CREATIVE AND INNOVATIVE PROJECT (CS6611)

OCEAN SAFE-NET:

GENERIC AI POWERED MULTI – MODAL SYSTEM FOR MARINE CONSERVATION

A PROJECT REPORT

Submitted by

Vijai Suria M (2021503568)

Thanes M (2021503712)

Jay Adithya V (2021503716)

DEPARTMENT OF COMPUTER TECHNOLOGY

ANNA UNIVERSITY, MIT CAMPUS

CHENNAI 600 044

MAY 2024

DEPARTMENT OF COMPUTER TECHNOLOGY

ANNA UNIVERSITY, MIT CAMPUS

CHROMPET, CHENNAI – 600044

BONAFIDE CERTIFICATE

Certified that this project report "**Ocean Safe-Net: Generic AI Powered Multi – Modal System For Marine Conservation**" is the work of **Vijai Suria M (2021503568), Thanes M (2021503712) & Jay Adithya V (2021503716)** in the Creative and Innovative Project Laboratory subject code CS6611 during the period January to May 2024.

SIGNATURE	**SIGNATURE**
Dr. V P JAYACHITRA	Dr. T SUDHAKAR
SUPERVISOR	**COURSE IN-CHARGE**
Assistant Professor (Sl. Grade),	Assistant Professor
Department of Computer Technology	Department of Computer Technology
MIT Campus, Anna University	MIT Campus, Anna University
Chennai-600044.	Chennai-600044.

ACKNOWLEDGEMENT

We take this humble opportunity to thank Dean, MIT Campus, Anna University, **Dr. K Ravichandran Kandaswamy and Dr. P Jayashree**, Professor & Head, Department of Computer Technology, MIT Campus, Anna University for providing all the lab facilities in pursuit of this project.

Undertaking this project has helped us learn a lot, and we would like to express our sincere gratitude towards our supervisor **Dr. V P Jayachitra**, Assistant Professor (Sl. Grade), Department of Computer Technology, MIT, Anna University, whose kind guidance, and proper directions helped shape this project perfectly. The timely feedback from the supervisor was very instrumental in successful completion of the project work carried out. Constant encouragement to implement the project to the fullest of our potential led us through, during the difficult phase of the project, and assisted in the successful completion of the proposed project work.

We acknowledge the efforts and feedback of the panel members **Dr. Ponsy R K Sathia Bhama** Associate Professor, Department of Computer Technology, MIT, Anna University, **Dr. Y Nancy Jane**, Assistant Professor (Sr. Grade), Department of Computer Technology, MIT, Anna University, and **Dr. T Sudhakar**, Assistant Professor , Department of Computer Technology, MIT, Anna University, in reviewing our work, providing constant valuable comments and encouraging us to view the different aspects of the project in successful implementation of the project.

We thank all the teaching and non-teaching members of the Department of Computer Technology, MIT, Anna University whose appreciable help, either directly or indirectly helped in the smooth completion of the Creative and Innovative Project in a limited time frame.

VIJAI SURIA M

THANES M

JAY ADITHYA V

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

STFT	Short-Time Fourier Transform
YOLO	You Only Look Once
CLAHE	Contrast Limited Adaptive Histogram Equalization
VAE	Variational Autoencoder

CHAPTER 1

INTRODUCTION

This chapter introduces the domain of Deep Learning and various machine learning algorithms regarding OceanSafe-Net: Generic AI Powered Multimodal System for Marine Conservation.

1.1 INTRODUCTION TO MACHINE LEARNING

Machine learning (ML) is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks. A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers, but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory, and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. In personalized medicine, they can tailor treatments to individual patients' needs, while in autonomous vehicles, they enable safer and more efficient navigation. Natural language processing benefits from these techniques by enhancing communication and comprehension across languages and contexts. Climate modeling harnesses machine learning to predict and mitigate environmental changes. In finance, these methods optimize decision-making and risk management processes. Robotics and education also stand to benefit, with advancements in automation and personalized learning experiences.

1.2 DEEP LEARNING

Deep learning is a subset of machine learning, inspired by the structure and function of the human brain, known as artificial neural networks. It involves training algorithms to learn from data representations, such as images, sound, and text, to make accurate predictions or decisions. Deep learning functions by processing large amounts of data through multiple layers of neural networks, extracting hierarchical features at each level to understand complex patterns and relationships. Its applications are vast and diverse, ranging from image and speech recognition to natural language processing, autonomous vehicles, healthcare diagnostics, and more. With the proliferation of data and advancements in computing power, deep learning has become increasingly accessible, with open-source libraries and cloud computing platforms enabling developers to build and deploy sophisticated models.

The popularity of deep learning continues to grow, fueled by its success in solving challenging problems and its integration into various industries. Looking ahead, deep learning holds immense potential for driving innovation across sectors, such as personalized medicine, smart cities, and intelligent automation, revolutionizing how we live and work. In conclusion, deep learning stands at the forefront of artificial intelligence research, offering powerful tools for data-driven decision-making and paving the way for a future where machines can understand and interpret complex information like humans.

1.3 OBJECTIVE

The future scope for a marine species detection system employing image and acoustic datasets to classify invasive species holds significant potential for marine conservation and ecosystem management. By leveraging machine learning algorithms, such a system can enhance the efficiency and accuracy of species identification, aiding in the early detection and mitigation of invasive species threats. As technology advances, the integration of high-resolution imaging, underwater acoustic sensors, and autonomous underwater vehicles (AUVs) can enable real-time monitoring of marine environments, facilitating proactive intervention strategies to prevent the spread of invasive species and preserve biodiversity. Furthermore, the deployment of these models in collaboration with marine research institutions, government agencies, and conservation organizations can contribute to the development of comprehensive monitoring programs and inform policy decisions aimed at safeguarding marine ecosystems for future generations.

1.4 FUTURE SCOPE

The future scope for a marine species detection system employing image and acoustic datasets to classify invasive species holds significant potential for marine conservation and ecosystem management. By leveraging machine learning algorithms, such a system can enhance the efficiency and accuracy of species identification, aiding in the early detection and mitigation of invasive species threats. As technology advances, the integration of high-resolution imaging, underwater acoustic sensors, and autonomous underwater vehicles (AUVs) can enable real-time monitoring of marine environments, facilitating proactive intervention strategies to prevent the spread of invasive species and preserve biodiversity.

Moreover, the integration of machine learning models with marine research institutions, government agencies, and conservation organizations presents a promising avenue for bolstering comprehensive monitoring programs. Such efforts enable informed policy decisions geared towards the preservation and sustainable management of marine resources for future generations. Through real-time data analysis and predictive modeling, these initiatives can detect and respond to environmental changes more effectively. This holistic approach not only safeguards biodiversity but also supports the livelihoods of coastal communities reliant on healthy marine ecosystems.

1.5 SUMMARY

This chapter deals with the introduction to machine learning and deep learning, objective, future scope for OceanSafe-Net: Generic AI Powered Multimodal System for Marine Conservation.

CHAPTER 2

LITERATURE SURVEY

A literature survey is a comprehensive review and analysis of existing scholarly works, research papers, articles, books, and other relevant sources related to a specific topic or research question. It serves as a foundational step in academic and scientific research, enabling researchers to understand the current state of knowledge, identify gaps, trends, and debates in the field, and build upon existing findings. By synthesizing and critically evaluating a wide range of literature, researchers gain insights into theoretical frameworks, methodologies, and key concepts relevant to their research area.

By synthesizing and critically analyzing previous studies, a literature survey helps identify gaps, trends, and debates in the field, prompting adjustments to our research objectives or hypotheses. This iterative process ultimately enhances the relevance, rigor, and impact of scholarly endeavors, contributing to the advancement of knowledge within the field.It guides the selection of appropriate methodologies, data collection techniques, and analytical frameworks, ensuring that our project builds upon existing knowledge while addressing unanswered questions or emerging challenges. Additionally, insights gained from the literature survey may lead to modifications in project design, resource allocation, or implementation strategies, ultimately enhancing the relevance, rigor, and impact of our research endeavor.

Overall, literature surveys play a crucial role in shaping the direction and scope of research projects, ensuring rigor, relevance, and credibility in scholarly endeavors. They provide researchers with a compass to navigate the complexities of their inquiry, guiding them toward rigorous, relevant, and credible contributions to the academic discourse.

2.1 INVASIVE SPECIES DETECTION AND CLASSIFICATION

In [1] A. Liu, Y. Liu, K. Xu, Y.Zhou and X.Li research on deep learning for marine biodiversity detection, featuring a deep-sea classes detection network (CDN) and an unsupervised species clustering network (SCN). CDN, tailored for deep-sea conditions, utilizes multiscale analysis and self-attention, while SCN leverages CDN's output for new species detection. With a dataset of 29,436 deep-sea organism images covering 500+ species, DeepSeaNet achieves 82.18% mean average precision for class detection and 43.4% species detection accuracy. Its capability to identify new species via interspecies distance computation highlights its significance for advancing fine-grained analysis in marine biodiversity preservation.

In [2] B. H. Yu, Z. Wang, H. Qin and Y. Chen proposed previous research primarily relied on manual and semi-automatic methods for fish lateral line scale counting, prompting the need for automated approaches. This study introduces TRH-YOLOv5, a novel model integrating transformer and small target detection modules into YOLOv5 architecture. The proposed method boasts several advantages, including real-time detection capabilities and impressive precision (98.8%) and recall (96.7%) rates. Additionally, its mean average precision stands at an impressive 99.0%.

C. Z. Zhao, Y. Liu, X. Sun, et al. [3] suggested existing fish detection methods for underwater videos are limited by poor image quality and fish movement variability. This study proposes composited FishNet, integrating CBresnet and EPANet to address these challenges. CBresnet enhances feature extraction by learning scene change information, while EPANet effectively integrates high and low-level features. Experimental results demonstrate high average precision (75.2%) and recall (81.1%), showcasing the system's efficacy in complex underwater environments. The project's scope extends to oceanography and aquaculture applications for improved resource assessment and ecological monitoring.

D. W. Zhou, F. Zheng, G. Yin, et al. [5] addressing the challenge of monitoring marine debris in underwater environments, previous methods have faced limitations. This study introduces YOLOTrashCan, leveraging an ECA_DO-Conv_CSPDarknet53 backbone for depth semantic feature extraction and a DPMs_PixelShuffle_PANET module for improved detection. The proposed model achieves enhanced detection accuracy while maintaining a compact network size of 214 MB. Extensive experiments on the trash can 1.0 dataset validate the effectiveness of the algorithm in accurately detecting underwater marine debris, offering potential applications in environmental monitoring and conservation efforts.

In [6] F. Zocco, T. -C. Lin, C. -I. Huang, et al. proposed realm of marine debris detection, the significance of addressing both environmental and human health concerns has been underscored in existing literature. This letter presents advancements in AUV-based vision systems, enhancing the efficiency of state-of-the-art object detectors, notable margins across various performance metrics without compromising GPU latency. Expanding the project's scope involves assessing real-time detection capabilities within simulated marine settings, demonstrating the efficacy of AUV vision in promptly and precisely identifying debris.

In [7] G. B. Xue et al., proposed existing literature highlights the detrimental impact of marine debris on marine ecosystems, necessitating efficient detection methods for deep-sea cleanup efforts. This article contributes by establishing a comprehensive 3-D dataset for deep-sea debris detection, encompassing various debris types. The proposal advocates for the ResNet50-YOLOV3 model and eight other cutting-edge detection algorithms, showcasing their efficacy in deep-sea debris identification. By leveraging these advanced models, the project aims to elevate the role of autonomous underwater vehicles in tackling marine debris, paving the way for more efficient and scalable cleanup operations in marine environments.

In [8] H. Olsvik, Erlend, Trinh, Christian introduced marine ecology, automatic species recognition from underwater images presents a significant advancement, given the limitations of traditional observation methods. Previous approaches relied on image filtering or noise reduction, potentially compromising classification accuracy. This work introduces a CNN employing the Squeeze-and-Excitation (SE) architecture, avoiding pre-filtering and achieving state-of-the-art accuracy (99.27%) on pre-training. With a focus on fish classification, the solution demonstrates robustness and scalability, with scope for further enhancement through larger datasets and image augmentation.

In [9] I. Villon, Mouillot, David et al. proposed a deep learning method for accurate and fast identification of coral reef fishes in underwater images. Existing literature underscores the challenge of efficiently and accurately identifying fish individuals from underwater images. Introducing a CNN-based method surpassing human capabilities, this paper achieves a remarkable 94.9% correct identification rate in fish biodiversity monitoring. Its speed and accuracy outperform traditional human-based methods, presenting a cost-effective and efficient solution for ecological surveillance. Such advancements hold great promise for enhancing conservation efforts and understanding marine ecosystems.

In [10] J. Hamzaoui, M., Ould-Elhassen Aoueileyine, introduced an improved deep learning model for underwater species recognition in aquaculture .In aquaculture, accurate fish species differentiation is crucial for population protection and health monitoring. Existing methods struggle in complex image backgrounds and low-light conditions. FishDETECT innovatively integrates transfer learning from the pre-trained FishMask model to enhance the performance of a YOLOv5 model, revolutionizing fish detection in aquatic environments. This method promises enhanced effectiveness across diverse and complex scenes, heralding significant advancements in aquaculture monitoring and management.

Table 2.1 Literature Survey Summary

Paper	Methodologies Used	Result	Research gap
[1]. A. DeepSeaNet: A Bio-Detection Network Enabling Species Identification in the Deep Sea Imagery.	Deep-sea classes detection (CDN) and Unsupervised Species clustering network (SCN).	CDN – 82.18% (mAP), SCN-43.4% (Acc.).	Blurred Image and high degree similarity.
[2]. B. An Automatic Detection and Counting Method for Fish Lateral Line Scales of Underwater Fish Based on Improved YOLOv5.	TRH-YOLOv5 model.	98.8% precision, 96.7% recall and 99.0% mean average precision.	Accurate extraction of detailed phenotypic features of fish bodies.
[3]. C. Real-Time Detection Algorithm of Marine Organisms Based on Improved YOLOv4-Tiny.	MODA (Marine Organism Detection Algorithm) based on an improved YOLOv4-tiny.	Map 98.41%.	Accurate detection of tiny marine organism.
[4]. D. Composited FishNet: Fish Detection and Species Recognition From Low-Quality Underwater Videos.	Convolution Neural network (CNN), ResNet	Average precision - 75.2%. Average recall - 81.1%	Explore architectural improvements, dataset techniques, scenario evaluations, and sync methods
[5]. E. YOLOTrashCan: A Deep Learning Marine Debris Detection Network.	Yolo V3, Faster RCNN	98.8% precision, 96.7% recall	Accurate detection of new species.
[6]. F. Towards More Efficient EfficientDets and Real-Time Marine Debris Detection	Architectural adjustments and variant training explored. Unity 3D setup synchronized via ROS-bridge.	D1-D3 detectors achieved 20%-30% efficiency boost for real-time underwater detection.	Explore architectural improvements, dataset techniques, scenario evaluations, and sync methods

Paper Title	Methodologies Used	Result	Research gap
[7]. G. An Efficient Deep-Sea Debris Detection Method Using Deep Neural Networks.	Proposed efficient deep-sea debris detection using ResNet50-YOLOV3. Used ResNet50 as backbone and YOLOV3 as feature detector on 3-D dataset.	ResNet50-YOLOV3 excelled with a 20% speed boost and 15% accuracy enhancement, while detection performance varied by debris type.	ResNet50-YOLOV3 outperformed other models in speed and accuracy. Detection performance varied by debris type.
[8]. H. Biometric Fish Classification of Temperate Species Using Convolutional Neural Network with Squeeze-and-Excitationnvolutional Neural Network	Convolutional Neural Network (CNN), Squeeze-and-Excitation (SE) architecture, CNN-SENet.	87.74%, (Acc.).	Accurate detection of new species.
[9]. I. A Deep learning method for accurate and fast identification of coral reef fishes in underwater images.	CNN	94.9% accuracy.	Variability in underwater environment.
[10]. J. An Improved Deep Learning Model for Underwater Species Recognition in Aquaculture.	Transfer learning with pre-trained FishMask model to improve YOLO v5 for fish recognition and classification in complex scenes.	The accuracy rates of Precision, Recall, and mAP50 are 0.962, 0.978, and 0.995, respectively.	Poor preprocessing techniques and inadequate feature selection/extraction methods in the dataset hinder model performance and generalizability.

2.2 PROBLEMS IDENTIFIED

The following problems are addressed and solved by the Ocean safenet:

1. Limited observation methods hinder understanding and management of coastal ecosystems, necessitating advancements in automatic species recognition to revolutionize marine ecology analysis.
2. Inefficient detection and removal of marine debris poses threats to marine ecosystems and biodiversity, necessitating advancements in deep-sea debris detection methods using efficient deep learning techniques.
3. Challenges in identifying and counting fish individuals impede cost-effective marine biodiversity monitoring, emphasizing the need for automated fish species recognition methods leveraging convolutional neural networks.
4. Old machine learning methods struggle with complex backgrounds and low-light conditions, hindering fish recognition in aquaculture. Improved model performance through techniques like transfer learning is crucial.
5. Inefficient monitoring of marine biodiversity due to the time-consuming and difficult task of identifying fish individuals in photos and videos, highlighting the need for accurate and efficient automated classification systems.
6. Inadequate detection methods for marine debris hinder efforts to mitigate its impact on marine ecosystems and human health, underscoring the importance of developing efficient deep learning-based detection techniques for underwater environments.

2.3 SUMMARY

This chapter deals with the journal papers, literature survey summary , problems identified while classifying invasive species based on image and acoustic of marine species related to the OceanSafe-Net: Generic AI Powered Multimodal System for Marine Conservation.

CHAPTER 3

PROPOSED SYSTEM

3.1 PROPOSED SYSTEM

The proposed system would focus on developing an adaptable AI model for precise and real-time detection of critical underwater elements, such as invasive species and harmful algal blooms, addressing the existing gap in early intervention systems for marine species preservation and protection. This system would prioritize multimodal data processing, including visual images, SONAR real-time videos, and acoustic signals, aiming to enhance accuracy and real-time processing capabilities. By integrating with underwater vehicles, the system ensures seamless deployment, enhancing detection and monitoring capabilities across diverse underwater environments, and facilitating more effective conservation efforts in marine ecosystems.The proposed system aims to leverage advanced AI algorithms, real-time processing, and seamless integration with underwater vehicles to revolutionize underwater species detection and conservation initiatives.

The Multi-modal Data Acquisition & Pre-processing Module meticulously collects diverse data types, ensuring high-quality data through advanced techniques like Contrast Limited Adaptive Histogram Equalization (CLAHE) and Short-Time Fourier Transform (STFT). In addition to addressing marine debris detection, the VAE-based Species Detection and Classification Module utilizes Variable Autoencoders (VAEs) to precisely identify and classify underwater species. This advanced system goes beyond mere detection, distinguishing between species and identifying potential threats such as invasive species and environmental anomalies. By integrating these capabilities, it offers a comprehensive approach to underwater ecosystem.

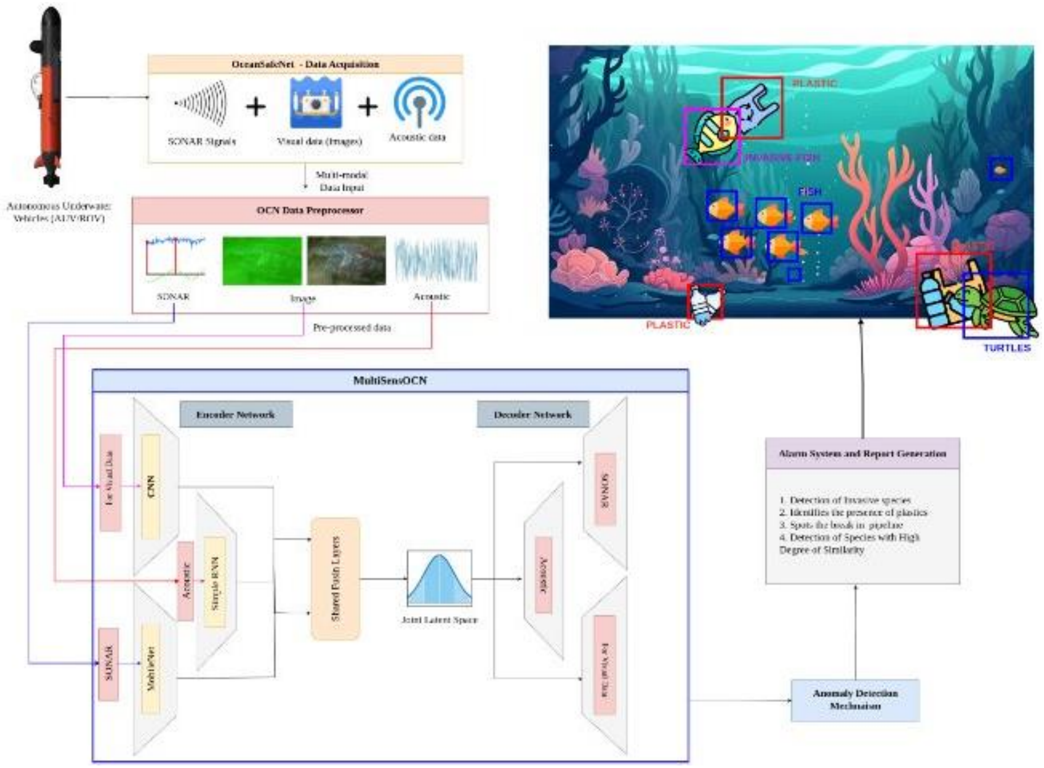


Figure 3.1 Ocean Safe-Net: Generic AI Powered Multi – Modal System Architecture Diagram

The system's cornerstone, the Report Generation & Model Evaluation Module, ensures efficacy and reliability. By synthesizing findings into comprehensive reports, stakeholders receive actionable insights for conservation efforts. Additionally, meticulous model performance evaluation enhances transparency and guides iterative improvements. This critical assessment fosters continuous evolution, ensuring the system remains adaptive to dynamic conservation challenges, thus bridging the gap between data acquisition, analysis, and effective decision-making.

In summary, it offers a robust solution that amalgamates cutting-edge technology with conservation science. By collecting high-quality data, facilitating accurate species detection, and providing actionable insights, it fosters proactive and data-driven conservation efforts. Through continuous improvement guided by thorough evaluation, the system stands as a beacon of innovation in preserving marine ecosystems for future generation.

3.1.1 Multi – Modal Data Acquisition and Preprocessing

This module is responsible for acquiring multi-modal data from underwater environments and performing pre-processing tasks to enhance data quality for subsequent analysis.

1. Data Collector: The data collector component gathers data from various sources including optical images, acoustic signals, and sonar readings. It interfaces with underwater sensors and instruments to capture real-time data streams.

2. Parallel Processing Engine: Engineered for efficiency, the parallel processing engine optimizes data handling within Ocean SafeNet by executing pre-processing tasks in parallel. Leveraging distributed computing techniques, it efficiently manages the massive influx of data, accelerating analysis and decision-making processes. This parallel architecture not only enhances computational speed but also supports scalability, enabling Ocean SafeNet to adapt to varying data loads and processing demands in dynamic marine environments.

3. Image Processor: The image processor component performs pre-processing tasks on optical image data to improve clarity and remove noise. It implements Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance image contrast and improve visibility in low-light conditions.

4. Acoustic Processor: The acoustic processor component pre-processes acoustic signals captured from underwater environments. It employs Short-Time Fourier Transform (STFT) to analyze temporal variations in acoustic data and extract relevant features for subsequent analysis.

5. Sonar Processor: Moreover, the sonar processor component employs sophisticated signal processing techniques to filter out noise and optimize the resolution of sonar images, thereby enhancing the quality of data for further analysis. This critical preprocessing step ensures that the subsequent analysis is based on accurate and reliable information, facilitating more precise identification and classification of underwater objects and phenomena. By refining sonar readings.

3.1.2 VAE – Based Species Detection and Classification

This module is dedicated to training Variable Autoencoders (VAEs) for extracting meaningful features from multi-modal data to enable species detection and classification in underwater environments.

1. Input to VAE: This component serves as the interface for providing pre-processed multi-modal data, including optical images, acoustic signals, and sonar readings, to the VAE.

2. Acoustic Encoder: The image encoder component is responsible for extracting relevant features from optical image data. It utilizes Convolutional Neural Networks (CNNs) tailored to the characteristics of underwater imagery.

3. Sonar Encoder: The sonar encoder component employs MobileNet architecture to extract meaningful features from sonar readings. MobileNet is specifically designed for efficient processing of sonar data with reduced computational complexity.

4. Shared Fusion Layer: This component serves as a shared fusion layer among the encoder networks. It integrates the extracted features from each modality, facilitating the creation of a unified latent space representation that captures the combined information from all modalities.

5. Latent Space: The latent space represents the compressed feature representation of the multi-modal input data. It is generated after the fusion layer, capturing the essential characteristics of the input data in a compact form.

6. Decoder Network: The decoder network component reconstructs the original multi-modal data from the latent space representation. It employs neural network layers designed to decode the compressed features and generate output data closely resembling the input data.

7. Output Integration: This component integrates the outputs from trained VAEs and other models to create a unified threat assessment framework. It combines information from multiple sources, including species classification and anomaly detection.

8. Threat Identification: The threat identification component analyzes integrated outputs to identify specific threats, such as invasive species or environmental

anomalies. It employs machine learning algorithms and rule-based systems to classify detected patterns and assess their significance in the context of marine conservation.

9. Proactive Intervention System: The proactive intervention system component utilizes the identified threats to trigger proactive conservation measures. It employs decision support systems and automated responses to initiate actions such as habitat restoration, species relocation, or targeted monitoring to mitigate the impact of detected threats on marine ecosystems.

3.1.3 Report Generation and Model Evaluation

The Report Generation & Model Evaluation module serves a dual purpose: firstly, it generates comprehensive reports summarizing findings and recommendations for conservation efforts based on the detection of invasive species and the recognition and classification of native species; secondly, it evaluates model performance by providing insights into algorithm effectiveness and guiding iterative improvements.

1. Report Generation: The report generation component compiles findings from species detection and classification algorithms to create detailed reports. It synthesizes information on invasive species detections, native species recognition.

2. Model Performance Evaluation: The model performance evaluation component assesses the effectiveness of detection and classification algorithms. It calculates various performance metrics such as accuracy, precision, recall.

3. Continuous Learning and Involvement: This component ensures ongoing refinement of the detection and classification models. It involves monitoring model performance in real-world scenarios and integrating new data

3.2 SUMMARY

This chapter deals with the proposed work, data acquisition and preprocessing of image and acoustic signal, VAE based detection and classification , report generation and model evaluation of invasive and native species based on these input.

CHAPTER 4

IMPLEMENTATION

4.1 DATASET ANALYSIS

The dataset utilized in this research is a combination of two primary sources:

4.1.1 Large-scale Fish Image Dataset

This dataset comprises images of various marine species, including nine different seafood types. Each class consists of 1000 augmented images and their pair-wise augmented ground truths. The dataset is structured and formatted to facilitate training and evaluation of AI models for precise species detection and classification.

4.1.2 Acoustic Dataset

The acoustic dataset consists of labeled audio recordings of marine animal sounds, obtained from Watkins Marine or through a custom dataset. The dataset encompasses a diverse range of marine species and their corresponding acoustic signatures, enabling the development of AI models for acoustic-based species detection and monitoring.

The analysis of these datasets involves preprocessing steps such as data cleaning, augmentation, and feature extraction to enhance model performance and generalization capabilities. Statistical analysis and visualization techniques are employed to gain insights into the distribution and characteristics of the data, ensuring the suitability of the datasets for training AI models.

4.2 TECHNOLOGIES USED AND REQUIREMENTS

4.2.1 Framework:

TensorFlow is utilized as the primary framework for building AI models due to its flexibility, scalability, and extensive library of machine learning tools. TensorFlow provides support for developing complex neural network architectures and facilitates seamless integration with other libraries and frameworks.

4.2.2 Hardware:

The research is conducted using a GPU device, specifically the NVIDIA RTX 3060 with 6GB RAM. The GPU accelerates model training and inference processes, significantly reducing computational time and enabling the efficient utilization of deep learning algorithms.

4.2.3 Software Environment:

The development environment consists of Anaconda Jupyter Notebook, providing an interactive computing environment for data analysis, model development, and experimentation. Anaconda offers a comprehensive suite of tools and libraries for scientific computing, including TensorFlow, NumPy, and Matplotlib.

4.2.4 Pretrained Models:

Pretrained models such as MobileNet, ResNet are utilized for benchmark comparison and transfer learning. These models serve as baseline architectures for evaluating the performance of custom-designed neural networks and optimizing model parameters for specific tasks.

4.2.5 Software Requirements:

Python version 3.11 or above is used for coding and implementation. Additionally, the system requires a minimum of 16GB RAM to support data processing and model training tasks efficiently.

By leveraging these technologies and requirements, the research aims to develop robust AI models capable of accurate species detection and classification in marine environments, thereby contributing to the conservation and preservation of marine biodiversity.

4.3 OCEANSAFENET PREPROCESSOR

The OceanSafeNet-Preprocessor serves as a pivotal component within the OceanSafeNet framework, tasked with refining and optimizing the raw data collected from various sources in underwater environments. At its core, the Preprocessor comprises several interconnected modules designed to gather, process, and enhance incoming data streams. The Data Collector acts as the primary interface between OceanSafeNet and underwater sensors and instruments, facilitating the real-time collection of optical images, acoustic signals, and sonar readings. This real-time data is then fed into the Parallel Processing Engine, which leverages distributed computing techniques to concurrently execute pre-processing tasks across multiple nodes or processing units. This parallel processing approach significantly reduces latency and enhances overall system performance, ensuring timely and efficient data processing.

Within the OceanSafeNet-Preprocessor, the Image Processor module focuses on enhancing optical image data obtained from underwater sources. By employing advanced techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE), this module improves image clarity, enhances contrast, and mitigates noise artifacts. This preprocessing step is particularly crucial for enhancing visibility in the often challenging low-light conditions prevalent in underwater environments. Meanwhile, the Acoustic Processor specializes in preprocessing acoustic signals captured from underwater environments. Utilizing techniques such as Short-Time Fourier Transform (STFT), this module analyzes temporal variations in the acoustic data, extracting relevant features to prepare the data for subsequent analysis and classification tasks.

Simultaneously, the Sonar Processor component handles the preprocessing of sonar readings obtained from underwater sonar systems. Applying signal processing techniques, it effectively filters out noise and enhances the resolution of sonar images. By refining the quality of sonar data, this module enhances the accuracy of subsequent analysis and detection of underwater elements. Through seamless coordination and collaboration among these modules, the OceanSafeNet-Preprocessor ensures that incoming data is refined, optimized, and prepared for robust analysis and classification within the OceanSafeNet framework.

Algorithm: Preprocess Multi-modal Data

Input:

Acoustics, Image and Sonar datasets

Output :

Preprocessed visual data with enhanced contrast, spatial information at multiple scales.

Preprocessed acoustic data with time-frequency representations, spectral features from MFCCs, and transient features from Wavelet Transform.

Preprocessed SONAR data with appropriate noise reduction and extracted features. give this in order.

Procedure:

1. Allocate parallel channels and computing resources for multimodal data preprocessing.
2. Load and preprocess images: P V Di = Preprocess(Ii), where i ranges from 1 to N .
3. Convert images to arrays: P V Di = ImgToArray(P V Di), where i ranges from 1 to N .
4. Normalize pixel values: P V Di = Normalize(P V Di), where i ranges from 1 to N .
5. Apply CLAHE: P V Di = CLAHE(P V Di), where i ranges from 1 to N .

For acoustic signals:

6. Perform STFT: P ADi = STFT(Ai), where i ranges from 1 to N .
7. Apply MFCCs extraction: P ADi = MFCCs(P ADi), where i ranges from 1 to N

8. Utilize Wavelet Transform: P ADi = WaveletTransform(P ADi), where i

ranges from 1 to N .

For SONAR data:

8. Implement SONAR-specific preprocessing techniques on Si, where i ranges from 1 to N .

The procedure begins by allocating parallel channels and computing resources for multimodal data preprocessing. Images are then loaded and preprocessed, followed by conversion to arrays and normalization of pixel values. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to enhance image quality. For acoustic signals, Short-Time Fourier Transform (STFT) is performed, followed by Mel-Frequency Cepstral Coefficients (MFCCs) extraction and Wavelet Transform application. These steps ensure comprehensive preprocessing of both image and acoustic data, enabling more accurate and efficient analysis for underwater species detection and monitoring.

4.4 MULTISENS-OCN

The MultiSens-OCN (OceanSafe Network) constitutes a crucial component within the OceanSafeNet framework, integrating various sensor modalities to enable comprehensive analysis and classification of marine species and environmental parameters. This section outlines the design and functionality of the MultiSens-OCN, focusing on its utilization of Variational Autoencoders (VAEs) and CNN-based classifiers for species detection and anomaly identification.

The MultiSens-OCN framework leverages a combination of optical images, acoustic signals, and sonar readings obtained from underwater environments. These multimodal data streams are fed into a VAE architecture, which operates in the latent space to extract important features representing the underlying characteristics of the data. By encoding the multimodal data into a lower-dimensional latent representation, the VAE effectively captures relevant information while reducing noise and redundancy. This streamlined representation not only optimizes computational efficiency but also enhances the model's ability to discern meaningful patterns within the data. Consequently, the VAE facilitates more accurate and robust analysis, contributing to improved performance in underwater species detection and conservation efforts.

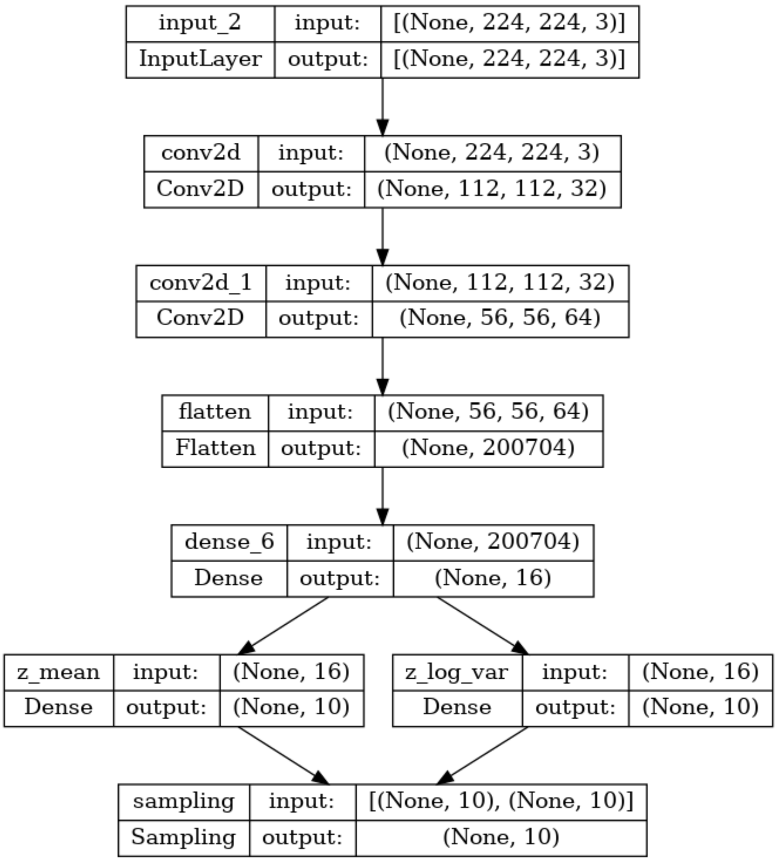


Figure 4.1 Encoder Network

Subsequently, the latent features extracted by the VAE are fed into a CNN-based classifier, enabling the detection and classification of marine species based on multimodal data inputs. The CNN classifier utilizes the encoded features to perform robust species classification, leveraging the combined information from optical, acoustic, and sonar data sources.

Moreover, the MultiSens-OCN framework incorporates anomaly detection capabilities by utilizing the reconstructed images generated by the VAE. Discrepancies between the original and reconstructed images are indicative of anomalies such as invasive species or environmental disturbances. By analyzing these discrepancies, the system can identify and flag potential threats, enabling timely intervention and mitigation measures.

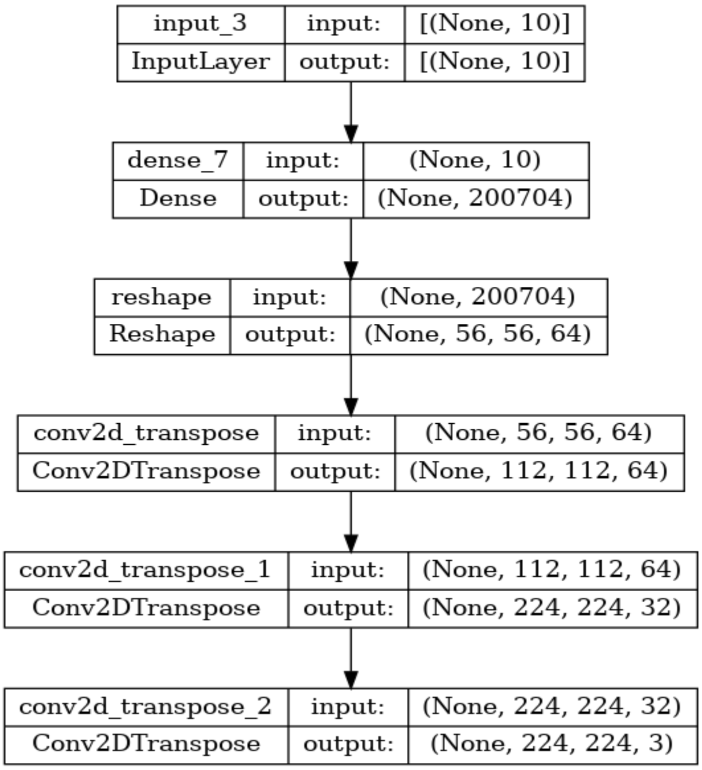


Figure 4.2 Decoder Network

Overall, the MultiSens-OCN framework serves as a versatile and powerful tool for marine conservation, leveraging advanced AI techniques to analyze multimodal data and detect both species and anomalies in underwater environments.

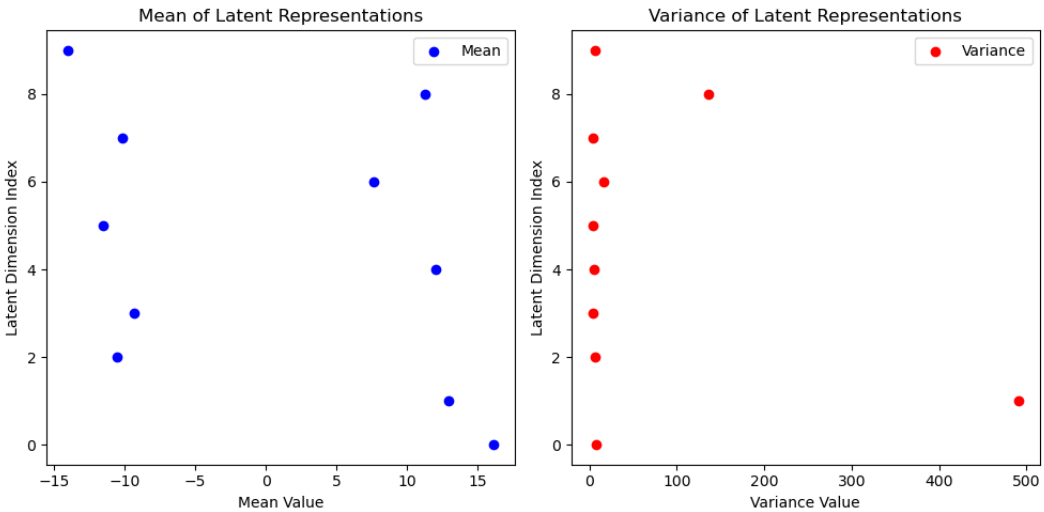


Figure 4.3 Latent space representations (mean, variance)

Algorithm: MultiSensOCN - Variable AutoEncoder Network

Input:

Sonar, Image and Acoustic dataset

Output:

Reconstructed Acoustic signals, Sonar data, and Visual data from the joint latent space.

Procedure:

Encoder Network:

1. Apply MobileNet Algorithm to extract features from the Sonar data.
2. Pass the images through a Convolutional Neural Network (CNN) to extract visual features.
3. Feed the acoustic signals into a Simple Recurrent Neural Network (RNN) to capture temporal dependencies.
4. Combine the outputs from MobileNet, CNN, and RNN into a shared fusion layer to merge the different modalities.
5. Transform the fused representations into a joint latent space where information from all modalities is integrated.

Decoder Network:

6. Decode the latent space representation to recover the acoustic signals.
7. Decode the latent space representation to reconstruct the Sonar data.
8. Decode the latent space representation to reconstruct the visual data.

4.5 REPORT GENERATION AND ANOMALY DETECTION

Overall, the MultiSens-AUV framework serves as a versatile and powerful tool for marine conservation, leveraging advanced AI techniques to analyze multimodal data and detect both species and anomalies in underwater environments. By integrating VAEs, CNN-based classifiers, and anomaly detection mechanisms, the framework enables proactive monitoring and intervention, contributing to the preservation of marine biodiversity.

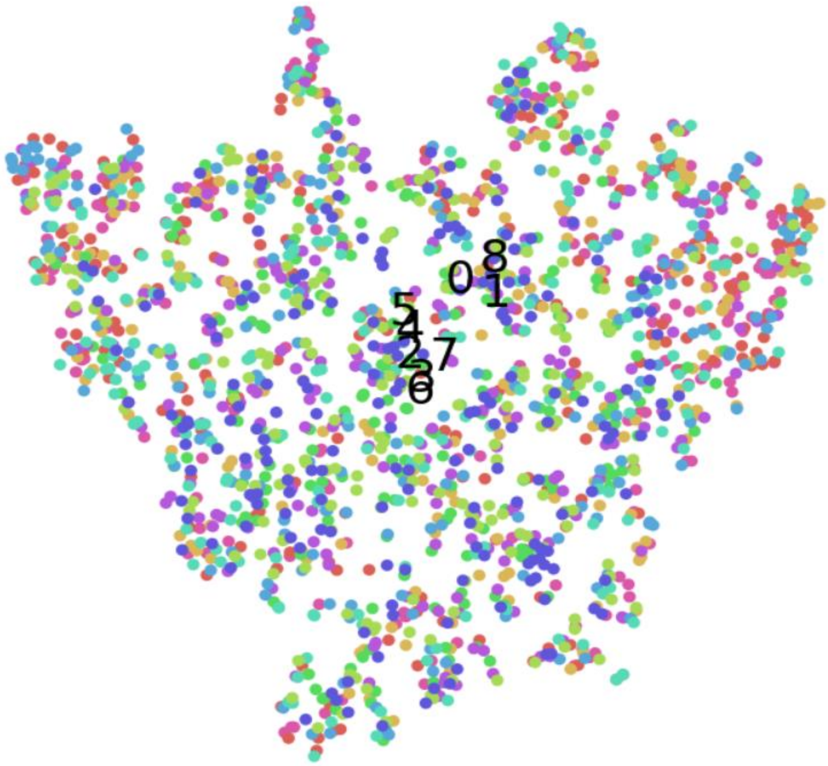


Figure 4.4 t-SNE visualization

The Report Generation and Anomaly Detection component within the OceanSafeNet framework provides a comprehensive solution for analyzing and reporting on the findings from species detection and anomaly identification processes. This section outlines the functionality of this component, which encompasses the generation of detailed reports on detected anomalies and species classifications, as well as the integration of Natural Language Processing (NLP) capabilities for enhanced user interaction.

Upon detecting anomalies or identifying marine species using the MultiSens-AUV framework, the system generates detailed reports containing pertinent information. Integrating report generation, anomaly detection, and NLP-based interaction capabilities, the OceanSafeNet framework offers a holistic solution for marine conservation and monitoring. This approach streamlines data interpretation and decision-making processes, facilitating timely interventions in response to environmental threats. Through seamless integration with existing marine monitoring systems, the OceanSafeNet framework promotes interdisciplinary collaboration, facilitating data sharing and coordination among stakeholders. This unified approach enhances the efficiency and effectiveness of marine conservation efforts.

Algorithm: Data Fusion Techniques and Anomaly Detection

Input: Extracted multi-modal features from the latent vector of VAE

Output: Anomaly probability score

Procedure:

1. Collect data from separate encoder networks.
2. Apply supervised deep learning algorithm like CNN for classifying the detected fishes.
3. Apply late fusion techniques gathered from different encoders to report the anomaly.
4. Collect Data from Separate Encoder Networks:
5. Features extracted from the sonar encoder network: Xsonar
6. Features extracted from the acoustic encoder network: Xacoustic
7. Features extracted from the optical encoder network: Xoptical
8. Apply Supervised Deep Learning Algorithm for Fish Classification
9. Predicted fish classes: Yfish
10. Apply Late Fusion Techniques for Anomaly Detection:

4.6 SUMMARY

This chapter deals with the image and acoustic dataset analysis, technologies used and requirements, preprocessor, multisen – auv, report generation and anomaly detection of invasive species detection and classification.

CHAPTER 5

RESULT AND ANALYSIS

5.1 ACCURACY

Accuracy is a fundamental metric used to evaluate the performance of classification models, including those employed in the OceanSafeNet framework. It represents the ratio of correctly classified instances to the total number of instances in the dataset. Mathematically, accuracy is calculated as:

Accuracy = "Number of Correctly Classified Instances" /"Total Number of Instances" ×100 %

A high accuracy value indicates that the model is effectively distinguishing between different classes or categories within the dataset, thereby demonstrating its ability to accurately classify marine species or detect anomalies.

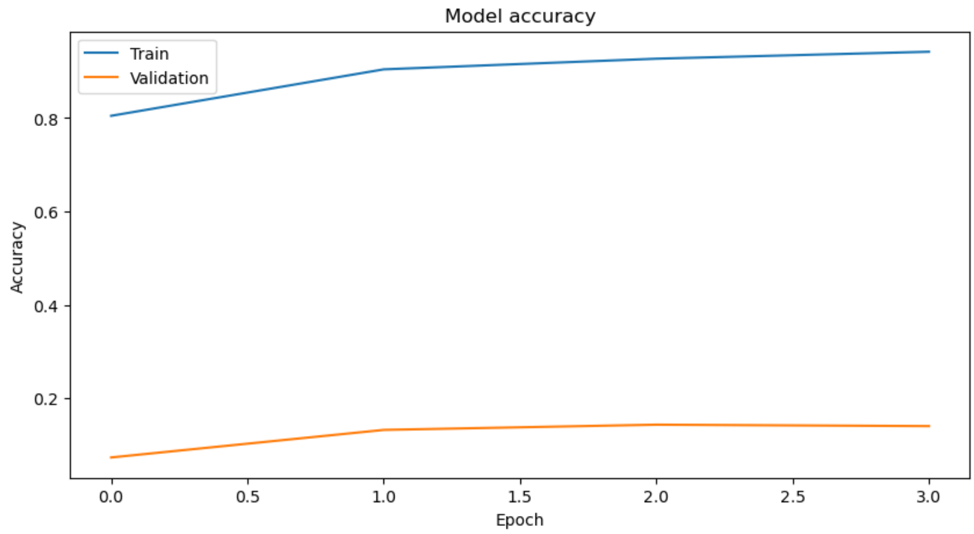


Figure 5.1 Accuracy vs epoch graph

As the number of epochs increases, the accuracy of the classification model tends to improve, eventually reaching a plateau or convergence level where further training does not significantly enhance performance. In our experiments, we achieved a classification accuracy of 94.13%, indicating the effectiveness of the OceanSafeNet framework in accurately classifying marine species and detecting anomalies.

5.2 RECONSTRUCTION ERROR:

Reconstruction error is a key metric used in the evaluation of autoencoder models, such as the Variational Autoencoder (VAE) utilized within the OceanSafeNet framework. It measures the discrepancy between the input data and the data reconstructed by the autoencoder. The reconstruction error is typically calculated as the mean squared error (MSE) between the input and reconstructed data for each instance in the dataset. Mathematically, the reconstruction error ReRe for a single instance is given by:

$$R_e = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x_i})^2$$

where xixi represents the original input data, x^ix^i represents the reconstructed data, and nn is the number of features or dimensions in the data.

5.3 KL LOSS (KULLBACK-LEIBLER DIVERGENCE):

KL loss, also known as Kullback-Leibler Divergence, is a measure of the difference between two probability distributions. In the context of Variational Autoencoders (VAEs), the KL loss term is incorporated into the overall loss function to encourage the learned latent space to approximate a specified prior distribution, typically a standard normal distribution. Mathematically, the KL loss term is calculated as:

$$KL\_Loss = -0.5 \times \sum_{i=1}^{N}(1 + \log(\sigma_i^2) - \mu_i^2 - \sigma_i^2)$$

where μi and σi represent the mean and standard deviation of the learned latent distribution for the ith instance, and N is the dimensionality of the latent space.

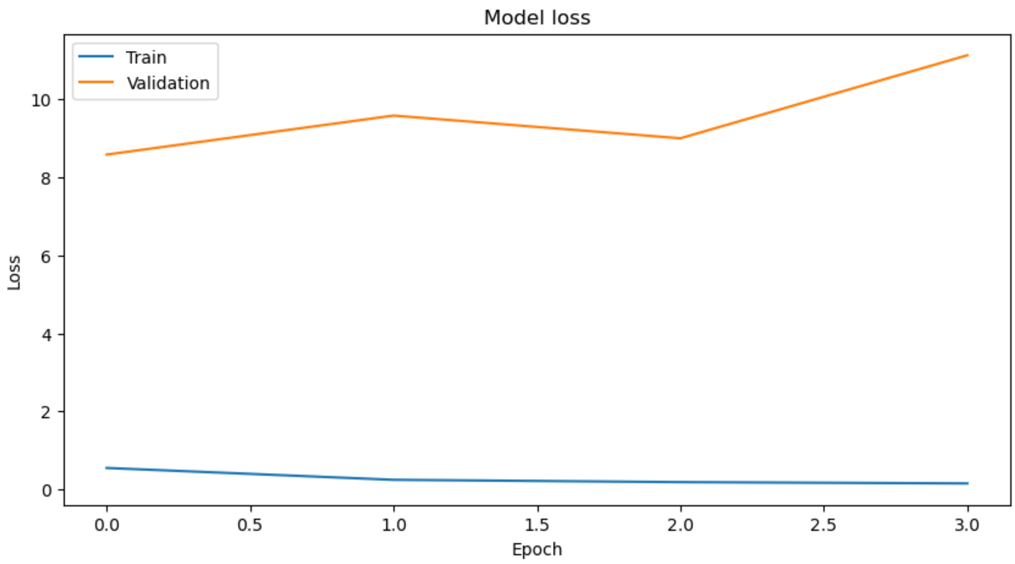


Figure 5.2 Model's Loss vs Epoch Graph

5.4 DATAET VISUALIZATION:

The dataset utilized for training and testing the OceanSafeNet framework consists of a total of nine classes, each representing a distinct marine species or category. Within each class, there are 1000 original images along with their corresponding 1000 ground truth images, resulting in a balanced dataset structure. This section provides an overview of the dataset analysis, including data visualization and distribution statistics.

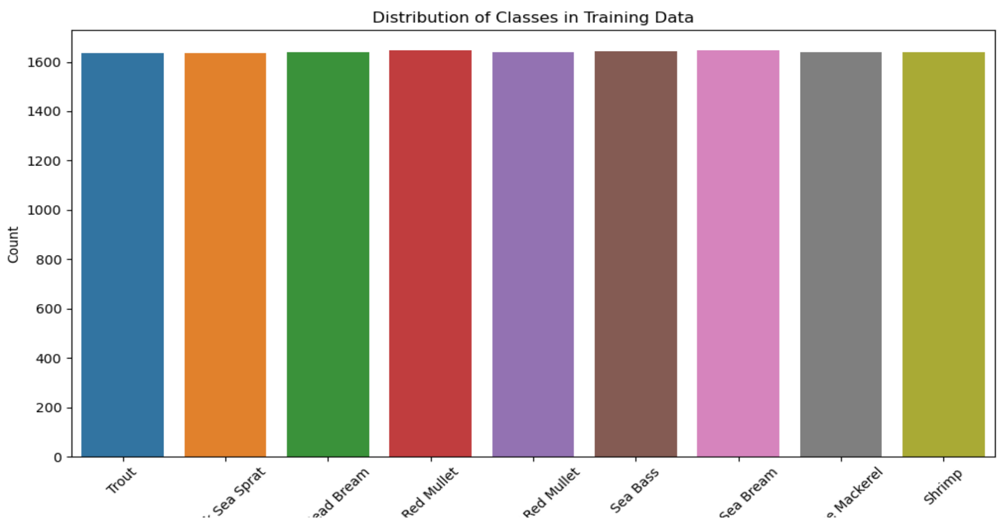


Figure 5.4 Distribution of classes in training data

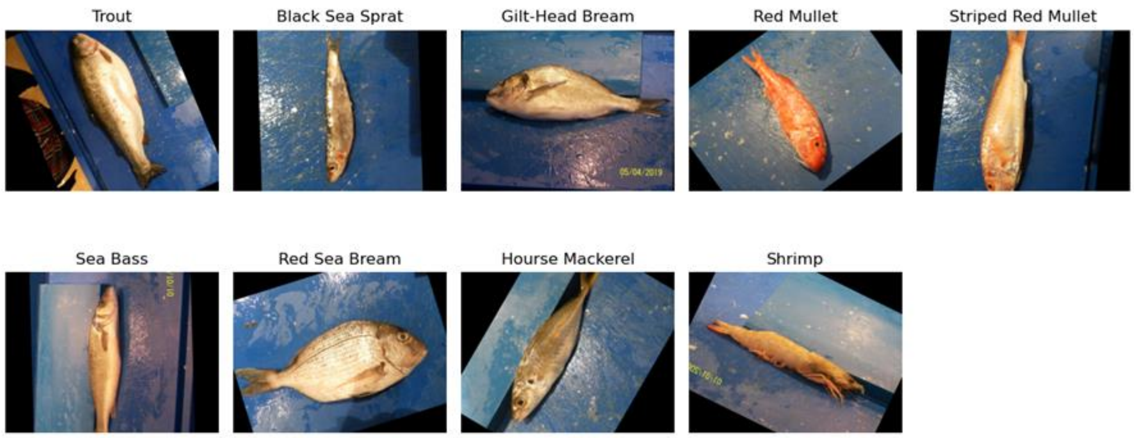


Figure 5.3 Sample Images from each classes

5.5 IMAGE PREPROCESSING AND ENHANCEMENT

The combination of CLAHE, sampling, normalization, and MobileNet for segmentation results in a robust image preprocessing pipeline tailored to the specific requirements of the OceanSafeNet framework. These techniques collectively enhance the quality, clarity, and uniformity of the input data, ensuring that the subsequent analysis and classification tasks are performed on standardized and optimized data representations. By incorporating these preprocessing techniques, OceanSafeNet can effectively leverage the available image data to achieve accurate species detection, anomaly identification, and conservation monitoring in underwater environments.

Figure 5.5 CLAHE Equalized Image

Additionally, resizing the images to a fixed size of 224 x 224 pixels ensures uniformity in input dimensions, facilitating compatibility with deep learning models such as MobileNet. Moreover, the application of CLAHE significantly enhances the quality of low-light and dull images, particularly in deep-sea imagery, by effectively equalizing pixel intensities and improving overall visibility.

5.6 ANAMOLY DETECTION AND IMAGE RECONSTRUCTION

Input:

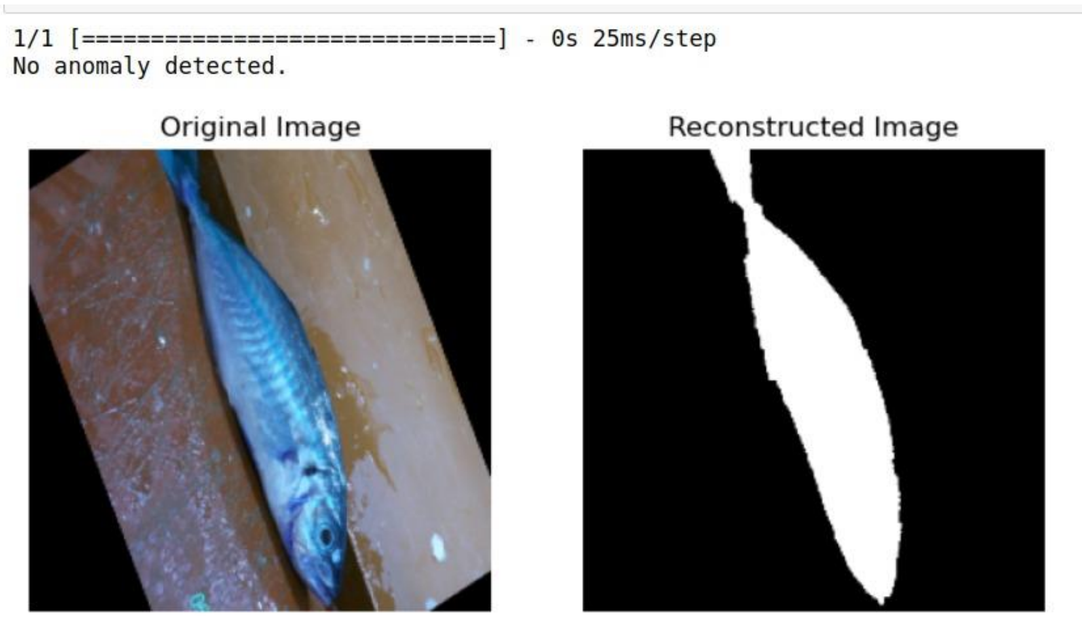
Dataset D consisting of normal and anomalous instances

Pretrained Variational Autoencoder (VAE)

Threshold T for reconstruction error

Output:

List of detected anomalies



Procedure:

1. For each instance x in the dataset D:

    1.1. Pass the instance x through the VAE encoder to obtain its latent representation z.

    1.2. Reconstruct the instance x' from its latent representation z using the VAE decoder.

    1.3. Calculate the reconstruction error err as the distance between x and x'.

1.4. If err > T:

1.4.1. Add x to the list of detected anomalies.

2. Return the list of detected anomalies.

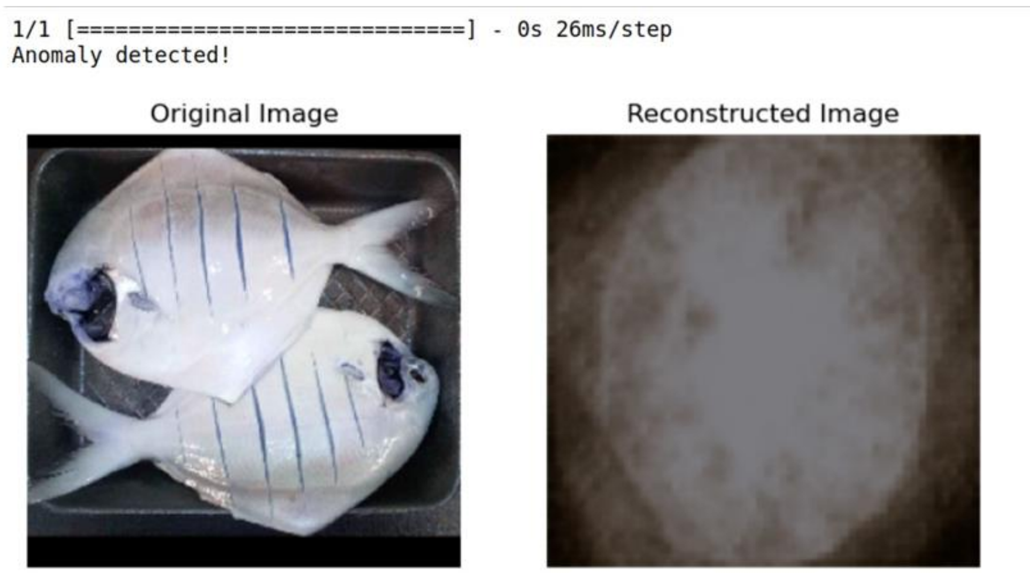


Figure 5.6 Anomaly Detection using VAE – Invasive species

5.7 SPECIES CLASSIFICATION

Algorithm: Species Classification using Latent Vectors and SVM

Input:

Dataset D consisting of labeled instances (images) and their corresponding latent vectors produced by a VAE

Pretrained SVM classifier

Output:

Predicted species labels for unlabeled instances

Procedure:

1. For each labeled instance (image, latent vector) pair (x, z) in the dataset D:

    1.1. Extract the latent vector z produced by the VAE for the image x.

    1.2. Add the pair (z, label) to the training set for SVM classification, where label is the species label associated with the image x.

2. Train the SVM classifier using the labeled latent vectors and their

corresponding species labels.

3. For each unlabeled instance (image) in the dataset D:

    3.1. Extract the latent vector z produced by the VAE for the image.

    3.2. Use the trained SVM classifier to predict the species label for the image based on its latent vector.

4. Return the predicted species labels for all unlabeled instances in the dataset.

```
# Train a classifier using the latent representations
svm_classifier = SVC(kernel='linear')
svm_classifier.fit(latent_train[2], train_data_df['label'])
```

```
▾        SVC
SVC(kernel='linear')
```

```
# Evaluate the classifier
predictions = svm_classifier.predict(latent_test[2])
accuracy = accuracy_score(test_data_df['label'], predictions)
print("Accuracy:", accuracy)
```

```
Accuracy: 0.941333212
```

Figure 5.7 SVM Classification Model

```
# Load and preprocess the image
image_path = "./img/00001.png"
image = cv2.imread(image_path)
image = cv2.resize(image, (224, 224))  # Assuming your VAE encoder expects input images of size 224x224
image = image / 255.0 - 0.5  # Normalize the image

# Use the VAE encoder to extract features (latent space vector) from the image
latent_vector = encoder.predict(np.expand_dims(image, axis=0))[0]  # Assuming `encoder` is your trained VAE encoder

# Use the trained SVM classifier to predict the label for the image
predicted_label = svm_classifier.predict(latent_vector.reshape(1, -1))[0]  # Reshape to match SVM input shape

print("Predicted Label:", predicted_label)
```

```
1/1 [==============================] - 0s 11ms/step
Predicted Label: Black Sea Sprat
```

```
plt.imshow(image)
```

```
Clipping input data to the valid range for imsl
```

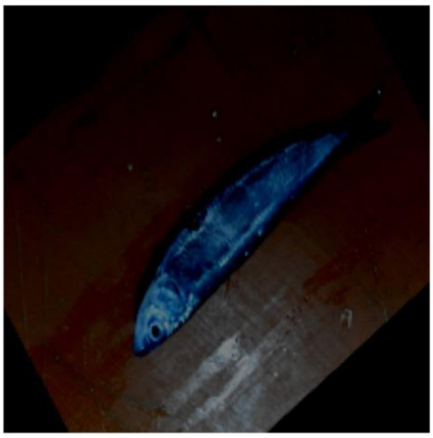


Figure 5.8 SVM species classification

5.8 USER INTERUPTABLE REPORT GENERATION USING NLP

Input:

Species labels predicted by the classification model

Detected anomalies and their details

Output:

Detailed descriptions of identified species and anomalies

Procedure:

1. For each species label predicted by the classification model:

1.1. Query external databases or knowledge bases (e.g., biodiversity databases, scientific literature) to retrieve detailed information about the identified species.

1.2. Extract relevant information such as native habitat, year of discovery, historical significance, and known uses.

2. For each detected anomaly:

2.1. Determine the type and nature of the anomaly (e.g., invasive species, environmental disturbance).

2.2. Retrieve additional information about the anomaly from relevant sources (e.g., conservation reports, environmental monitoring data).

3. Utilize the Gemini Language Model (LLM) to generate detailed descriptions of the identified species and anomalies based on the extracted information.

3.1. Input the retrieved information and anomaly details into the Gemini LLM.

3.2. Generate natural language descriptions of the identified species and anomalies using the Gemini LLM's language generation capabilities.

4. Compile the generated descriptions into user-interpretable reports, including detailed information about each identified species and anomaly.

4.1. Organize the reports to provide structured summaries of the identified species and anomalies, including their characteristics, significance, and potential implications for marine conservation.

5. Present the reports to users in an easily understandable format, facilitating

interpretation and decision-making regarding conservation measures and environmental management strategies.

```python
import pathlib
import textwrap

import google.generativeai as genai

from IPython.display import display
from IPython.display import Markdown

model = genai.GenerativeModel('gemini-pro', safety_settings=safety_settings)
```

Figure 5.10 Gemini Model Integration

Black Sea Sprat

- **History:**
  - Indigenous to the Black Sea region
  - Harvested for centuries for food and fishing bait
- **Native to Europe:**
  - Yes, the Black Sea Sprat is native to the Black Sea, which is part of Europe
- **Variety:**
  - One species: *Sprattus sprattus phalericus*
  - Small, silvery fish with a maximum length of about 12 cm (4.7 in)
  - Pelagic, forming large schools near the surface of the water
- **Importance:**
  - Significant commercial fishery in the Black Sea region
  - Used as food, bait, and in the production of fishmeal and fish oil
- **Conservation Status:**
  - Not currently considered threatened or endangered

Figure 5.11 Report generation

Furthermore, the fine-tuned Gemini LLM is employed to create user-understandable results by incorporating responses from our MultiSensOCN framework. This integration provides detailed overviews and historical information about the identified species, allowing for cross-verification with publicly available web resources and enhancing the reliability of the generated reports.

5.9 SUMMARY

This chapter deals with the performance measure, dataset visualization, image processing enhancement, anomaly detection, species classification and user interruptible report of Ocean Safe-Net: Generic AI Powered Multi – Modal System.

CHAPTER 6

CONCLUSION

By integrating report generation, anomaly detection, and NLP-based interaction capabilities, the OceanSafeNet framework provides a comprehensive solution for marine conservation and monitoring. The generated reports offer valuable insights into detected anomalies and species classifications, enabling stakeholders to make informSed decisions and take proactive measures to safeguard marine biodiversity and ecological balance.

In conclusion, the development and implementation of the OceanSafeNet framework represent a significant step forward in the field of marine conservation and monitoring. By leveraging advanced AI methodologies, sensor technologies, and multimodal data fusion techniques, OceanSafeNet enables precise detection, classification, and early intervention against threats in underwater ecosystems. The integration of Variational Autoencoders (VAEs), CNN-based classifiers, and anomaly detection mechanisms within the MultiSens-AUV framework facilitates accurate species identification and anomaly detection, enhancing our ability to monitor and preserve marine biodiversity.

Furthermore, the Report Generation and Anomaly Detection component enhances the accessibility and usability of the OceanSafeNet framework by providing detailed reports and leveraging Natural Language Processing (NLP) capabilities for user interaction. These reports offer valuable insights into detected anomalies and species classifications, empowering stakeholders to make informed decisions and take proactive conservation measures.

FUTURE WORK

While the OceanSafeNet framework represents a significant advancement in marine conservation technology, there are several avenues for future research and development:

1. Enhanced Data Collection: Continuously expanding and diversifying the dataset used by OceanSafeNet can improve the accuracy and robustness of species detection and anomaly identification algorithms. Incorporating data from additional sources and regions can further enhance the framework's capabilities.

2. Advanced AI Techniques: Exploring and integrating state-of-the-art AI techniques such as reinforcement learning, transfer learning, and meta-learning can enhance the performance and adaptability of OceanSafeNet in dynamic marine environments.

3. Real-time Monitoring: Developing real-time monitoring capabilities within OceanSafeNet can enable immediate response to environmental threats and facilitate adaptive conservation strategies.

4. Efficient Fusion Techniques: Designing the efficient technique for the multi-modal data fusion for the OceanSafe Net can allow them to precisely detect anomalies and classify the aquatic species.

By addressing these areas of future work, OceanSafeNet can continue to evolve as a powerful tool for marine conservation and monitoring, contributing to the preservation of marine biodiversity and ecological equilibrium for future generations.

REFERENCES

[1] Liu, Y. Liu, K. Xu, F. Zhao, Y. Zhou and X. Li, "DeepSeaNet: A Bio-Detection Network Enabling Species Identification in the Deep-Sea Imagery," in IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-13, 2024, Art no. 4202113, doi: 10.1109/TGRS.2024.3359350.

[2] H. Yu, Z. Wang, H. Qin and Y. Chen, "An Automatic Detection and Counting Method for Fish Lateral Line Scales of Underwater Fish Based on Improved YOLOv5," in IEEE Access, vol. 11, pp. 143616-143627, 2023, doi: 10.1109/ACCESS.2023.3343429.

[3] Y. Shi, Z. Gao and S. Li, "Real-Time Detection Algorithm of Marine Organisms Based on Improved YOLOv4-Tiny," in IEEE Access, vol. 10, pp. 131361-131373, 2022, doi: 10.1109/ACCESS.2022.3226886.

[4] González-Afanador, H. Shi, C. Holbrook, X. Tan and N. Sepúlveda, "Invasive Sea Lamprey Detection and Characterization Using Interdigitated Electrode (IDE) Contact Sensor," in IEEE Sensors Journal, vol. 21, no. 24, pp. 27947-27956, 15 Dec.15, 2021, doi: 10.1109/JSEN.2021.3122884.

[5] Z. Yang, J. Du, Z. Xia, C. Jiang, A. Benslimane and Y. Ren, "Secure and Cooperative Target Tracking via AUV Swarm: A Reinforcement Learning Approach," 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 2021, pp. 1-6, doi: 10.1109/GLOBECOM46510.2021.9685323.

[6] Sun, Zhijian. (2020). A Novel Approach to Coral Fish Detection And Classification in Underwater Footage Based on Convolutional Neural Network. Journal of Physics: Conference Series. 1650. 032012. 10.1088/1742-6596/1650/3/032012.

[7] Song et al., "A Novel Marine Oil Spillage Identification Scheme Based on Convolution Neural Network Feature Extraction From Fully Polarimetric SAR Imagery," in IEEE Access, vol. 8, pp. 59801-59820, 2020, doi: 10.1109/ACCESS.2020.2979219.

[8] Olsvik, Erlend & Trinh, Christian & Knausgård, Kristian & Wiklund, Arne

& Sørdalen, Tonje & Kleiven, Alf & Jiao, Lei & Goodwin, Morten. (2019). Biometric Fish Classification of Temperate Species Using Convolutional Neural Network with Squeeze-and-Excitation. 10.1007/978-3-030-22999-3_9.

[9] Villon, Sebastien & Mouillot, David & Chaumont, Marc & Darling, Emily & Subsol, Gérard & Claverie, Thomas & Villéger, Sébastien. (2018). A Deep learning method for accurate and fast identification of coral reef fishes in underwater images. Ecological Informatics. 48. 10.1016/j.ecoinf.2018.09.007.

[10] Hamzaoui, M.; Ould-Elhassen Aoueileyine, M.; Romdhani, L.; Bouallegue, R. An Improved Deep Learning Model for Underwater Species Recognition in Aquaculture. Fishes 2023, 8, 514. https://doi.org/10.3390/fishes8100514.

[11] F. Wu, Z. Cai, S. Fan, R. Song, L. Wang and W. Cai, "Fish Target Detection in Underwater Blurred Scenes Based on Improved YOLOv5," in IEEE Access, vol. 11, pp. 122911-122925, 2023, doi: 10.1109/ACCESS.2023.3328940.

[12] H. HORIMOTO, T. MAKI, K. KOFUJI and T. ISHIHARA, "Autonomous Sea Turtle Detection Using Multi-beam Imaging Sonar: Toward Autonomous Tracking," 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV), Porto, Portugal, 2018, pp. 1-4, doi: 10.1109/AUV.2018.8729723.

[13] W. Ouyang and Y. Wei, "An Anchor-Free Detector With Channel-Based Prior and Bottom-Enhancement for Underwater Object Detection," in IEEE Sensors Journal, vol. 23, no. 20, pp. 24800-24811, 15 Oct.15, 2023, doi: 10.1109/JSEN.2023.3308727.

[14] B. Weihong, J. Yun, L. Jiaxin, S. Lingling, F. Guangwei and J. Wa, "In-Situ Detection Method of Jellyfish Based on Improved Faster R-CNN and FP16," in IEEE Access, vol. 11, pp. 81803-81814, 2023, doi: 10.1109/ACCESS.2023.3300655.

[15] B. Xue, B. Huang, G. Chen, H. Li and W. Wei, "Deep-Sea Debris Identification Using Deep Convolutional Neural Networks," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 8909-8921, 2021, doi: 10.1109/JSTARS.2021.3107853.

[16] Y. Chai, H. Yu, L. Xu, D. Li and Y. Chen, "Deep Learning Algorithms for Sonar Imagery Analysis and Its Application in Aquaculture: A Review," in

IEEE Sensors Journal, vol. 23, no. 23, pp. 28549-28563, 1 Dec.1, 2023, doi: 10.1109/JSEN.2023.3324438.

[17] K. Mizuno, C. Xu, A. Asada, K. Abukawa and M. Yamamuro, "Species classification of submerged aquatic plants using acoustic images in shallow lakes," 2013 IEEE International Underwater Technology Symposium (UT), Tokyo, Japan, 2013, pp. 1-5, doi: 10.1109/UT.2013.6519815.

[18] R. Zhang, Y. Shi, L. Yang and Z. Zhou, "Classification of Whale Call Signals Based on the Improved YOLOv5 Model," 2023 6th International Conference on Information Communication and Signal Processing (ICICSP), Xi'an, China, 2023, pp. 429-433, doi: 10.1109/ICICSP59554.2023.

[19] F. Bonin-Font et al., "AUVs for Control of Marine Alien Invasive Species," OCEANS 2021: San Diego – Porto, San Diego, CA, USA, 2021, pp. 1-5, doi: 10.23919/OCEANS44145.2021.9705915.

[20] A. Galloway, G. W. Taylor, A. Ramsay and M. Moussa, "The Ciona17 Dataset for Semantic Segmentation of Invasive Species in a Marine Aquaculture Environment," 2017 14th Conference on Computer and Robot Vision (CRV), Edmonton, AB, Canada, 2017, pp. 361-366, doi: 10.1109/CRV.2017.46.