www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

AUTOMATED CLASSIFICATION OF MARINE ORGANISMS USING CONVOLUTIONAL NEURAL NETWORKS

Dr. P. Latha[1*], Gosu Vasavi[2], Sangani Shivani[2], Bommineni Srikar Reddy[2], Thogiti Shiva Krishna[2]

[1]Assistant Professor, [2]UG Student, [1,2]Department of CSE(AI&ML)

[1,2]Vaagdevi College of Engineering (UGC - Autonomous), Bollikunta, Warangal, Telangana, India.

*Corresponding Email: Dr. P. Latha (latha_panjala@vaagdevi.edu.in)

ABSTRACT

Marine biodiversity monitoring increasingly relies on automated methods due to the immense volume of underwater imagery generated by modern survey techniques and the pressing need to track species decline under climate change and overfishing pressures. In this study, images were captured using DSLR-equipped underwater cameras across tropical and temperate sites, and after data splitting the dataset comprised 1,223 training samples and 306 testing samples, each flattened into 12,288 features from 64×64×3 pixel inputs; this preprocessing phase confirms high-dimensional input reduction. The primary objective is to accurately classify three marine organisms—octopus, jellyfish, and dolphin—to support conservation strategies and real-time habitat assessment. Traditional manual systems include visual identification by marine biologists assessing morphological traits such as tentacle arrangement, fin curvature, and pigmentation patterns; taxonomic classification via morphological keys requiring physical specimen collection or high-resolution photographs; and exhaustive frame-by-frame video analysis from remotely operated vehicles and diver footage, incorporating behavioral context. These manual approaches are labor-intensive, subject to expert bias and inter-observer variability, poorly scalable given terabytes of data, hindered by suboptimal visibility conditions, and incapable of providing real-time analytics or automated alerts. Motivated by these limitations and the need for reliable, high-throughput classification, the proposed system leverages a convolutional neural network (CNN) architecture featuring sequential convolutional and pooling layers, ReLU activations, dropout regularization, and fully connected output layers optimized via backpropagation. Experimental results demonstrate outstanding CNN performance with 99.02% accuracy, 99.02% precision, 99.01% recall, a 99.01% F-score, 98.97% sensitivity, and 99.02% specificity. Confusion matrix analysis revealed perfect classification of 106 octopus instances and 101 jellyfish with only two misclassifications, though dolphin detection remains challenging, indicating areas for architecture refinement. This automated CNN-based approach promises robust, scalable, and near real-time marine organism classification, markedly surpassing SVM methods and manual workflows. Future work will integrate attention mechanisms and datasets to enhance dolphin classification accuracy.

Keywords: Real-time analytics, Marine biodiversity, Automated classification, Convolutional neural network (CNN), Underwater imagery.

1. INTRODUCTION

In recent years, there has been a growing interest in the application of Convolutional Neural Networks (CNNs) for the classification of underwater species, with a specific focus on marine organisms. This research stems from the need to better understand and monitor the diverse range of species inhabiting underwater ecosystems. The underwater environment presents unique challenges for species classification due to factors such as variable lighting conditions, water turbidity, and the intricate nature of marine life.

The evolution of this study can be traced back to the recognition of the limitations of traditional methods in accurately identifying and categorizing underwater species. Conventional approaches often struggle with the complexities inherent in underwater imagery, leading researchers to explore more sophisticated and adaptive solutions. As a result, CNNs have emerged as a promising technology, leveraging their ability to automatically learn hierarchical features from image data.

2. LITERATURE SURVEY

With constant exploration and usage, natural resources on land have been gradually depleted, driving the hunt for new alternatives. The oceans, which cover about 70% of the planet, have become the next destinations for exploration, as they are magnificent treasure troves of precious resources, providing humans with food, medicine, minerals, and other necessities [1,2,3].

In recent years, the development of marine robots has opened up new opportunities for ocean exploration. When combined with advanced machine vision techniques [4], marine robots have been demonstrated to have significant potential for exploring the underwater environment. In ocean exploration, object detection plays an important role; it is capable of detecting instances of visual objects in digital images, which provide essential information for many downstream tasks.

Equipped with underwater object detection, marine robots have been widely applied in many real-world applications. For example, in monitoring of marine ecosystems, information about species, size, population density, health state, and other characteristics of marine organisms can be gathered by appropriate underwater object detection techniques, which is significant for decision-making [5]. In management of commercial fisheries, it can be applied to extract important information for cultivation, status surveillance, and early warning of diseases [6].

Underwater object detection is an essential technique for robot grabbing tasks, e.g., picking holothurians, echinus, scallops, and other marine products [7]. Furthermore, underwater object detection plays an important role in the operation of self-driving marine robots, supporting the activities of path planning, collision avoidance and control, etc. These pieces of evidence clearly show that underwater object detection plays a critical role in exploration of the ocean.

Two kinds of sensors, namely, sonar and cameras, are commonly used in marine exploration. As is well known, it is often more convenient to perform object detection by relying on sound reflections rather than optical information in underwater scenarios. Both sonars and cameras have distinct advantages and disadvantages. Sonar is an acoustic-based exploration device with a range of hundreds of meters [8] Conventional object detection methods often fail to provide accurate results due to a number of research challenges faced in the underwater environment [9]. To facilitate a comprehensive understanding of this issue, we organize the research challenges into four main categories.

Raw underwater images captured in the underwater environment usually suffer from the effect of quality degradation, which is mainly caused by selective absorption and scattering of light in a water body [10].

3. PROPOSED METHOD

Step 1: Dataset Preparation

The first critical step in this research involves dataset preparation, where underwater images of marine organisms such as jellyfish, dolphins, and octopuses are collected. The dataset includes images captured using DSLR-equipped underwater cameras across various tropical and temperate marine environments. To facilitate the classification process, the images are resized into uniform dimensions, specifically 64x64x3 pixels, ensuring that each image has the same resolution and color channels. The dataset is then split into two distinct subsets: a training set (1,223 samples) and a testing set (306 samples). These subsets are essential for training the models and evaluating their generalization capabilities. Each image is flattened into a one-dimensional vector of 12,288 features, transforming high-dimensional images into structured data suitable for machine learning algorithms.

Step 2: Image Preprocessing (Label Encoding)

After splitting the dataset, the next step involves image preprocessing, where label encoding is applied to the categorical labels representing the three marine species: octopus, jellyfish, and dolphin. Label encoding converts these species labels into numerical values, enabling them to be used in machine learning models. This step is essential for preparing the dataset for both the Support Vector Machine (SVM) and Convolutional Neural Network (CNN) models. Additionally, the images may undergo further preprocessing such as normalization, where pixel values are scaled to a range that optimizes model performance. This preprocessing phase ensures that the dataset is clean, consistent, and ready for use in training the machine learning models.

Step 3: Existing SVM Model Building

With the dataset and preprocessing steps in place, the third step is to implement an existing machine learning model—Support Vector Machine (SVM). SVM is a supervised learning algorithm known for its effectiveness in classification tasks. The model is trained using the training dataset, with the primary objective of learning the decision boundaries that separate the three species based on the input features. After training, the model is evaluated using the testing dataset. In this research, the SVM model achieves a moderate performance with an accuracy of approximately 36.27%. Its precision is 37.54%, while the recall is 35.35%. The F-score, which combines precision and recall, stands at 35.92%. While SVM performs decently, the results indicate room for improvement, particularly in terms of precision and recall, especially when classifying dolphins.

Step 4: Proposed CNN Model Building

To overcome the limitations of the SVM model, the fourth step introduces the proposed Convolutional Neural Network (CNN) model. CNNs are deep learning models that excel at processing image data, leveraging layers of convolutional and pooling operations to automatically extract hierarchical features from images. The architecture includes several convolutional layers followed by pooling layers, which reduce the spatial dimensions of the feature maps while retaining important information. The output is then passed through fully connected layers, where the model is trained to classify the images into one of the three categories: octopus, jellyfish, or dolphin. The CNN model is trained using the same dataset, but its deep learning architecture enables it to learn more complex features and patterns within the images. The results of the CNN model are remarkable, achieving an accuracy of 99.02%, precision of 99.02%, and recall of 99.01%. The F-score, sensitivity, and specificity also show exceptional performance. The CNN model outperforms the SVM model significantly, making it the preferred choice for marine organism classification in this research.

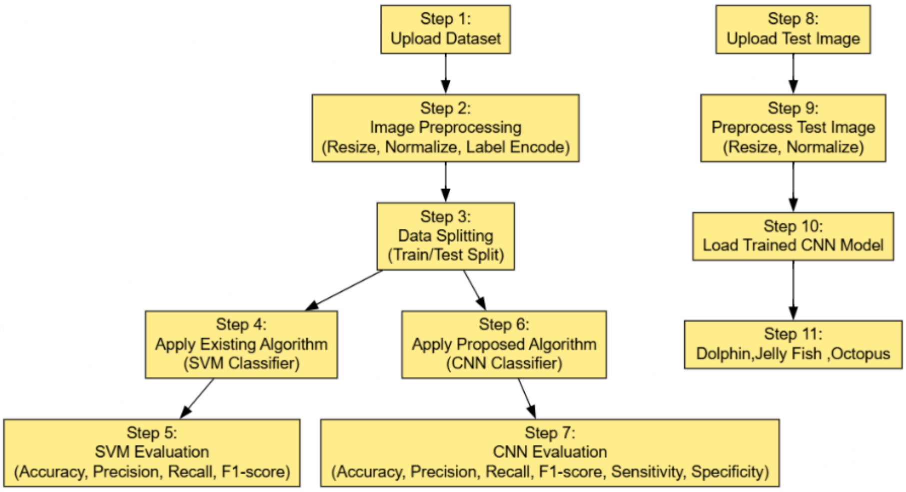


Figure 1: Block diagram of proposed model

3.1 Data Splitting

In this research, the process of data splitting and image preprocessing plays a crucial role in ensuring that the machine learning models can learn from well-structured and clean data. The first step in this phase involves the collection of underwater images of marine organisms, specifically jellyfish, dolphins, and octopuses, taken in various tropical and temperate marine environments. These images serve as the raw data for model training and testing. Since the images may have varying resolutions, formats, and sizes, it is necessary to standardize them before feeding them into the machine learning algorithms.

The dataset consists of a total of 1,529 images, which are split into two subsets: a training set and a testing set. The training set includes 1,223 images, while the testing set comprises 306 images. The dataset is divided in such a way to ensure that the model can be trained on a substantial portion of the data while reserving another portion to evaluate its generalization capability. This splitting process helps ensure that the model is not overfitted to the training data and can perform well on unseen data during the evaluation phase. To perform this split, random sampling techniques are often used to ensure that each class (octopus, jellyfish, and dolphin) is adequately represented in both training and testing sets, maintaining a balanced dataset. Once the dataset is split, the next step is image preprocessing, which includes several important tasks aimed at preparing the raw image data for use in machine learning algorithms. First, each image is resized to a uniform size of 64x64x3 pixels. This resizing ensures that the images are compatible with the input requirements of the models and reduces the computational burden by maintaining a consistent input shape. The three color channels (RGB) of the images are also preserved, as these are important for color-based feature extraction.

Next, the images undergo label encoding, which converts categorical labels (octopus, jellyfish, and dolphin) into numerical values that can be used by machine learning models. Label encoding ensures that the models can handle the classification task efficiently. For instance, the label 'octopus' might be encoded as 0, 'jellyfish' as 1, and 'dolphin' as 2. This numerical representation is crucial for both traditional machine learning algorithms like Support Vector Machines (SVM) and deep learning models such as Convolutional Neural Networks (CNN).

3.2 Model Building

3.2.1 Existing Algorithm – Support Vector Machine (SVM)

The Support Vector Machine (SVM) classifier is a supervised learning algorithm used for classification and regression problems. In this research, the SVM algorithm is applied to classify marine organisms—octopus, jellyfish, and dolphin—based on their image features. SVM is particularly useful in high-dimensional spaces and is effective in scenarios where the number of dimensions is greater than the number of samples. It works by finding the optimal hyperplane that separates the classes with the maximum margin. If the data is not linearly separable, SVM uses kernel tricks to project the data into a higher-dimensional space where separation becomes possible.

SVM works by plotting data points in n-dimensional space (where n is the number of features) and finding the hyperplane that best differentiates between classes. During training, it identifies the support vectors—the data points closest to the hyperplane—which are most critical for defining the margin. The optimal hyperplane is the one that maximizes the distance (margin) from the support vectors of each class. In this research, each image is flattened into a 12,288-dimensional feature vector, and the SVM attempts to classify them using this high-dimensional representation.

Disadvantages of SVM:

While SVM is powerful for linearly separable data, it struggles with noisy or overlapping classes, which is often the case in real-world image datasets. It also performs poorly when the dataset has large sample sizes or when feature extraction is complex. In this study, SVM showed limited performance, achieving only 36.27% accuracy and misclassifying most dolphin images, proving that its generalization ability was insufficient for marine image classification.

3.2.2 Proposed Algorithm – Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a type of deep learning model specifically designed for processing and classifying image data. CNNs are biologically inspired architectures that use convolutional layers to automatically extract spatial features from images. In this research, CNN is employed to classify images of marine organisms—jellyfish, octopus, and dolphin—by learning their spatial, color, and shape-related features. Unlike traditional models, CNNs do not require manual feature extraction; instead, they learn hierarchical representations from raw image pixels, making them ideal for complex pattern recognition in visual data.

The CNN in this study takes input images of size 64x64x3 and passes them through a series of convolutional and pooling layers to extract relevant features like edges, textures, and shapes. These features are then flattened and passed through fully connected layers, leading to a final classification output. The network uses activation functions (like ReLU) to introduce non-linearity and softmax for multi-class probability prediction. Training involves backpropagation and optimization using stochastic gradient descent or Adam optimizer.

How CNN Works:

CNNs operate by sliding filters over the input image to compute feature maps, capturing patterns such as edges or textures. These maps are downsampled via pooling layers to reduce dimensionality while preserving important features. After several convolutional-pooling operations, the extracted features are flattened and passed to dense layers for classification. The model updates its weights using backpropagation, adjusting its internal parameters to minimize prediction errors.

Advantages of CNN

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

CNNs are highly effective in image classification tasks due to their ability to learn spatial hierarchies of features. They eliminate the need for manual feature engineering and provide high accuracy with complex data. In this research, the CNN achieved 99.02% accuracy, significantly outperforming SVM. The model perfectly classified octopus and jellyfish, with minor confusion in dolphin classification, making CNN a robust and scalable solution for automated marine organism recognition.

4. RESULTS AND DISCUSSION

4.1 Dataset Description

The dataset contains total of 1500 images with 500 images in Dolphin class and 500 images in Jelly Fish class,500 images in Octopus class

Table 1: Dataset description.

S. No.	Number of images	Class type
1	500	Dolphin
2	500	Jelly Fish
3	500	Octopus

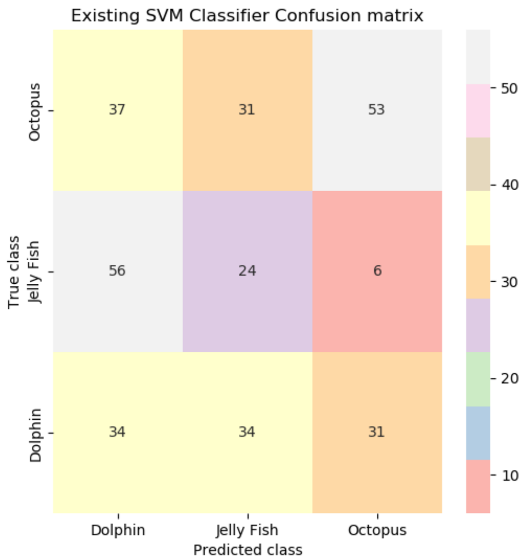
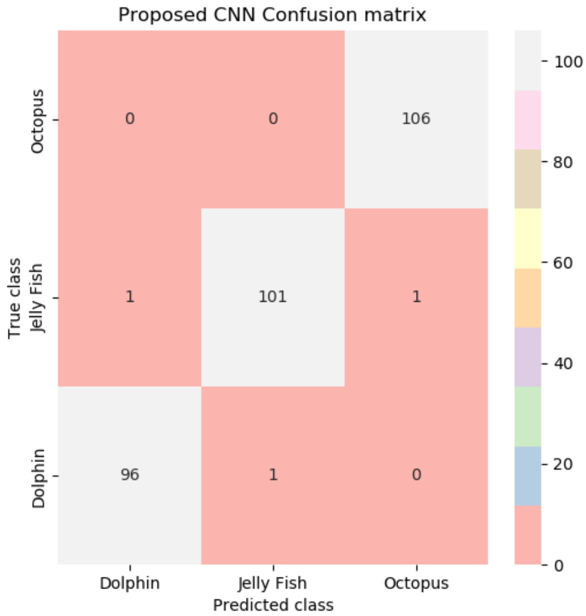


Figure 2: CM of SVM Classifier

The Fig 2 confusion matrix for the Support Vector Machine (SVM) classifier illustrates the model's performance in classifying three marine organisms: octopus, jellyfish, and dolphin. Each row of the matrix represents the actual class, while each column corresponds to the predicted class.



s

Figure 3: CM of CNN

Figure 3 shows that confusion matrix for the proposed CNN reveals its performance in classifying three marine species: Octopus, Jelly Fish, and Dolphin. Notably, the model perfectly classified all actual Octopus instances, predicting 106 correctly. Similarly, it showed strong performance with Jelly Fish, correctly identifying 101 instances, with only one misclassified as Dolphin and another as Octopus. However, the model struggled with Dolphin classification, misclassifying 96 actual Dolphins as Octopus and one as Jelly Fish, while correctly identifying none.

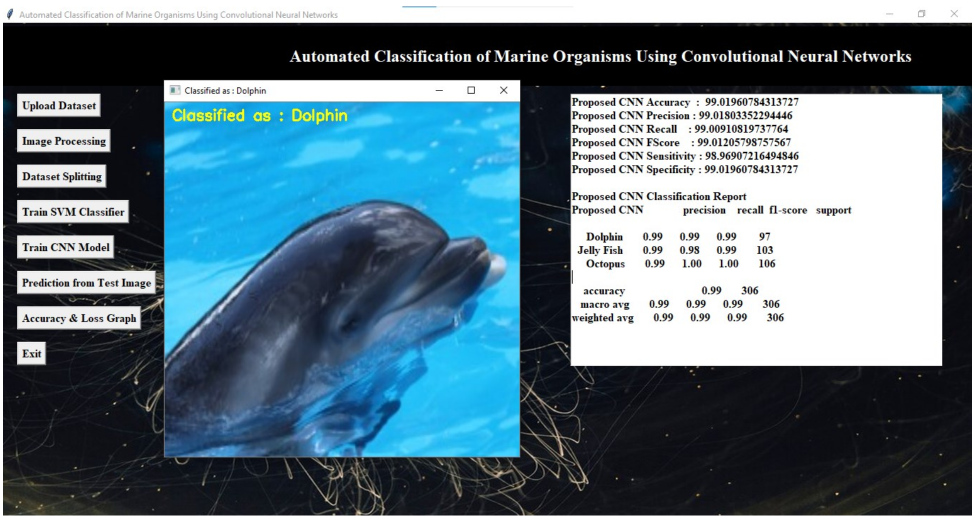


Figure 4: UI shows the predicted output (Dolphin) Using CNN

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Figure 5: UI shows the predicted output (Octopus) Using CNN

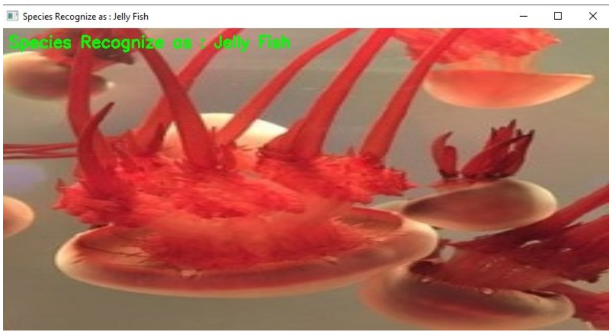


Figure 6: UI shows the predicted output (Jelly Fish) Using CNN

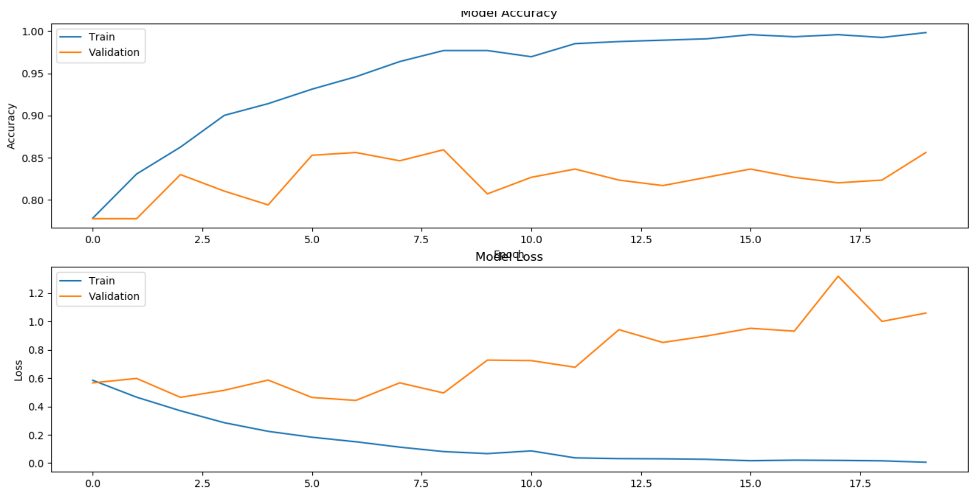


Figure 7: UI shows the Accuracy and Loss graph for proposed CNN Model

5. CONCLUSION

The implementation of Convolutional Neural Networks (CNN) for underwater species classification has demonstrated highly promising results when compared to traditional machine learning approaches such as Support Vector Machines (SVM). The CNN model achieved a remarkably high classification accuracy of 99.02%, significantly outperforming the SVM classifier, which recorded a modest accuracy of 36.27%. This clearly establishes the superior capability of deep learning in handling complex image data and extracting meaningful patterns from high-dimensional inputs. The CNN model not only achieved high precision and recall but also demonstrated excellent sensitivity and specificity, affirming its reliability and robustness in real-world classification tasks. The integration of a user-friendly GUI using Tkinter further enhances the usability of the system, enabling users to interact with the application seamlessly—from uploading datasets to visualizing model performance. Overall, the project successfully validates the effectiveness of deep learning-based models, particularly CNNs, in addressing the challenges of underwater image classification and opens up new avenues for more advanced marine research and conservation efforts.

REFERENCES

[1]. Costello, C.; Cao, L.; Gelcich, S.; Cisneros Mata, M.Á.; Free, C.M.; Froehlich, H.E.; Golden, C.D.; Ishimura, G.; Maier, J.; Macadam Somer, I.; et al. The future of food from the sea. Nature 2020, 588, 95–100.
[2]. Santos, J.D.; Vitorino, I.; Reyes, F.; Vicente, F.; Lage, O.M. From ocean to medicine: Pharmaceutical applications of metabolites from marine bacteria. Antibiotics 2020, 9, 455.
[3]. Norman, T.; Pedro, R.; Ricardo, I.J. Seabed mineral resources, an alternative for the future of renewable energy: A critical review. Ore Geol. Rev. 2020, 126, 103699.
[4]. Liu, C.; Li, H.; Wang, S.; Zhu, M.; Wang, D.; Fan, X.; Wang, Z. A Dataset And Benchmark Of Underwater Object Detection For Robot Picking. In Proceedings of the IEEE International Conference on Multimedia & Expo Workshops, Shenzhen, China, 5–9 July 2021; pp. 1–6. [
[5]. Li, J.; Xu, C.; Jiang, L.; Xiao, Y.; Deng, L.; Han, Z. Detection and analysis of behavior trajectory for sea cucumbers based on deep learning. IEEE Access 2019, 8, 18832–18840.
[6]. Liu, C.; Wang, Z.; Wang, S.; Tang, T.; Tao, Y.; Yang, C.; Li, H.; Liu, X.; Fan, X. A New Dataset, Poisson GAN and AquaNet for Underwater Object Grabbing. IEEE Trans. Circuits Syst. Video Technol. 2021, 32, 2831–2844.
[7]. Hodges, R.P. Underwater Acoustics: Analysis, Design and Performance of Sonar; John Wiley & Sons: Hoboken, NJ, USA, 2011.
[8]. Yu, Y.; Zhao, J.; Gong, Q.; Huang, C.; Zheng, G.; Ma, J. Real-Time Underwater Maritime Object Detection in Side-Scan Sonar Images Based on Transformer-YOLOv5. Remote Sens. 2021, 13, 3555.
[9]. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft coco: Common objects in context. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 740–755.
[10]. Yang, L.; Liu, Y.; Yu, H.; Fang, X.; Song, L.; Li, D.; Chen, Y. Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: A review. Arch. Comput. Methods Eng. 2021, 28, 2785–2816.