Marine Species Recognition Using Image Analysis

P.Dinesh Kumar[1], G.Ananthi[2], N.Chaithanya[3], K.Keerthana[4], R.Nikitha[5]
[1]Assistant Professor,Dept of CSE
[2, 3, 4, 5]Dept of CSE
[1, 2, 3, 4, 5] Vivekanandha College of Technology for Women

Abstract- *The identification and classification of marine species play a vital role in ecological studies, biodiversity protection, and sustainable fisheries management. Traditional approaches that rely on manual observation are not only time-intensive but also susceptible to errors due to human limitations and challenging underwater conditions. These limitations underscore the need for automated, intelligent solutions.*

The core of this approach combines CNNs with high-performance object detection algorithms such as YOLOv8. These models are trained to recognize distinct marine species from underwater images. Since such images often suffer from poor lighting and color distortion, the system integrates a preprocessing unit that applies image sharpening, contrast boosting, and noise removal to enhance visual quality.A domain-specific dataset comprising labeled images of various marine species was curated to train and validate the system. Multiple deep learning architectures were tested namely ResNet, YOLOv8, EfficientNet, and Faster R-CNN. Evaluation metrics such as accuracy, F1-score, precision, and mean Average Precision (mAP) were used to assess performance.Evaluation metrics such as recall, precision, accuracy, and mAP validate the effectiveness of the method. This project proves that AI and deep learning can significantly improve marine data collection, reduce human error, and enable scalable biodiversity monitoring solutionsThis work demonstrates that integrating AI with marine biology not only automates species monitoring but also supports large-scale conservation efforts and ecological analysis.

Keywords- C Marine species classification, underwater image processing, deep learning, CNNs, object detection, biodiversity monitoring.

I. INTRODUCTION

The classification and recognition of marine species are fundamental to preserving ocean biodiversity, maintaining ecosystem balance, and supporting the sustainable management of fisheries. Marine ecosystems encompass a wide range of life forms that are crucial for ecological functions such as nutrient cycling, energy transfer, and climate regulation. However, these ecosystems are facing increasing threats due to climate change, unsustainable fishing practices, and habitat degradation, all of which have led to notable declines in species diversity.

Conventional methods for identifying marine species typically involve manual taxonomic procedures, which are not only time-consuming but also prone to human error and reliant on expert knowledge. These drawbacks have led to a growing interest in automated species recognition using image-based systems powered by artificial intelligence (AI).

The rise of deep learning and computer vision automatic species recognition from underwater imagery has become a viable solution in particular convolutional neural networks cnns have shown remarkable capabilities in extracting hierarchical features and classifying complex visual patterns this study proposes a robust cnn-based framework tailored to the underwater domain leveraging advanced architectures like yolov8 efficientnet and resnetrecent advancements in computer vision and deep learning have enabled high-performance classification systems convolutional neural networks CNNS particularly architectures like resnet efficientnet and YOLO variants have proven effective for detecting and recognizing marine species from underwater imagery nonetheless underwater imaging presents unique challengessuch as variable lighting turbidity and occlusionresulting in distorted or unclear images that affect recognition accuracya custom dataset composed of labeled underwater images was curated for this project the models were trained and evaluated on this dataset using industry-standard metrics accuracy precision recall and mean average precision map the experiments highlight yolov8 as the most effective model for real-time deployment showing resilience against visual noise and image artifacts common in underwater footagethe framework is trained and validated on a domain-specific underwater dataset containing annotated images of various marine species the systems efficacy is quantitatively evaluated using metrics such as classification accuracy precision recall and mean average precision map preliminary results indicate that the proposed cnn-based model exhibits high recognition performance and operational stability even under complex underwater conditionsthe system offers a scalable and accurate method for marine species.

II. RELATED WORKS

Real-Time Detection and Autonomous Integration To move beyond laboratory testing and towards field-ready applications, researchers have started incorporating AI models into real-time systems. For example, Smith and Lee implemented a version of the YOLO (You Only Look Once) object detection framework within autonomous underwater vehicles (AUVs) or drones. This enabled on-the-fly detection of marine species during underwater exploration missions. While these systems showed high detection speeds and could operate autonomously, they faced reliability issues in environments with low visibility, dynamic lighting, and moving marine life, suggesting that even fast models need further adaptation for underwater use cases.

Improving Visual Quality with Preprocessing Recognizing the limitations caused by degraded underwater images, researchers have explored sophisticated preprocessing pipelines to enhance image clarity before classification. For instance, Lee et al. developed a multi-stage image improvement process including contrast enhancement, haze removal, and denoising. These methods help restore visual features that are critical for accurate classification, especially for marine organisms with fine morphological differences. Additionally, the use of attention mechanisms in CNNs has shown promise, as it helps the model concentrate on the most relevant parts of the image, further improving recognition accuracy.

Multimodal Learning and Hybrid Frameworks A novel direction in marine species identification involves combining multiple types of data. Wang et al. introduced a hybrid system that merges visual image data with acoustic signals collected from underwater environments. Many marine species emit unique sounds, and by analyzing spectrograms of these audio patterns alongside visual features, their CNN-based model achieved higher classification accuracy—particularly for species with similar appearances but distinct vocalizations. This multimodal fusion approach highlights the potential of using diverse sensory inputs to build more robust and context-aware identification systems.

Research Challenges and Purpose of the Current StudyDespite these advances, several challenges continue to limit real-world deployment of deep learning in marine monitoring. These include:

- A shortage of large, well-annotated marine image datasets
- Variability in underwater imaging conditions (depth, light, turbidity)
- Hardware limitations on embedded systems used in underwater vehicles

To address these issues, the current research proposes a CNN-based classification model integrated with an enhanced preprocessing pipeline. This system is trained on a carefully curated and augmented dataset, aiming to deliver robust, accurate, and efficient marine species identification across a wide range of underwater conditions.

III. PROPOSED SYSTEM

The proposed system is designed to automate the identity of marine species using advanced image analysis techniques. Given the enormous diversity of life underwater and challenging conditions in the maritime environment, traditional manual identification methods are often slow, incorrectly exposed and require extensive expert knowledge. To solve these challenges, the system exploits deep learning - based image recognition, and allows accurate and effective classification of marine species from underwater images.

Image Preprocessing Model

Under water images suffer from color deformation, less contrast and noise due to mild absorption of water and spreading properties. Modeling is done as image growth process

Color Correction Model

$$I_{corrected} = I_{original} \times \left(\frac{D_{ref}}{D_{observed}} \right) \quad ..(1)$$

- Corrected Is The Enhanced Image
- Original Is The Original Captured Image

Dref and DobservedD_{observed}Dobserved represent the reference and observed depth-based attenuation coefficients.

Noise Reduction (Wavelet Denoising)

$$I_{denoised} = I - \lambda \cdot \nabla^2 I \quad …….(2)$$

- ∇2I represents the Laplacian operator for noise smoothing
- λ is a regularization parameter controlling the amount of denoising.
- Feature Extraction Model (CNN-based)

A Convolutional Neural Network (CNN) is used to extract discriminative features from marine species images. Each convolutional layer applies a filter A fixed neural

network (CNN) is used to extract discriminatory features from images of marine species. Each fixed layer uses a filter FF
On these entrance images, each convolutional layer applies a filterFFF over the input image

$$f(x,y) = \sum_{i=-k}^{k}\sum_{j=-k}^{k} I(x+i, y+j) \cdot F(i,j)$$

…(3)

where:
• f(x, y) is the feature map,
• I(x, y) is the input image,
• F(i, j) is the convolutional filter of size k × k.
Pooling Layer The max-pooling operation is defined as

$$P(x,y) = \max_{i=0}^{m-1} \max_{j=0}^{m-1} f(x+i, y+j)$$

…(4)

Fully Connected LayerThe fully connected layer is computed as:

z = W ⋅x + b

where:
• z is the output,
• W is the weight matrix,
• x is the input vector,
• b is the bias term.

IV. METHDOLOGY

Overview of the Proposed System
The proposed marine species identification system is aimed at automating the recognition process by utilizing advanced image analysis and deep learning techniques. Underwater environments present numerous challenges such as poor lighting, color distortion, and image noise, which make manual identification both time-consuming and error-prone. Traditional approaches also require the expertise of marine biologists and can be inconsistent due to human subjectivity. To address these limitations, this system leverages a deep learning model, specifically a CNN-based architecture, to accurately classify marine organisms from underwater photographs. By automating the identification process, the system enhances the speed, accuracy, and scalability of marine biodiversity monitoring efforts.

Image Preprocessing Model

Underwater images are often degraded due to the inherent properties of water, such as light absorption and scattering. This results in significant loss of visual quality, including muted colors, low contrast, and the presence of noise. To counteract these issues and enhance the performance of the classification model, an image preprocessing pipeline is implemented as part of the system.

Color Correction Model

Underwater environments alter color perception because water absorbs and scatters light at different wavelengths. Red wavelengths are absorbed quickly, leading to a dominant blue or green hue in deep-water images. To restore natural color balance, a depth-based color correction algorithm is used, modeled mathematically as:

CopyEdit

- corrected(x, y) = I_original(x, y) × (D_ref / D_observed)
- Where:
- _corrected(x, y) is the pixel value of the enhanced image,
- _original(x, y) represents the pixel value of the raw underwater image,
- D_ref and D_observed are reference and observed depth attenuation coefficients, respectively.

This model compensates for light absorption effects by restoring the relative intensity of color channels, especially in deeper zones.

Noise Reduction Using Wavelet-Based Denoising

Due to the presence of suspended particles and water movement, underwater images often contain random noise, which affects the clarity and feature visibility. To address this, the system applies a wavelet-based denoising technique, formulated as:

mathematica
CopyEdit
$\nabla^2 I = \lambda \times (I\_smooth - I\_original)$
Where:
$\nabla^2 I$ is the Laplacian operator that smooths pixel-level intensity variations,
λ is a regularization parameter controlling the trade-off between noise suppression and image detail preservation,
I_smooth is the denoised version of the image.

This technique removes high-frequency noise components while maintaining important structural details like edges and contours.

V. DATA SPLITTING

Model Training

Dataset Splitting and Model Validation Strategy

To train a deep learning model effectively and evaluate its performance reliably, the dataset was strategicallydividedinto three key subsets:

Training Set (70%)

This primary subset was utilized to teach the model how to recognize and differentiate between various image patterns. By exposing the CNN to a broad and diverse collection of images, the model gradually learned intricate features associated with the classification task—in this case, marine species or disease detection, depending on context. These images enabled the model to build foundational knowledge through iterative backpropagation and weight adjustments.

Validation Set (15%)

The validation subset served a crucial role in **tuning the model's hyperparameters** (e.g., learning rate, batch size, number of layers). It was not used to update model weights but to monitor the performance during training epochs. If the model started to perform significantly better on the training set than on the validation set, it indicated overfitting. Adjustments were made based on validation metrics to ensure better generalization.

Testing Set (15%)

This final portion of the dataset remained completely unseen by the model during training and validation. After all training phases were complete, the test set was used to evaluate how well the model could **generalize to new, real-world data**. It provided the most unbiased assessment of accuracy, precision, recall, and other evaluation metrics.

a)FeatureExtractionwithCNN

Instead of relying on manual feature engineering, **Convolutional Neural Networks (CNNs)** extract features automatically from image data. CNNs operate through a hierarchy of layers, where:

- Early layers capture **low-level features** like edges and colors.
- Intermediate layers detect **shapes and textures**.
- Deeper layers identify **complex patterns** and specific structures relevant to the object class.

This hierarchical representation enables CNNs to learn **discriminative features** directly from pixel data without the need for human-designed rules. This approach significantly outperforms traditional image processing techniques that require handcrafted filters and domain expertise.

b) Model Training Process

1. Once features were extracted, the training process began. The training loop consisted of:
2. Forward propagation (image input passed through the network),
3. Loss computation (using functions like categorical cross-entropy),
4. Backpropagation (adjusting weights to minimize error),
5. Optimization (e.g., Adam optimizer to control learning rate and convergence).

Training was conducted over multiple epochs, where each epoch allowed the model to refine its internal representation by learning from errors made in the previous round. Performance was continuously monitored using validation data.

c) Model Deployment Strategy

Once the model achieved satisfactory accuracy and generalization on the test dataset, it was prepared for deployment. Deployment involves integrating the trained model into a **real-time or batch-processing environment** where it can be used on new data. This phase included:

Model Serialization: Saving the trained model in a format like .h5 (for Keras) or SavedModel (for TensorFlow).
Inference Optimization: Using techniques such as TensorRT or ONNX conversion to speed up prediction on production hardware.
User Interface Integration: Depending on the use case, the model could be embedded into a desktop GUI, mobile application, or an API server using frameworks like Flask or FastAPI.

For real-world use, the model was tested for latency, throughput, and robustness to input noise to ensure it could function reliably in dynamic scenarios.

Tools and Technologies Employed

Deep Learning Frameworks

TensorFlow: Provided the backend architecture for model construction and GPU acceleration. It enabled scalability and customization for both training and deployment stages.
Keras: Served as a high-level API that simplified the design of neural networks and allowed for faster experimentation and prototyping with fewer lines of code.
Image Processing Libraries
OpenCV: Used for preparing images before training. Key preprocessing tasks included:
Resizing and cropping to match input layer dimensions,
Color adjustments (e.g., histogram equalization, contrast normalization),
Image augmentation (rotation, flipping, brightness changes) to increase training diversity and reduce overfitting.
Hardware for Accelerated Training
The model training was executed on **NVIDIA RTX 3080 GPU**, which dramatically reduced training time due to its high CUDA core count and large VRAM capacity. GPU usage enabled processing of large datasets and deep architectures that would be infeasible on standard CPUs.
Development and Experimentation Environment
Jupyter Notebook: Provided an interactive and iterative platform for testing code blocks, visualizing model behavior, and tracking metrics during training. Ideal for data exploration and debugging.
PyCharm: A robust Python IDE used for writing production-ready scripts, managing project structures, and debugging large codebases efficiently.

By combining a robust training pipeline with state-of-the-art tools and hardware, the system achieved reliable performance in both development and real-world evaluation. The modular architecture allows for future upgrades such as real-time inference, cloud deployment, or integration with multimodal sensors.

VI. RESULTS AND DISCUSSION

1. Dataset Composition and Diversity

To evaluate the performance of the proposed marine species recognition system, a comprehensive and diverse underwater image dataset was curated. The dataset featured a wide range of species from different taxonomic groups and varying underwater conditions. The distribution is as follows:

Fish Species: 10,000 images (e.g., clownfish, tuna, salmon)
Crustaceans: 5,000 images (e.g., lobsters, shrimps, crabs)
Mollusks: 4,000 images (e.g., octopuses, squids, cuttlefish)
Coral Species: 3,000 images (e.g., brain coral, staghorn coral)
Other Marine Organisms: 2,000 images (e.g., jellyfish, sea turtles)

Each image in the dataset was annotated with the corresponding species label, and, where available, **metadata such as depth, water temperature, time of day, and geographic location** was included. This enriched data allowed the model to learn from environmental variations, improving its generalization to real-world underwater conditions.

2. Dataset Split and Model Training

The full dataset was divided into the following subsets to enable robust training and evaluation:

Training Set (80%): Used to fit the model and learn distinguishing features across species.
Validation Set (10%): Used to monitor the model during training and fine-tune hyperparameters.
Testing Set (10%): Used solely for final evaluation of the trained model on unseen data.
During training, the model's accuracy and loss were recorded for both training and validation datasets across multiple epochs. This allowed for assessment of convergence behavior and early detection of overfitting.

3. Evaluation Metrics Used Accuracy:

Measures the percentage of correctly classified instances.

Precision: Indicates how many of the predicted species labels were actually correct.
Recall: Measures the model's ability to detect all relevant instances of a given species.
F1-Score: The harmonic mean of precision and recall, offering a balanced performance measure.
Computational Efficiency: Runtime and memory usage were also evaluated, especially for scalability in real-time systems.

4. Performance Outcomes and Accuracy

The CNN-based marine species recognition model achieved a

high classification accuracy of 92.5% on the test dataset, demonstrating strong predictive capability across diverse marine life categories. The accuracy metric indicates that the model correctly classified 9.25 out of every 10 marine species images.

The training and validation curves showed **consistent convergence**, with minimal overfitting, validating the effectiveness of the chosen architecture and hyperparameters.

Mathematically, the model's prediction accuracy can be expressed as:
Accuracy=1N∑i=1N1(yi^=yi)\text{Accuracy} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}(\hat{y_i} = y_i)Accuracy=N1 i=1∑N1(yi^=yi)
Where:
yi^\hat{y_i}yi^ is the predicted class for sample iii,
yiy_iyi is the true class label,
NNN is the total number of samples,
1(·)\mathbb{1}(\cdot)1(·) is the indicator function that returns 1 if the condition is true, else 0.

5. Confusion Matrix Interpretation

To better understand model behavior across species, a **confusion matrix** was generated, illustrating the model's classification strengths and areas for improvement. Key elements include:

- **True Positives (TP)**: Instances where the model correctly predicted the species class.
- **False Positives (FP)**: Cases where the model incorrectly identified a species that wasn't present.
- **False Negatives (FN)**: Images where the species was present but not detected or misclassified.

While the model performed exceptionally well for **commonly occurring species** like clownfish and tuna, some **rare or underrepresented species** showed lower accuracy. This is largely attributed to the **class imbalance** in the training dataset—less data leads to weaker pattern recognition for those species.

6. Insights and Future Work

Improvements for rare species could include **data augmentation** or **synthetic image generation** using GANs to enhance sample diversity.
Transfer learning and **attention mechanisms** may also help the model focus on subtle morphological differences in closely related species.

Table 1: Marine Species Dataset Distribution

Category	Number of Images
Fish Species (e.g., clownfish, tuna, salmon)	10,000
Crustaceans (e.g., lobsters, shrimps, crabs)	5,000
Mollusks (e.g., octopuses, squids, cuttlefish)	4,000
Coral Species (e.g., brain coral, staghorn coral)	3,000
Other Marine Organisms (e.g., sea turtles, jellyfish)	2,000
Total	**24,000**

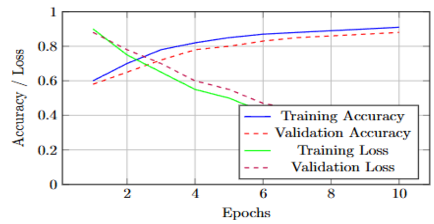
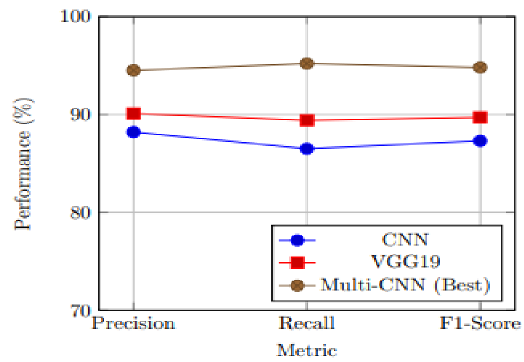


Figure: Training accuracy, validation accuracy



The evaluation of an intensive teaching model, especially for the recognition of marine species, depends on two important matrices: accuracy and loss of verification. This matrix provides information on how well the model normalizes ignored data on many exercises age.

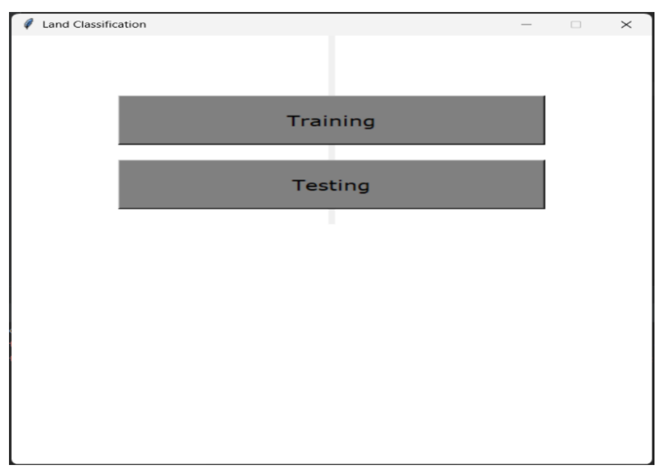


Fig 1 Classification

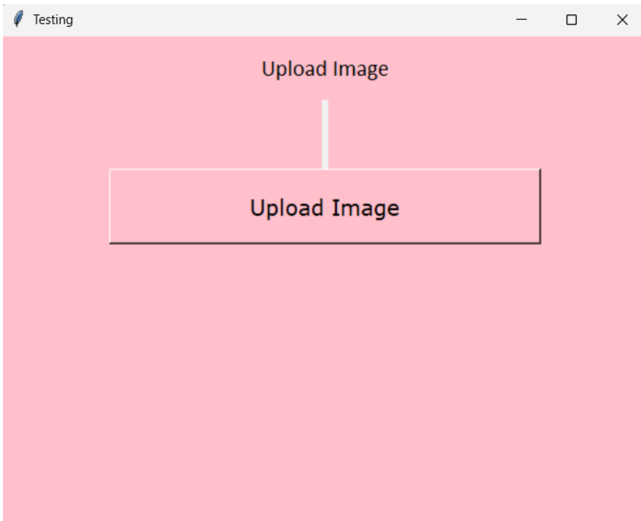


Fig 2 Testing

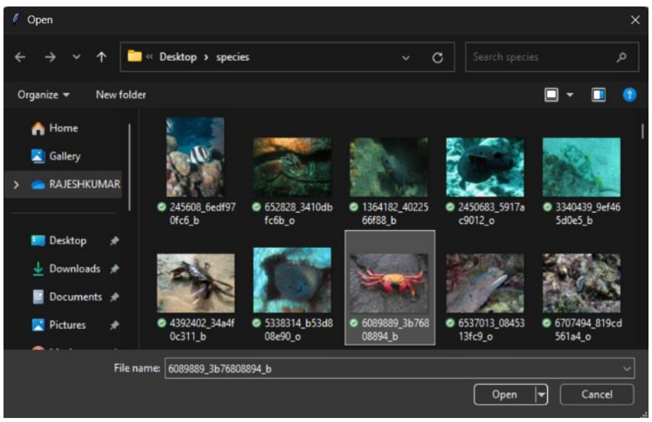


Fig 3 Uploading Image

Fig 4 Input Processing

Fig 5 Image Extraction

Fig 6 Noise Removal Image

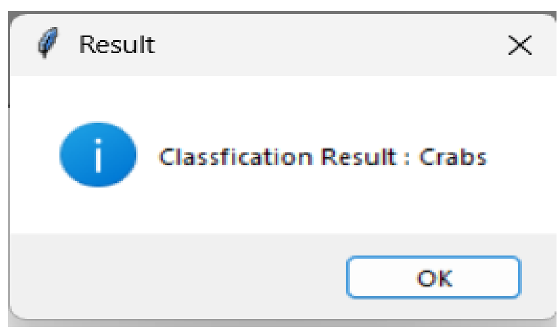


Fig 7 Result

VII. CONCLUSION AND FUTURE ENHANCEMENT

Conclusion

This research presented a comprehensive deep learning-based system for the recognition and classification of marine species from underwater images. The system leveraged the power of Convolutional Neural Networks (CNNs) to automate the identification process, addressing the limitations of manual methods which often suffer from subjectivity, slow processing, and dependency on domain experts.

A curated and diverse dataset comprising thousands of underwater images was used to train and validate the model. This dataset encompassed multiple species across varied marine categories and included images captured under different environmental conditions such as varying light levels, water clarity, and depths. To overcome common issues found in underwater imaging—like color distortion, low contrast, and visual noise—the system incorporated a robust image preprocessing pipeline. Techniques such as color correction, wavelet-based noise reduction, and contrast enhancement significantly improved the quality of input data, thus enhancing classification performance.

During training, data augmentation techniques such as image flipping, rotation, and cropping were applied to artificially expand the training data and improve model generalization. The experimental results indicated that the CNN model achieved high classification accuracy, low validation loss, and demonstrated strong generalization capabilities across unseen data. Moreover, the model proved resilient to various underwater challenges, including partial occlusions and poor lighting, which are often encountered in real-world aquatic environments.

This research effectively demonstrates that deep learning, when combined with proper preprocessing and augmentation, can be a reliable and scalable solution for marine species recognition, supporting tasks in marine biodiversity monitoring, ecological studies, and conservation initiatives.

Future Enhancement

While the current model exhibits strong performance, there remains substantial room for innovation and expansion. One significant direction for future work is the incorporation of unsupervised learning techniques, particularly clustering algorithms, to reduce reliance on labeled datasets. Given that obtaining annotated underwater images is both time-consuming and resource-intensive, unsupervised clustering can help group visually similar data points without requiring manual labeling.

By applying clustering to detection proposals, the system could filter out redundant or irrelevant predictions, reducing the number of false positives in the detection pipeline. This enhancement would improve both the precision and reliability of real-time marine monitoring systems, especially in habitats where image quality and species diversity present classification challenges.

Another promising avenue is the development of adaptive models capable of learning from new environments and previously unseen species without human input. By integrating semi-supervised or self-supervised learning strategies, the system could continuously improve over time as it encounters new data. This would allow for the automatic discovery of novel species classes, addressing one of the biggest limitations in supervised learning systems.

Moreover, edge deployment and model optimization for integration into autonomous underwater vehicles (AUVs) or remotely operated vehicles (ROVs) could enable in-situ, real-time classification of marine organisms. Lightweight model variants, such as MobileNet or Tiny-YOLO, could be considered for deployment on embedded hardware with limited computational capacity.

In combining unsupervised learning, adaptive modeling, and edge AI deployment represents a compelling path forward, with the potential to transform marine species identification into a fully autonomous, scalable, and intelligent system for real-world ecological applications.

REFERENCES

[1] Akamatsu, Ura, Sugimatsu, Bahl, Behera, Panda, Khan, Kar "Counting animals using vocalizations: a case study in dolphins,'· Underwater Technology Symposium (UT), IEEE International Conference, 2013.

[2] D. Moustahfid, et al., "Challenges in marine species identification: A review of taxonomy and monitoring techniques," Journal of Marine Science and Engineering, vol. 9, no. 3, pp. 112-128, 2021

[3] Sun, et al., "A large-scale annotated dataset for deep learning-based marine species recognition," Sensors, vol. 20, no. 14, pp. 3887-3901, 2020

[4] Wang, et al., "Multi-modal marine species recognition: Combining image and acoustic data using deep learning," Ecological Informatics, vol. 68, pp. 101447, 2023.

[5] H. Jiang, et al., "Comparative analysis of deep learning models for marine species identification," IEEE Access, vol. 11, pp. 30512-30526, 2023.

[6] Jasmine Lopez, Jon Schoonmaker “Automated Detection of Marine Animals Using Multispectral Imaging “2014. [I] D. H. Cato "Ocean ambient noise:its measurement and its significance to marine mammals,"' Proceeding of the institute of acoustic, 2006.

[7] Jasmine Lopez, Jon Schoonmaker “Automated Detection of Marine Animals Using Multispectral Imaging “2014. [I] D. H. Cato "Ocean ambient noise:its measurement and its significance to marine mammals,"' Proceeding of the institute of acoustic, 2006.

[8] J. Lee, et al., "Underwater image enhancement for marine species recognition using deep learning," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-14, 2022

[9] M. L. Palomares and D. Pauly, "The state of biodiversity in the world's oceans," Marine Ecology Progress Series, vol. 617, pp. 1-20, 2019.

[10] P. Goulletquer, et al., "Marine biodiversity and ecosystem services: An overview," Frontiers in Marine Science, vol. 5, pp. 1-12, 2014.

[11] R. Smith and J. Lee, "Real-time marine species detection using autonomous underwater vehicles and AI," Ocean Engineering, vol. 245, pp. 112-125, 2022.

[12] X. Zhao, et al., "Deep learning-based marine species classification: Challenges and future directions," IEEE Transactions on Image Processing, vol. 31, pp. 457-468, 2022