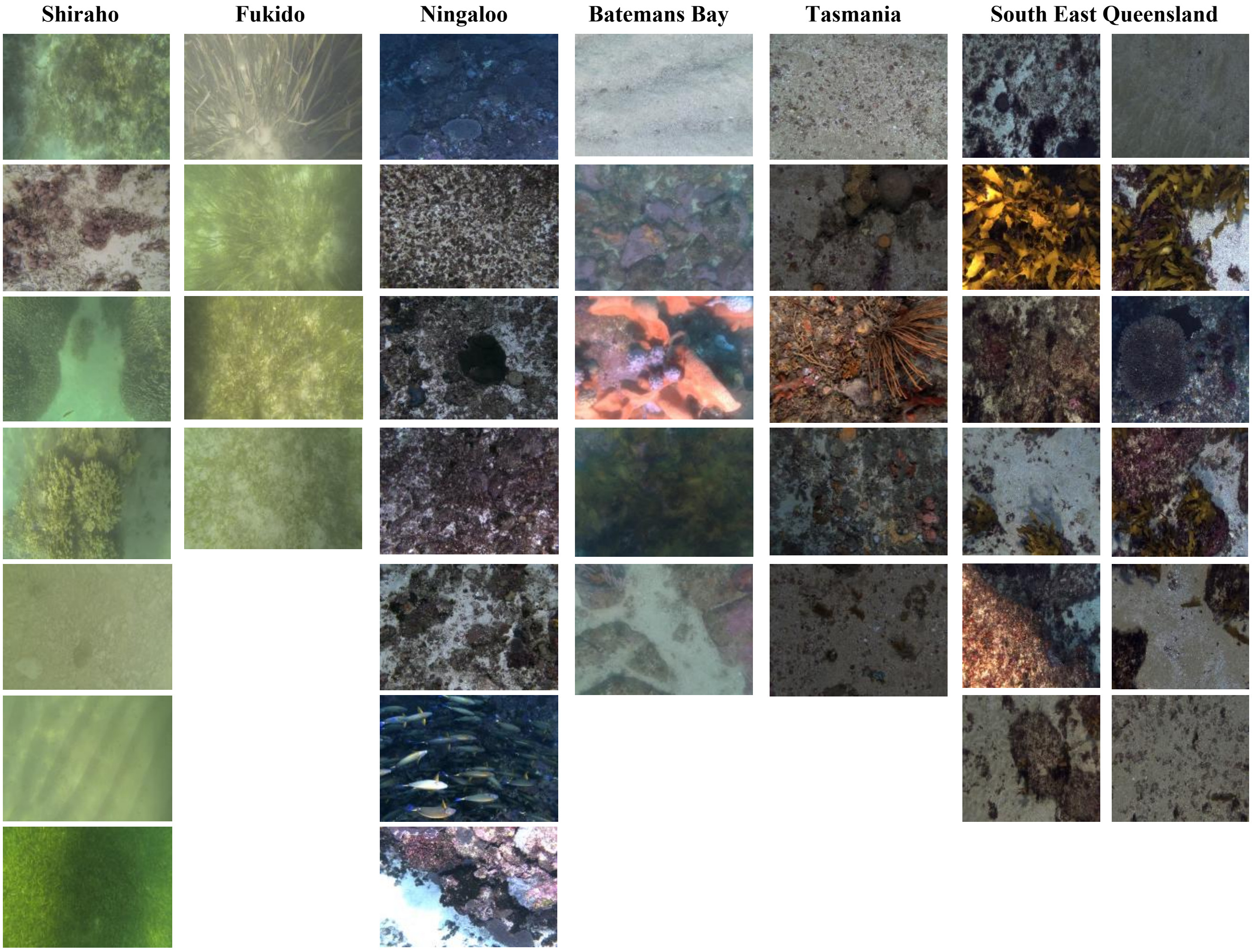


Figure 2. Sample images of the six different datasets used in this study

Table 1. Details of the datasets used in this study

Dataset	Location	Survey Year	# of classes	# of Images
Batemans Bay	Australia	2010	5	8870
Tasmania	Australia	2018	5	8763
Ningaloo	Australia	2019	7	19708
South East Queensland	Australia	2019	12	20088
Fukido	Japan	2016	4	1500
Shiraho	Japan	2016	7	3000

2.2. Methodology

The proposed framework in this study consists of two sequential stages. In the first stage, a SVM [72] classifier with a 3rd order polynomial kernel function was employed to classify benthic habitats images. The input attributes for this classification were derived from pre-trained CNNs, Gray Level Co-occurrence Matrix (GLCM) texture parameters [73], color spaces, and the BOF [74] approach. In the subsequent stage, the correctly classified images served as training data for multiple CNN architectures to perform semantic segmentation of benthic habitats.

2.2.1. Benthic Habitats Single Label Classification

The following steps illustrate the single label classification stage of the proposed framework:

- Manually labeled benthic images from each dataset (Table 1) served as inputs to evaluate various approaches.
- Descriptors from the ResNet-50 FC1000 layer [75], HSV color space, GLCM texture parameters, and BOF shape features were integrated for automated single label classification.
- The extracted attributes produced from the abovementioned approaches were utilized as input to train the SVM classifier, with the resulting outputs representing the predicted image labels.
- To validate the SVM classifier, a random sampling strategy was employed, dividing the images within each class to 60% of images used for training and a separate independent 40% for testing.

- Additional images were then classified using the validated SVM classifier, and each classified image was individually verified.

By the end of this stage, the framework had automatically classified underwater images into benthic habitat classes, classifying each image based on its dominant class. For benthic habitats classification, 250 BOF attributes were extracted utilizing a block width of 32 and a grid step of 16. The strongest 80% of features from each class were retained based on the grid point selection method. In addition, 26 GLCM texture attributes were computed for each image including entropy, correlation, homogeneity, energy, and other texture measures. The classification accuracy for each class was determined by dividing the number of correctly classified test images by the total number of test images in that class. The overall accuracy (OA) was then calculated as the mean classification accuracy across all tested images.

2.2.2. Benthic Habitat Semantic Segmentation

Correctly classified underwater images from the evaluated datasets were subsequently used as inputs for benthic habitat semantic segmentation as follows:

- A total of 750 benthic images were selected and equally divided to represent six major dominant benthic cover classes—corals, blue corals, algae, brown algae, seagrasses, and sediments—representing the categories present in all tested datasets.
- Manual digitization was meticulously performed on these images and the resulted digitized images were converted into raster format using ARC GIS software. To ensure accuracy, two additional experts independently reviewed and compared the manually classified images to the original ones prior to the evaluation of the proposed methods.
- These images were employed as inputs to evaluate pretrained ResNet-50 for benthic habitat segmentation. Within each class, images were split into 70% training images and separate 30% testing images for the evaluation process.
- Finally, the proposed framework produced two outputs for each image: (i) a single-label classification identifying the dominant benthic class, and (ii) a semantically segmented image depicting the spatial distribution of benthic habitats.

For benthic habitats segmentation, the optimal training parameters for the ResNet-50 network that yielded the best results were as follows: optimization method—stochastic gradient descent with momentum (SGDM); momentum—0.9; L2 regularization—0.0005; learning rate—0.001; mini-batch size—8; and maximum number of epochs—25. The training session consumed approximately 110 minutes using a single GPU NVIDIA RTX 3060. To assess segmentation performance, three widely adopted metrics were utilized: Mean Pixel Accuracy (MPA), F1-score, and Mean Intersection over Union (mIoU) [76]. The mIoU measures the average overlap ratio between the predicted and ground truth classes, whereas the F1-score evaluates the correspondence between the predicted object boundaries and their ground truth counterparts. Additionally, the MPA represents the average percentage of correctly classified pixels across all classes. All experiments were conducted in the MATLAB environment, and the overall workflow of the proposed framework is illustrated in Figure 3.

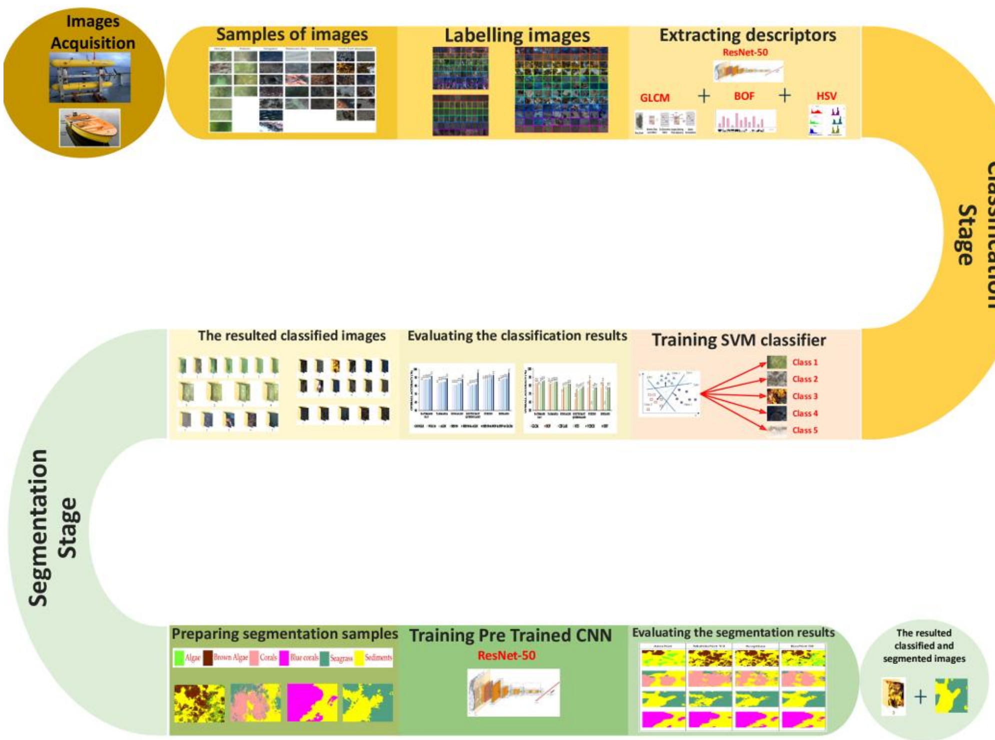


Figure 3. The general workflow of the methodology used in this study for benthic habitats single label classification and segmentation

3. Results

3.1. Benthic Habitats Single Label Classification

Descriptors derived from various approaches—such as multiple convolutional layers from pretrained CNNs (e.g., loss3 [77], AlexNet FC7 [78], VGG16 FC7 [50], and ResNet-50 FC1000)—as well as different color spaces (i.e., HSV, YIQ, YCbCr, and CIELAB), GLCM texture features, BOF shape parameters, and combinations thereof were evaluated for automated single-label classification. The overall classification accuracies obtained using the SVM classifier with GLCM, BOF, and various color space attributes are summarized in Table 2 and Figure 4. In addition, Table 3 and Figure 5 present the overall accuracies achieved using the SVM classifier with descriptors extracted from different CNN convolutional layers across the evaluated datasets. The best results are shown in bold in all tables. Moreover, an "Improvement over baseline (%)" column, was added to Table 3 to demonstrate how much the combined descriptor configuration (RES 50 FC1000 & BOF & HSV & GLCM) improves upon the best single CNN model (the baseline = highest among GoogleNet, VGG16, AlexNet, or ResNet-50 individually). The improvement over baseline was computed as follows:

$$\text{Improvement over baseline} = \frac{\text{Combined Accuracy} - \text{Baseline Accuracy}}{\text{Baseline Accuracy}} \times 100\% \tag{1}$$

Table 2. The overall accuracy (in %) for the compared texture, shape, and color spaces attributes on various datasets. Numbers of attributes were presented in each descriptor

Dataset (No. of classes)	GLCM (26)	BOF (250)	CIB LAB (256)	YIQ (256)	YCBCR (256)	HSV (256)
Batemans Bay (5)	83	83	78	83	85	**86**
Tasmania (5)	83	82	84	84	85	**85**
Ningaloo (7)	70	78	79	79	82	**82**
South East Queensland (12)	64	76	74	76	77	**79**
Fukido (4)	78	**87**	54	74	78	78
Shiraho (7)	78	**84**	54	73	78	78

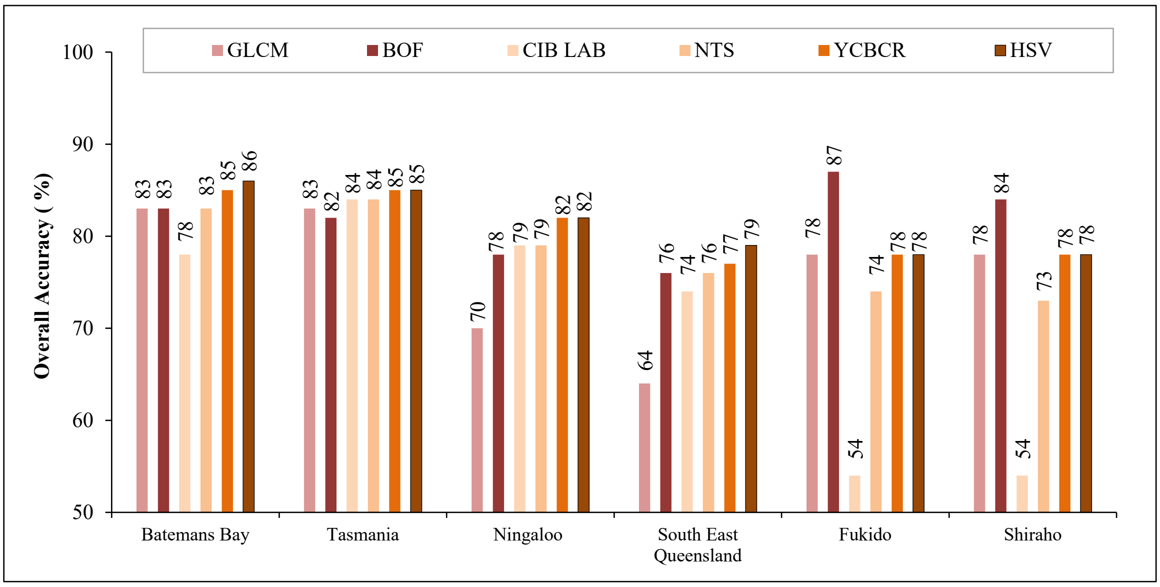


Figure 4. Comparison of overall accuracies of different texture, shape, and color spaces attributes for all datasets

Table 3. The overall accuracy (in %) for the evaluated pretrained CNNs and integrated descriptors on various datasets. Numbers of attributes were presented in each descriptor

Dataset (No. of classes)	GOOGLE NET LOSS 3 (1000)	VGG16 FC7 (4096)	ALEX FC7 (4096)	RES 50 FC1000 (1000)	RES 50 & ALEX FC 7 (5096)	RES 50 FC1000 & BOF & HSV & GLCM (1532)	Improvement over baseline (%)
Batemans Bay (5)	85	87	88	88	88	**92**	+4.55%
Tasmania (5)	81	84	84	85	85	**89**	+4.71%
Ningaloo (7)	81	81	81	83	83	**88**	+6.02%
South East Queensland (12)	77	80	81	81	82	**96**	+18.51%
Fukido (4)	87	88	88	89	89	**93**	+4.49%
Shiraho (7)	84	87	88	89	89	**96**	+7.87%

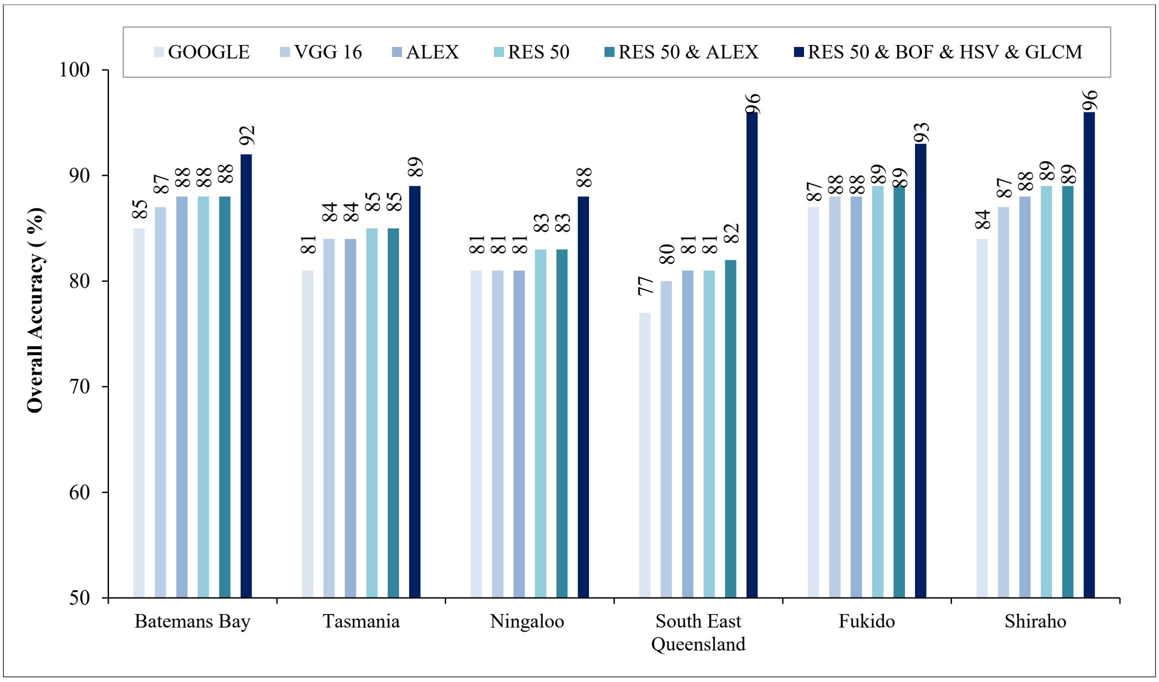


Figure 5. Comparison of overall accuracies of various pretrained CNNs and combined descriptors for all datasets

3.2. Benthic Habitats Semantic Segmentation

Four pretrained CNN architectures—ResNet-50, Xception [79] , AlexNet, and MobileNetV2 [80]—were evaluated for benthic habitat semantic segmentation. The corresponding Mean Pixel Accuracy (MPA), F1-score, and Mean Intersection over Union (M IoU) results for each network are presented in Table 4. Examples of segmented benthic habitat images produced by the evaluated CNNs are illustrated in Figure 6, while Figure 7 displays the confusion matrix of ResNet-50 based on an independent 30% testing subset.

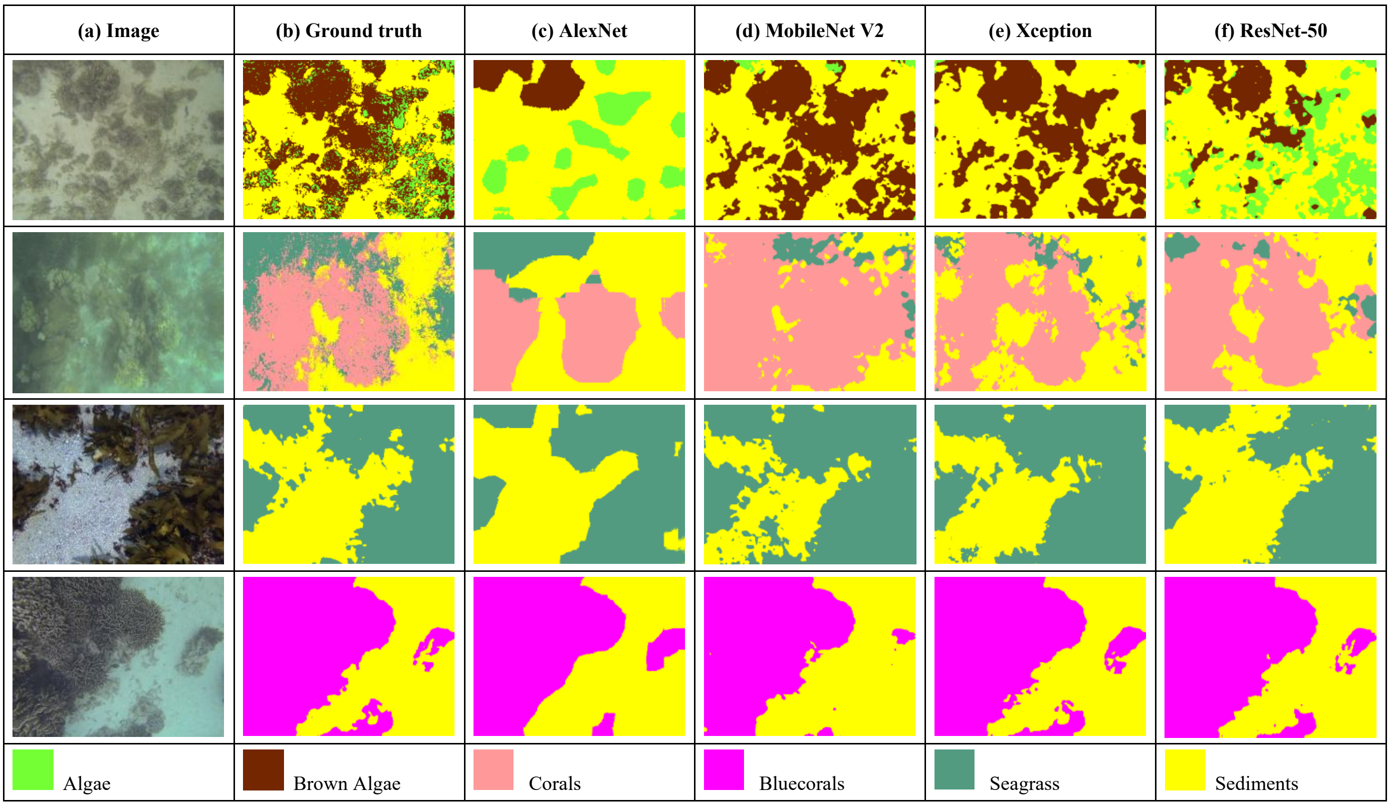


Figure 6. Comparison between the outputs from the evaluated pre-trained CNNs for benthic habitats semantic segmentation using an independent 30% subset of testing images

Table 4. Comparison of different pre-trained CNNs semantic segmentation performance

Pretrained CNNs	MPA	M IoU	F1-score
ResNet-50	**84**	**72**	**62**
Xception	79	67	57
MobileNet V2	78	63	56
AlexNet	68	50	42

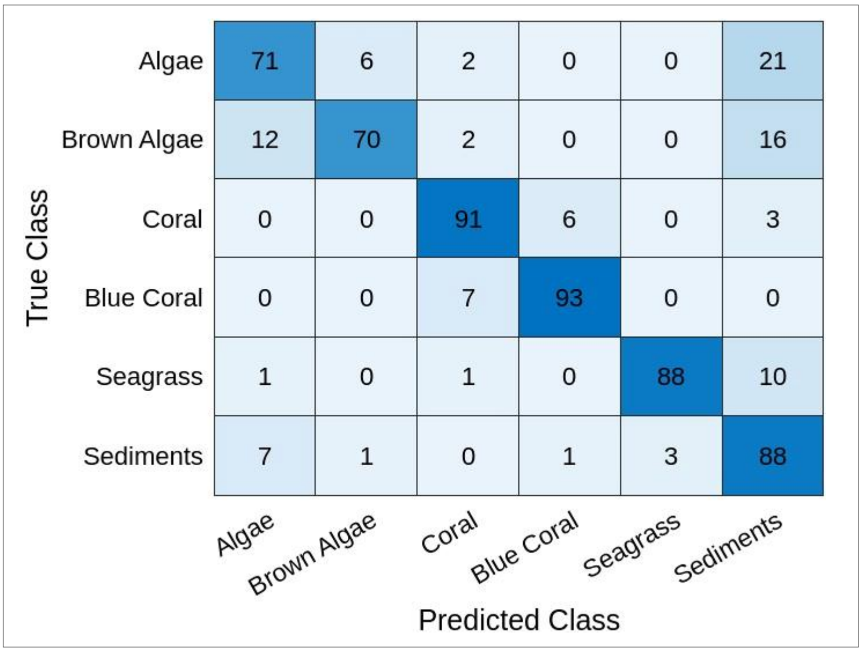


Figure 7. Confusion matrix of the ResNet-50 semantic segmentation

4. Discussion

From a theoretical standpoint, the design of the proposed framework is motivated by the inherent multi-scale and multi-modal nature of benthic habitats. These environments exhibit complex visual properties—fine-textured substrates, color-varied vegetation, and irregular three-dimensional structures—that cannot be comprehensively modeled by a single feature type. The integration of BOF, HSV, GLCM, and ResNet-50 features therefore reflects a theoretically grounded strategy for combining complementary representations: BOF and GLCM capture local geometric and textural cues, HSV provides illumination-stable chromatic information, and ResNet-50 descriptors encode deeper semantic features through residual learning. This layered representation is consistent with established theories in hierarchical feature fusion, which suggest that merging low-, mid-, and high-level descriptors enhances discrimination in visually heterogeneous settings. Furthermore, the two-stage design of the framework—global single-label classification followed by pixel-wise semantic segmentation—follows a theoretical paradigm in computer vision where coarse global context helps constrain and improve fine-grained spatial predictions. These theoretical considerations collectively explain the robustness and generalizability of the proposed approach across diverse underwater imaging conditions.

For benthic habitat classification, various attributes derived from shape, texture, and multiple color spaces (e.g., CIE XYZ, YIQ, and YCbCr) were assessed. Additionally, the performance of individual color dimensions (e.g., Hue or Saturation alone) from each tested color space was analyzed. However, these descriptors produced considerably lower overall accuracy. Principal Component Analysis (PCA) was also applied to reduce redundant features, but this led to a decline in overall accuracy across all experiments. Among the tested color spaces, the HSV descriptor achieved the highest accuracy, marginally outperforming YCbCr. It is noteworthy that color-based descriptors were more effective than shape or texture descriptors for high-quality AUV-acquired images (the four Australian datasets). In contrast, this relationship was reversed for low-quality and noisy images captured by the towed camera (Shiraho and Fukido datasets). These comparative results are presented in Table 2 and illustrated in Figure 4.

On the other hand, several layers from different CNN architectures—such as FC7 and FC8 from VGG19, FC8 from VGG16 and AlexNet, and combinations of these layers—were tested across all classification categories. However, these layers yielded relatively low overall accuracy (OA) values. The FC1000 layer from ResNet-50 achieved the highest performance among all tested layers, and the OA slightly improved when it was integrated with the FC7 layer from AlexNet. Moreover, the findings confirmed the superior performance of CNN-extracted features compared to traditional BOF shape and texture-based GLCM features, aligning with previous research [81, 82]. Also, the integrated feature configuration—combining ResNet-50 FC1000 descriptors with BOF, HSV color space, and GLCM texture attributes—

consistently achieved the highest classification accuracies across all datasets, as illustrated in Figure 5. The improvement over the best-performing single CNN baseline ranged from 4.5% to 17.1%, with the largest gain observed for the Southeast Queensland dataset, which contains the highest class diversity (12 classes), as shown in Table 3. This study further demonstrated that integrating diverse feature types outperformed approaches that relied solely on CNN descriptors, as reported in earlier works [83, 84]. These results highlight the complementary nature of the combined features, where BOF and GLCM capture local shape and texture variations, while HSV contributes color invariance, collectively enhancing the discriminative power of the CNN descriptors. The findings confirm that feature-level fusion substantially improves benthic habitat classification performance, particularly in complex or visually heterogeneous underwater environments. The obtained accuracy was comparable to results from previous studies [85-87], despite variations in the number of evaluated classes, water quality conditions, class composition, and image quality across datasets.

For benthic habitat semantic segmentation, several pre-trained CNN architectures—such as Inception, VGG16, and VGG19—were evaluated; however, these models produced relatively lower accuracy values. Multiple experiments were also conducted to determine the optimal training parameters for each CNN, with the maximum mini-batch size limited to 8 due to GPU memory constraints. Among all tested networks, ResNet-50 achieved notably higher Mean Pixel Accuracy (MPA), F1-score, and Mean Intersection over Union (M IoU) values, outperforming the other CNNs in segmenting benthic habitats, as presented in Table 4. The superior performance of ResNet architectures has also been validated in previous studies [63, 88, 89], primarily due to their residual learning framework, which effectively mitigates vanishing gradient issues and facilitates deeper feature extraction. As illustrated in Figure 6 and further confirmed by the confusion matrix in Figure 7, the lowest segmentation accuracies were recorded for brown algae and other algae, primarily due to their overlapping morphological and spectral characteristics. Moreover, all evaluated CNNs exhibited a tendency to misclassify algae as sediments, particularly in low-quality underwater imagery with high turbidity and uneven illumination. Such confusion underscores the persistent challenge of distinguishing visually similar benthic classes in optically degraded environments. In contrast, seagrasses, corals, and blue corals were segmented with high accuracy owing to their distinct shapes, colors, and textures. These findings highlight that both habitat complexity and image quality strongly influence CNN-based segmentation performance.

Furthermore, both the parametric paired t-test and its non-parametric counterpart, the Wilcoxon signed-rank test [90], were employed to evaluate the statistical significance of differences between the proposed single-label classification and segmentation methods for benthic habitats. In both tests, the decision is determined by the returned h-value, where h = 1 denotes rejection of the null hypothesis—indicating a statistically significant difference between compared algorithms—while h = 0 signifies no significant difference. The statistical significance analysis was conducted separately for each dataset to ensure an unbiased evaluation. The classification results revealed that only the integration of ResNet-50 FC1000, BOF, HSV, and GLCM descriptors yielded h = 1 at the 5% significance level for both tests, confirming a significant difference compared to other evaluated descriptors. Similarly, both statistical tests were applied to assess the segmentation results, where the null hypothesis was rejected (h = 1) exclusively for the ResNet-50 M IoU, F1-score, and MPA outcomes, demonstrating a significant difference between ResNet-50 and the other CNNs. Conversely, both tests failed to reject the null hypothesis for the remaining segmentation models, indicating no statistically significant differences among them.

The proposed framework demonstrates several key advantages. First, the trained SVM classifier can automatically categorize large volumes of underwater images acquired from various regions. Second, the classification system can be applied periodically to monitor temporal changes in complex marine ecosystems. Third, the framework achieves high accuracy with minimal logistical requirements, computational cost, and training data. Moreover, the proposed semantic segmentation module is fully automated, straightforward, accurate, and broadly applicable. However, certain limitations were identified. The segmentation performance decreased for images containing mixed substrates with visually similar classes or when the number of classes within a single image increased. Additionally, low-quality images with extensive shadowed areas remain challenging, as shadows significantly reduce segmentation accuracy. It should be noted that direct comparison between the accuracies achieved in this study and those reported in previous research is inherently challenging due to substantial variations in underwater image characteristics, water quality, and substrate diversity. Nevertheless, a review of recent literature reveals that previous studies have reported overall accuracies ranging from approximately 80% to 99%, depending on factors such as the number of evaluated classes, class composition, illumination and turbidity conditions, and image quality. Despite these inherent differences, the proposed framework achieved accuracies of 88–96% for single-label classification and 84% for semantic segmentation, results that are broadly comparable to—and in several cases competitive with—the best-performing methods reported in the literature. These outcomes highlight the robustness, generalizability, and efficiency of the proposed approach across diverse underwater environments and benthic habitat types.

Future research could address these issues by incorporating advanced image enhancement techniques to improve image quality and mitigate shadow effects. Furthermore, integrating CNN-based descriptors with 3D morphological features such as slope, roughness, and curvature may enhance benthic habitat detection from underwater image mosaics. Another promising direction involves ensemble learning by combining outputs from multiple CNN architectures to leverage their complementary strengths, potentially improving semantic segmentation accuracy. Finally, exploring new generations of CNN architectures may further enhance segmentation performance while reducing computational time and hardware requirements.

5. Conclusion

This study introduced an efficient and fully automated framework based on convolutional neural networks (CNNs) for the classification and semantic segmentation of benthic habitats from underwater imagery. The framework effectively integrates multiple complementary descriptors—namely, pre-trained ResNet-50 FC1000 features, Bag of Features (BOF), HSV color components, and Gray Level Co-occurrence Matrix (GLCM) texture measures— to enhance the discriminative power of benthic habitat image classification. Validation across six distinct study areas demonstrated that integrating these descriptors substantially improved the classification performance, achieving the highest overall accuracies and surpassing those of individual descriptors by a minimum margin of 4%. Statistical significance analysis confirmed that the integrated feature set consistently outperformed the individual descriptors at the 5% confidence level, emphasizing the effectiveness of multi-descriptor fusion for robust underwater image analysis. In the semantic segmentation stage, the ResNet-50 network exhibited superior predictive capability among the evaluated CNN architectures, yielding overall performance metrics of 84% MPA, 62% F1-score, and 72% M IoU. Significance testing of these results further verified the statistical superiority of ResNet-50 at the 5% significance level. Overall, these findings indicate that the proposed framework constitutes a robust and effective solution for large-scale benthic habitat mapping and monitoring. Its demonstrated capacity for accurate classification and detailed segmentation establishes a strong foundation for future developments in automated marine habitat assessment, with potential extensions involving more advanced deep architectures, adaptive enhancement methods, and multimodal data fusion to further improve performance under complex underwater conditions.

6. Declarations

6.1. Author Contributions

Conceptualization, H.M. and K.N.; methodology, H.M. and K.N.; software, H.M.; validation, K.N.; formal analysis, H.M.; investigation, H.M.; resources, K.N.; data curation, H.M.; writing—original draft preparation, H.M.; writing—review and editing, K.N.; visualization, H.M.; supervision, K.N.; project administration, K.N.; funding acquisition, K.N. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

Publicly available datasets were analyzed in this study. This data can be found here: The Australian benthic datasets is available at: http://imos-data.s3-website-ap-southeast-2.amazonaws.com/?prefix=IMOS/AUV/ (accessed on January 2026).

6.3. Funding

This research was financially supported partly by Nakamura Laboratory in Tokyo Institute of Technology and JSPS Grant-in-Aids for Scientific Research (No. 15H02268), and Science and Technology Research Partnership for Sustainable Development (SATREPS) program, Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA).

6.4. Conflicts of Interest

The authors declare no conflict of interest.

7. References

[1] Zhang, H., Gruen, A., & Li, M. (2022). Deep Learning for Semantic Segmentation of Coral Images in Underwater Photogrammetry. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 5(2), 343–350. doi:10.5194/isprs-annals-V-2-2022-343-2022.

[2] Agulles, M., Marbà, N., Duarte, C. M., & Jordà, G. (2024). Mediterranean seagrasses provide essential coastal protection under climate change. Scientific Reports, 14(1), 30269. doi:10.1038/s41598-024-81026-5.

[3] Higgins, E., Metaxas, A., & Scheibling, R. E. (2022). A systematic review of artificial reefs as platforms for coral reef research and conservation. PLOS ONE, 17(1), e0261964. doi:10.1371/journal.pone.0261964.

[4] James, R. K., Keyzer, L. M., van de Velde, S. J., Herman, P. M. J., van Katwijk, M. M., & Bouma, T. J. (2023). Climate change mitigation by coral reefs and seagrass beds at risk: How global change compromises coastal ecosystem services. Science of the Total Environment, 857, 159576. doi:10.1016/j.scitotenv.2022.159576.

[5] Ahmad, I., Guo, P., Zhao, M. X., Zhong, Y., Zheng, X. Y., Zhang, S. Q., Qiu, J. W., Shi, Q., Yan, H. Q., Tao, S. C., & Xu, L. J. (2024). Coral reefs of Pakistan: a comprehensive review of anthropogenic threats, climate change, and conservation status. Frontiers in Marine Science, 11. doi:10.3389/fmars.2024.1466834.

[6] Thirukanthan, C. S., Azra, M. N., Lananan, F., Sara', G., Grinfelde, I., Rudovica, V., Vincevica-Gaile, Z., & Burlakovs, J. (2023). The Evolution of Coral Reef under Changing Climate: A Scientometric Review. Animals, 13(5), 949. doi:10.3390/ani13050949.

[7] Ferretto, G., Vergés, A., Poore, A. G. B., Glasby, T. M., & Griffin, K. J. (2023). Habitat Provision and Erosion Are Influenced by Seagrass Meadow Complexity: A Seascape Perspective. Diversity, 15(2), 125. doi:10.3390/d15020125.

[8] Hossain, M. S., Muslim, A. M., Nadzri, M. I., Sabri, A. W., Khalil, I., Mohamad, Z., & Beiranvand Pour, A. (2021). Coral habitat mapping: a comparison between maximum likelihood, Bayesian and Dempster–Shafer classifiers. Geocarto International, 36(11), 1217–1235. doi:10.1080/10106049.2019.1637466.

[9] Zaki, N. H. M., Chong, W. S., Muslim, A. M., Reba, M. N. M., & Hossain, M. S. (2022). Assessing optimal UAV-data pre-processing workflows for quality ortho-image generation to support coral reef mapping. Geocarto International, 37(25), 10556-10580. doi:10.1080/10106049.2022.2037732.

[10] Devlin, M. J., Petus, C., & Oubelkheir, K. (2023). Special Issue Overview: Advances in Remote Sensing and Mapping for Integrated Studies of Reef Ecosystems in Oceania (Great Barrier Reef and Beyond). Remote Sensing, 15(10), 2505. doi:10.3390/rs15102505.

[11] Evans, S. N., Konzewitsch, N., Hovey, R. K., Kendrick, G. A., & Bellchambers, L. M. (2025). Selecting the best habitat mapping technique: a comparative assessment for fisheries management in Exmouth Gulf. Frontiers in Marine Science, 12. doi:10.3389/fmars.2025.1570277.

[12] Lee, C. B., Martin, L., Traganos, D., Antat, S., Baez, S. K., Cupidon, A., Faure, A., Harlay, J., Morgan, M., Mortimer, J. A., Reinartz, P., & Rowlands, G. (2023). Mapping the National Seagrass Extent in Seychelles Using PlanetScope NICFI Data. Remote Sensing, 15(18), 4500. doi:10.3390/rs15184500.

[13] Ginting, D. N. B., Wicaksono, P., & Farda, N. M. (2023). Mapping Benthic Habitat From Worldview-3 Image Using Random Forest Case Study: Nusa Lembongan, Bali, Indonesia. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-4/W6-2022, 123–129. doi:10.5194/isprs-archives-XLVIII-4-W6-2022-123-2023.

[14] Morsy, S., Suárez, A. B. Y., & Robert, K. (2023). 3D Mapping of Benthic Habitat Using XGBoost and Structure from Motion Photogrammetry. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-1/W1-2023, 1131–1136. doi:10.5194/isprs-annals-X-1-W1-2023-1131-2023.

[15] Mohamad, M. N., Reba, M. N. M., & Hossain, M. S. (2022). A screening approach for the correction of distortion in UAV data for coral community mapping. Geocarto International, 37(24), 7089–7121. doi:10.1080/10106049.2021.1958066.

[16] Marcaccio, J. V., Gardner Costa, J., Parker, S., & Midwood, J. D. (2025). High resolution satellite data and image segmentation produce accurate benthic substrate maps in clear waters of the great lakes. Applied Geomatics, 17(2), 343–356. doi:10.1007/s12518-025-00607-9.

[17] Ventura, D., Grosso, L., Pensa, D., Casoli, E., Mancini, G., Valente, T., Scardi, M., & Rakaj, A. (2023). Coastal benthic habitat mapping and monitoring by integrating aerial and water surface low-cost drones. Frontiers in Marine Science, 9. doi:10.3389/fmars.2022.1096594.

[18] Geisz, J. K., Wernette, P. A., & Esselman, P. C. (2024). Classification of Lakebed Geologic Substrate in Autonomously Collected Benthic Imagery Using Machine Learning. Remote Sensing, 16(7), 1264. doi:10.3390/rs16071264.

[19] Zhong, J., Li, M., Gruen, A., Gong, J., Li, D., Li, M., & Qin, J. (2024). Application of Photogrammetric Computer Vision and Deep Learning in High-Resolution Underwater Mapping: A Case Study of Shallow-Water Coral Reefs. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-2–2024, 247–254. doi:10.5194/isprs-annals-X-2-2024-247-2024.

[20] Hamylton, S. M. (2017). Mapping coral reef environments: A review of modern remote sensing methods and applications. Remote Sensing, 9(2), 118. doi:10.3390/rs9020118.

[21] Misiuk, B., & Brown, C. J. (2024). Benthic habitat mapping: A review of three decades of mapping biological patterns on the seafloor. Estuarine, Coastal and Shelf Science, 296, 1–27,. doi:10.1016/j.ecss.2023.108599.

[22] Mahmood, A., Ospina, A. G., Bennamoun, M., An, S., Sohel, F., Boussaid, F., Hovey, R., Fisher, R. B., & Kendrick, G. A. (2020). Automatic hierarchical classification of kelps using deep residual features. Sensors (Switzerland), 20(2), 1–20,. doi:10.3390/s20020447.

[23] Zotou, M., Sini, M., Trygonis, V., Greggio, N., Mazaris, A. D., & Katsanevakis, S. (2025). Towards Scalable Ecological Monitoring: Assessing AI-Based Annotation of Benthic Images. Journal of Marine Science and Engineering, 13(9), 1721. doi:10.3390/jmse13091721.

[24] Gómez-Ríos, A., Tabik, S., Luengo, J., Shihavuddin, A. S. M., Krawczyk, B., & Herrera, F. (2019). Towards highly accurate coral texture images classification using deep convolutional neural networks and data augmentation. Expert Systems with Applications, 118, 315–328. doi:10.1016/j.eswa.2018.10.010.

[25] Zhang, H., Li, M., Pan, X., Zhang, X., Zhong, J., & Qin, J. (2022). Novel approaches to enhance coral reefs monitoring with underwater image segmentation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVI-3/W1-2022, 271–277. doi:10.5194/isprs-archives-XLVI-3-W1-2022-271-2022.

[26] González-Rivero, M., Beijbom, O., Rodriguez-Ramirez, A., Holtrop, T., González-Marrero, Y., Ganase, A., Roelfsema, C., Phinn, S., & Hoegh-Guldberg, O. (2016). Scaling up ecological measurements of coral reefs using semi-automated field image collection and analysis. Remote Sensing, 8(1), 30. doi:10.3390/rs8010030.

[27] Halder, S., Islam, N., Ray, B., Andrews, E., Hettiarachchi, P., & Jackson, E. (2024). AI-based seagrass morphology measurement. Journal of Environmental Management, 369. doi:10.1016/j.jenvman.2024.122246.

[28] Hu, K. J., Pan, Y. T., Jiang, L. W., Lee, S. Der, & Kao, S. L. (2025). A robust underwater image enhancement algorithm. Journal of Supercomputing, 81(1), 244. doi:10.1007/s11227-024-06719-0.

[29] Xiao, C., Zhou, Z., & Hu, Y. (2025). A Lightweight Semantic Segmentation Model for Underwater Images Based on DeepLabv3+. Journal of Imaging, 11(5), 162. doi:10.3390/jimaging11050162.

[30] Paul, M. A., Kumar, K. S., Sagar, S., & Sreeji, S. (2023). LWDS: lightweight DeepSeagrass technique for classifying seagrass from underwater images. Environmental Monitoring and Assessment, 195(5), 614. doi:10.1007/s10661-023-11183-z.

[31] Langlois, L. A., Collier, C. J., & McKenzie, L. J. (2023). Subtidal seagrass detector: development of a deep learning seagrass detection and classification model for seagrass presence and density in diverse habitats from underwater photoquadrats. Frontiers in Marine Science, 10. doi:10.3389/fmars.2023.1197695.

[32] Jamil, S., Rahman, M., & Haider, A. (2021). Bag of features (Bof) based deep learning framework for bleached corals detection. Big Data and Cognitive Computing, 5(4), 53. doi:10.3390/bdcc5040053.

[33] Fallati, L., Saponari, L., Savini, A., Marchese, F., Corselli, C., & Galli, P. (2020). Multi-Temporal UAV Data and object-based image analysis (OBIA) for estimation of substrate changes in a post-bleaching scenario on a maldivian reef. Remote Sensing, 12(13), 2093. doi:10.3390/rs12132093.

[34] Nemani, S., Cote, D., Misiuk, B., Edinger, E., Mackin-McLaughlin, J., Templeton, A., Shaw, J., & Robert, K. (2022). A multi-scale feature selection approach for predicting benthic assemblages. Estuarine, Coastal and Shelf Science, 277. doi:10.1016/j.ecss.2022.108053.

[35] Awalludin, E. A., Hitam, M. S., Yussof, W. N. J. H. W., & Bachok, Z. (2017). Classification of coral reef components using color and texture features. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 9(3-4), 109-113..

[36] Srividhya, K., & Ramya, M. M. (2017). Object classification in underwater images using adaptive fuzzy neural network. 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 142–148. doi:10.1109/FSKD.2017.8392973.

[37] Jarina Raihan, A., Abas, P. E., & De Silva, L. C. (2021). Role of restored underwater images in underwater imaging applications. Applied System Innovation, 4(4). doi:10.3390/asi4040096.

[38] Ani Brown Mary, N., & Dharma, D. (2019). A novel framework for real-time diseased coral reef image classification. Multimedia Tools and Applications, 78(9), 11387–11425. doi:10.1007/s11042-018-6673-2.

[39] Ganesan, A., & Santhanam, S. M. (2022). A novel feature descriptor based coral image classification using extreme learning machine with ameliorated chimp optimization algorithm. Ecological Informatics, 68, 101527. doi:10.1016/j.ecoinf.2021.101527.

[40] Nazmi, Z., Shakoor, M. H., & Rahmani, M. (2023). Improved Multi-scale Local Binary Pattern for Feature Extraction and Coral Reef Classification. Journal of Machine Vision and Image Processing, 10(3), 31–45. doi:doi:20.1001.1.23831197.1402.10.3.2.4.

[41] Sotoodeh, M., Moosavi, M. R., & Boostani, R. (2019). A structural based feature extraction for detecting the relation of hidden substructures in coral reef images. Multimedia Tools and Applications, 78(24), 34513–34539. doi:10.1007/s11042-019-08050-w.

[42] Aung, T., Abdul Razak, R., & Rahiman Bin Md Nor, A. (2025). Artificial intelligence methods used in various aquaculture applications: A systematic literature review. Journal of the World Aquaculture Society, 56(1), e13107. doi:10.1111/jwas.13107.

[43] Zhang, J., Luo, F., Quan, X., Wang, Y., Shi, J., Shen, C., & Zhang, C. (2024). Improving wave height prediction accuracy with deep learning. Ocean Modelling, 188. doi:10.1016/j.ocemod.2023.102312.

[44] Mahmoud, A. S., Mohamed, S. A., Helmy, A. K., & Nasr, A. H. (2025). BDCN_UNet: Advanced shoreline extraction techniques integrating deep learning. Earth Science Informatics, 18(2), 187. doi:10.1007/s12145-024-01693-w.

[45] Xu, M., Chen, K., Guo, H., Huang, Y., Wu, M., Shi, Z., Zhang, C., & Guo, J. (2025). MFogHub: Bridging Multi-Regional and Multi-Satellite Data for Global Marine Fog Detection and Forecasting. 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 12637–12646. doi:10.1109/CVPR52734.2025.01179.

[46] Mittal, S., Srivastava, S., & Jayanth, J. P. (2023). A Survey of Deep Learning Techniques for Underwater Image Classification. IEEE Transactions on Neural Networks and Learning Systems, 34(10), 6968–6982. doi:10.1109/TNNLS.2022.3143887.

[47] Game, C. A., Thompson, M. B., & Finlayson, G. D. (2024). Machine learning for non-experts: A more accessible and simpler approach to automatic benthic habitat classification. Ecological Informatics, 81. doi:10.1016/j.ecoinf.2024.102619.

[48] Pavoni, G., Corsini, M., Pedersen, N., Petrovic, V., & Cignoni, P. (2021). Challenges in the deep learning-based semantic segmentation of benthic communities from Ortho-images. Applied Geomatics, 13(1), 131–146. doi:10.1007/s12518-020-00331-6.

[49] Fawad, Ahmad, I., Ullah, A., & Choi, W. (2023). Machine learning framework for precise localization of bleached corals using bag-of-hybrid visual feature classification. Scientific Reports, 13(1), 1–11. doi:10.1038/s41598-023-46971-7.

[50] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. doi:10.48550/arXiv.1409.1556.

[51] Yasir, M., Rahman, A. U., & Gohar, M. (2021). Habitat mapping using deep neural networks. Multimedia Systems, 27(4), 679–690. doi:10.1007/s00530-020-00695-0.

[52] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2261–2269. doi:10.1109/CVPR.2017.243.

[53] Chen, B., Zhao, W., Zhang, Q., Li, M., Qi, M., & Tang, Y. (2025). Semantic segmentation of underwater images based on the improved SegFormer. Frontiers in Marine Science, 12. doi:10.3389/fmars.2025.1522160.

[54] Mills, M. S., Ungermann, M., Rigot, G., den Haan, J., Leon, J. X., & Schils, T. (2023). Assessment of the utility of underwater hyperspectral imaging for surveying and monitoring coral reef ecosystems. Scientific Reports, 13(1), 21103. doi:10.1038/s41598-023-48263-6.

[55] Schürholz, D., & Chennu, A. (2023). Digitizing the coral reef: Machine learning of underwater spectral images enables dense taxonomic mapping of benthic habitats. Methods in Ecology and Evolution, 14(2), 596–613. doi:10.1111/2041-210X.14029.

[56] Šiaulys, A., Vaičiukynas, E., Medelytė, S., & Buškus, K. (2024). Coverage estimation of benthic habitat features by semantic segmentation of underwater imagery from South-eastern Baltic reefs using deep learning models. Oceanologia, 66(2), 286–298. doi:10.1016/j.oceano.2023.12.004.

[57] Raine, S., Marchant, R., Kusy, B., Maire, F., Sunderhauf, N., & Fischer, T. (2024). Human-in-the-Loop Segmentation of Multi-species Coral Imagery. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2723–2732. doi:10.1109/CVPRW63382.2024.00278.

[58] Liu, H., Chen, Y., Wang, R., Li, M., & Li, Z. (2025). MFA-Deeplabv3+: an improved lightweight semantic segmentation algorithm based on Deeplabv3+. Complex and Intelligent Systems, 11(10), 424. doi:10.1007/s40747-025-02028-y.

[59] Jagadish, S., S, P. K., Riad Al-Fatlawy, R., Ameer, S. A., & J, K. (2024). Random Forest based Coral Reef Health Classification Using Satellite Imagery and Geographical Information System. 2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), 1–5. doi:10.1109/iacis61494.2024.10722008.

[60] de Oliveira, L. M. C., Lim, A., Conti, L. A., & Wheeler, A. J. (2022). High-resolution 3D mapping of cold-water coral reefs using machine learning. Frontiers in Environmental Science, 10. doi:10.3389/fenvs.2022.1044706.

[61] Nieuwenhuis, B. O., Marchese, F., Casartelli, M., Sabino, A., van der Meij, S. E. T., & Benzoni, F. (2022). Integrating a UAV-Derived DEM in Object-Based Image Analysis Increases Habitat Classification Accuracy on Coral Reefs. Remote Sensing, 14(19). doi:10.3390/rs14195017.

[62] Zhong, J., Li, M., Zhang, H., & Qin, J. (2023). Fine-Grained 3D Modeling and Semantic Mapping of Coral Reefs Using Photogrammetric Computer Vision and Machine Learning. Sensors, 23(15), 6753. doi:10.3390/s23156753.

[63] Zuo, X., Jiang, J., Shen, J., & Yang, W. (2025). Improving underwater semantic segmentation with underwater image quality attention and muti-scale aggregation attention. Pattern Analysis and Applications, 28(2), 80. doi:10.1007/s10044-025-01460-7.

[64] Bewley, M., Friedman, A., Ferrari, R., Hill, N., Hovey, R., Barrett, N., Marzinelli, E. M., Pizarro, O., Figueira, W., Meyer, L., Babcock, R., Bellchambers, L., Byrne, M., & Williams, S. B. (2015). Australian sea-floor survey data, with images and expert annotations. Scientific Data, 2. doi:10.1038/sdata.2015.57.

[65] Johnson-Roberson, M., Pizarro, O., Williams, S. B., & Mahon, I. (2010). Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys. Journal of Field Robotics, 27(1), 21–51. doi:10.1002/rob.20324.

[66] Althaus, F., Hill, N., Ferrari, R., Edwards, L., Przeslawski, R., Schönberg, C. H. L., Stuart-Smith, R., Barrett, N., Edgar, G., Colquhoun, J., Tran, M., Jordan, A., Rees, T., & Gowlett-Holmes, K. (2015). A standardised vocabulary for identifying benthic biota and substrata from underwater imagery: The CATAMI classification scheme. PLoS ONE, 10(10), 0141039. doi:10.1371/journal.pone.0141039.

[67] CATAMI (2026). Collaborative and Annotation Tools for Analysis of Marine Imagery (CATAMI), Rome, Italy. Available online: https://catami.org/ (accessed on January 2026).

[68] Bravo, G., Trobbiani, G. A., Bigatti, G., Beltramino, L. E., & Irigoyen, A. J. (2025). Towed Video-Diver: A Useful Low-Cost Tool for Rapid Benthic Mapping and Biodiversity Monitoring. Ecologies, 6(1), 10. doi:10.3390/ecologies6010010.

[69] Lowe, S. C., Misiuk, B., Xu, I., Abdulazizov, S., Baroi, A. R., Bastos, A. C., Best, M., Ferrini, V., Friedman, A., Hart, D., Hoegh-Guldberg, O., Ierodiaconou, D., Mackin-McLaughlin, J., Markey, K., Menandro, P. S., Monk, J., Nemani, S., O'Brien, J., Oh, E., … Trappenberg, T. (2025). BenthicNet: A global compilation of seafloor images for deep learning applications. Scientific Data , 12(1), 230. doi:10.1038/s41597-025-04491-1.

[70] Adams, L. A., Karenyi, N., Parker, D., & Sink, K. (2023). Patterns and potential drivers of mesophotic communities of the warm-temperate Amathole shelf of South Africa. Estuarine, Coastal and Shelf Science, 295. doi:10.1016/j.ecss.2023.108562.

[71] Goldman, J. (213). GoPro Hero3+ Black Edition review: Top-of-the-line action CAM improves on excellent predecessor. CNET Networks, San Francisco, United States. Available online: https://www.cnet.com/reviews/gopro-hero3-plus-black-edition-review/ (accessed on January 2026).

[72] da Silveira, C. B. L., Strenzel, G. M. R., Maida, M., Gaspar, A. L. B., & Ferreira, B. P. (2021). Coral reef mapping with remote sensing and machine learning: A nurture and nature analysis in marine protected areas. Remote Sensing, 13(15). doi:10.3390/rs13152907.

[73] Latumakulita, L. A., Astawa, I. N. G. A., Mairi, V. G., Purnama, F., Wibawa, A. P., Jabari, N. A. M., & Islam, N. (2022). Combination of Feature Extractions for Classification of Coral Reef Fish Types Using Backpropagation Neural Network. International Journal on Informatics Visualization, 6(3), 643–649. doi:10.30630/joiv.6.3.1082.

[74] Nazir, S., Yousaf, M. H., & Velastin, S. A. (2018). Evaluating a bag-of-visual features approach using spatio-temporal features for action recognition. Computers and Electrical Engineering, 72, 660–669. doi:10.1016/j.compeleceng.2018.01.037.

[75] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. doi:10.1109/CVPR.2016.90.

[76] Yuval, M., Alonso, I., Eyal, G., Tchernov, D., Loya, Y., Murillo, A. C., & Treibitz, T. (2021). Repeatable semantic reef-mapping through photogrammetry and label-augmentation. Remote Sensing, 13(4), 1–19. doi:10.3390/rs13040659.

[77] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9. doi:10.1109/CVPR.2015.7298594.

[78] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84–90. doi:10.1145/3065386.

[79] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-ResNet and the impact of residual connections on learning. 31st AAAI Conference on Artificial Intelligence, AAAI 2017, 42, 4278–4284. doi:10.1609/aaai.v31i1.11231.

[80] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint, arXiv:1704.04861. doi:10.48550/arXiv.1704.04861.

[81] Sharan, S., Kininmonth, S., & Mehta, U. V. (2021). Automated CNN based coral reef classification using image augmentation and deep learning. International Journal of Engineering Intelligent Systems, 29(4), 253-261.

[82] Zhou, Z., Yang, X., Ji, H., & Zhu, Z. (2023). Improving the classification accuracy of fishes and invertebrates using residual convolutional neural networks. ICES Journal of Marine Science, 80(5), 1256–1266. doi:10.1093/icesjms/fsad041.

[83] Feng, Z., Zhang, P., & Huo, X. (2023). Underwater Target Classification Based on Feature Fusion and Gene Encoding of CNN-BIGRU-Attention. IEEE Access, 11, 139546–139556. doi:10.1109/ACCESS.2023.3341499.

[84] Xu, L., Bennamoun, M., An, S., Sohe, F., & Boussaid, F. (2018). Classification of Corals in Reflectance and Fluorescence Images Using Convolutional Neural Network Representations. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 1493–1497. doi:10.1109/ICASSP.2018.8462574.

[85] Mehrunnisa, Leszczuk, M., Juszka, D., & Zhang, Y. (2025). Improved Binary Classification of Underwater Images Using a Modified ResNet-18 Model. Electronics (Switzerland), 14(15), 2954. doi:10.3390/electronics14152954.

[86] Firdous, R. J., & Sabena, S. (2023). Collaborative CNN with Multiple Tuning for Automated Coral Reef Classification. Computer, Communication, and Signal Processing. AI, Knowledge Engineering and IoT for Smart Systems. ICCCSP 2023, IFIP Advances in Information and Communication Technology, 670. Springer, Cham., Switzerland. doi:10.1007/978-3-031-39811-7_7.

[87] Mahmood, A., Bennamoun, M., An, S., Sohel, F., & Boussaid, F. (2020). ResFeats: Residual network based features for underwater image classification. Image and Vision Computing, 93, 103811–103818,. doi:10.1016/j.imavis.2019.09.002.

[88] Balado, J., Olabarria, C., Martínez-Sánchez, J., Rodríguez-Pérez, J. R., & Pedro, A. (2021). Semantic segmentation of major macroalgae in coastal environments using high-resolution ground imagery and deep learning. International Journal of Remote Sensing, 42(5), 1785–1800. doi:10.1080/01431161.2020.1842543.

[89] Song, H., Mehdi, S. R., Zhang, Y., Shentu, Y., Wan, Q., Wang, W., Raza, K., & Huang, H. (2021). Development of coral investigation system based on semantic segmentation of single-channel images. Sensors, 21(5), 1–21. doi:10.3390/s21051848.

[90] Hooshmand Moghaddam, V., & Hamidzadeh, J. (2016). New Hermite orthogonal polynomial kernel and combined kernels in Support Vector Machine classifier. Pattern Recognition, 60, 921–935. doi:10.1016/j.patcog.2016.07.004.