CS6611 – CREATIVE INNOVATIVE PROJECT

OceanSafe-Net: GENRIC AI POWERED MULTI-MODAL SYSTEM FOR MARINE CONSERVATION

SECOND REVIEW

MENTOR: **Dr. JAYACHITRA.V P** GROUP**-10**

TEAM MEMBERS:

NAME	VIJAI SURIA M	THANES M	JAY ADITHYA V
REG. NO.	2021503568	2021503712	2021503716

ABSTRACT:

In the pursuit of marine conservation, "OceanSafeNet: GENRIC AI POWERED MULTI-MODAL SYSTEM FOR MARINE CONSERVATION" emerges as a cutting-edge AI-powered multimodal system tailored to detect and classify marine species, including invasive ones. By integrating sonar, acoustic, and optical image data, OceanSafeNet enables comprehensive analysis of underwater ecosystems. Employing Variable Autoencoders (VAEs) for multimodal data and an anomaly detection framework, the system facilitates early intervention against invasive species and environmental concerns. Additionally, OceanSafeNet-env, an adjunct module, continuously monitors vital environmental parameters, ensuring proactive conservation measures. This project introduces the Eco-System Focused Lightweight Model, prioritizing energy efficiency and adaptability. This model architecture dynamically prioritizes acoustic data processing in low luminosity or dark environments, minimizing environmental impact while ensuring accurate detection and monitoring of marine ecosystems. Through the fusion of advanced AI methodologies and sensor technologies, OceanSafeNet underscores a pivotal step toward safeguarding marine biodiversity and ecological equilibrium.

PROBLEM STATEMENT:

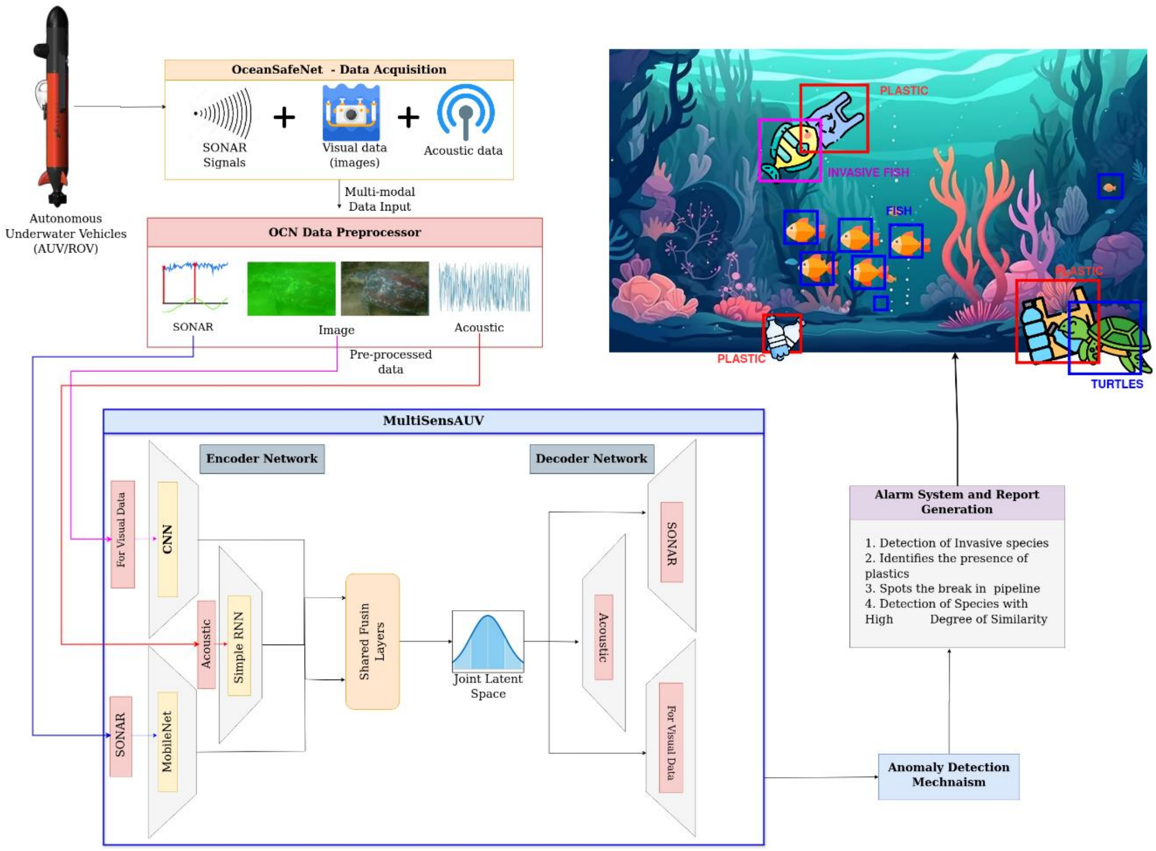
- Existing AI models lack precision in detecting critical underwater elements like invasive species and harmful algal blooms.
- There is a significant gap in accurate early intervention systems for detecting and alerting plastic pollution to preserve and protect marine environments.
- Current models struggle with multimodal data processing, including visual images, real-time videos, acoustic signals, and environmental/geospatial data, highlighting the need for real-time processing and high-accuracy models, which the current system lacks.
- There is a pressing need for a versatile model capable of integration with various underwater vehicles like Autonomous Underwater Vehicles (AUVs), Remotely Operated Vehicles (ROVs), and other underwater drones.

OBJECTIVE:

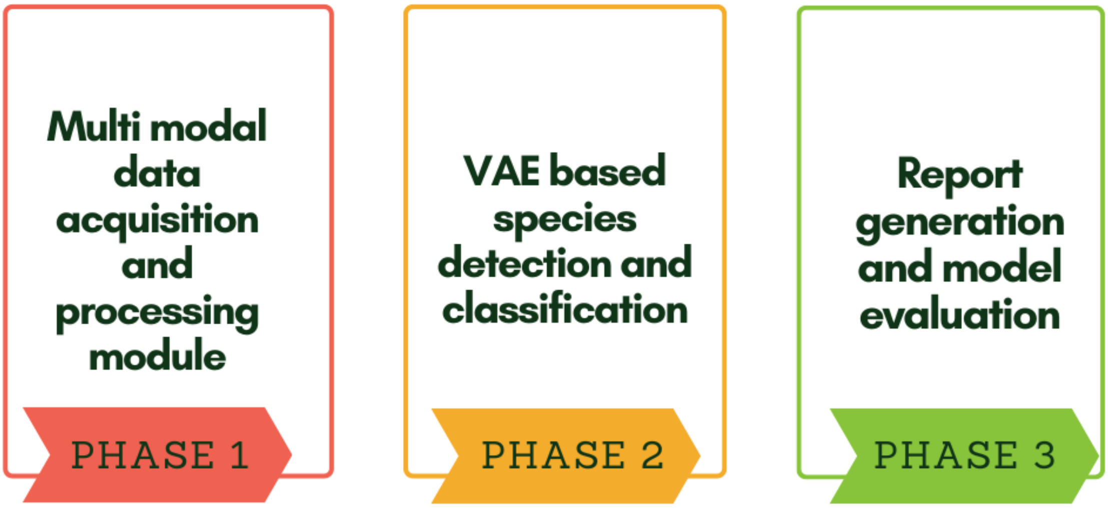
The objectives of our project are as follows:

- Develop AI models for precise detection and classification of underwater elements, focusing on invasive species and harmful algal blooms.
- Create a highly efficient early intervention systems for real-time monitoring, detection, and alerting of environmental threat.
- Achieve precise detection through multi-modal data fusion, using advanced computer vision and signal processing techniques for accurate results.
- Engineer a versatile and light weight model architecture for seamless integration with various underwater vehicles, facilitating widespread deployment and data collection.

SYSTEM BLOCK DIAGRAM:



MODULE DESIGN AND DESCRIPTION:



1. **Multi-modal Data Acquisition & Pre-processing Module:** This module oversees the collection of sonar, acoustic, and optical image data from underwater environments. It includes pre-processing techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and Short-Time Fourier Transform (STFT) to enhance data quality and relevance for subsequent analysis.
2. **VAE-based Species Detection and Classification Module:** Focused on training Variable Autoencoders (VAEs), this module aims to extract meaningful features from multi-modal data for underwater species detection and classification. Also, it integrates outputs from trained VAEs and other models to detect threats in underwater environments, including invasive species and environmental anomalies, facilitating proactive conservation measures.
3. **Report Generation & Model Evaluation:** This module generates comprehensive reports summarizing findings and recommendations for conservation efforts. It also evaluates model performance, providing insights into algorithm effectiveness and guiding iterative improvements.

Module 1: Multi-modal Data Acquisition & Pre-processing

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

Components:

- **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.
- **Parallel Processing Engine:** The parallel processing engine processes the collected data in parallel to expedite pre-processing tasks. It utilizes distributed computing techniques to handle large volumes of data efficiently.
- **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.

- **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.
- **Sonar Processor:** The sonar processor component handles pre-processing tasks for sonar readings. It applies signal processing techniques to filter noise and enhance the resolution of sonar images, improving the accuracy of subsequent analysis.

Design:

- The data collector component interfaces with underwater sensors and instruments to retrieve optical images, acoustic signals, and sonar readings in real-time.
- The parallel processing engine distributes data to the image processor, acoustic processor, and sonar processor for concurrent pre-processing.
- The image processor applies CLAHE to enhance image contrast and improve visibility in low-light conditions.
- The acoustic processor utilizes STFT to analyze temporal variations in acoustic data and extract relevant features.
- The sonar processor implements signal processing techniques to filter noise and enhance the resolution of sonar images.

Algorithm 1 Preprocess Multi-modal Data

1) **Require:**
   - Image data: $I = \{I_1, I_2, ..., I_N\}$
   - Acoustic signals: $A = \{A_1, A_2, ..., A_N\}$
   - SONAR data: $S = \{S_1, S_2, ..., S_N\}$
2) **Ensure:**
   - Preprocessed visual data: PVD $= \{PVD_1, PVD_2, ..., PVD_N\}$
   - Preprocessed acoustic data: PAD $= \{PAD_1, PAD_2, ..., PAD_N\}$
   - Preprocessed SONAR data: PSD $= \{PSD_1, PSD_2, ..., PSD_N\}$
3) **Tasks:**
   a) Allocate parallel channels and computing resources for multimodal data preprocessing.
   b) For visual data (images and videos):
      - Load and preprocess images: $PVD_i = \text{Preprocess}(I_i)$, where i ranges from 1 to N.
      - Convert images to arrays: $PVD_i = \text{ImgToArray}(PVD_i)$, where i ranges from 1 to N.
      - Normalize pixel values: $PVD_i = \text{Normalize}(PVD_i)$, where i ranges from 1 to N.
      - Apply CLAHE: $PVD_i = \text{CLAHE}(PVD_i)$, where i ranges from 1 to N.
   c) For acoustic signals:
      - Perform STFT: $PAD_i = \text{STFT}(A_i)$, where i ranges from 1 to N.
      - Apply MFCCs extraction: $PAD_i = \text{MFCCs}(PAD_i)$, where i ranges from 1 to N.
      - Utilize Wavelet Transform: $PAD_i = \text{WaveletTransform}(PAD_i)$, where i ranges from 1 to N.
   d) For SONAR data:
      - Implement SONAR-specific preprocessing techniques on S_i, where i ranges from 1 to N.
4) **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.

Module 2: VAE-based Species Detection and Classification

Overview:

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.

Components:

- **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.
- **Image 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.
- **Acoustic Encoder:** The acoustic encoder component extracts pertinent features from acoustic signals using a Simple Recurrent Neural Network (RNN). This architecture is chosen for its ability to capture temporal dependencies in acoustic data.
- **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.
- **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.
- **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.
- **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.
- **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, to provide a comprehensive view of potential threats in the underwater environment.
- **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.
- **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.

Design:

- The input to VAE component interfaces with pre-processed multi-modal data, including optical images, acoustic signals, and sonar readings.
- Separate encoder components are employed for each modality, namely the image encoder, acoustic encoder, and sonar encoder, tailored to the characteristics of their respective data types.
- A shared fusion layer integrates the extracted features from each modality, facilitating the creation of a unified latent space representation.
- The latent space captures the essential characteristics of the input data, enabling efficient encoding and subsequent decoding.

- The decoder network reconstructs the original multi-modal data from the latent space representation, generating output data closely resembling the input data.
- The output integration component combines outputs from trained VAEs, species classifiers, and anomaly detection algorithms to create a unified threat assessment framework.
- Threat identification analyzes integrated outputs to identify specific threats, leveraging machine learning algorithms and rule-based systems for pattern recognition and classification.
- The proactive intervention system utilizes identified threats to trigger proactive conservation measures, employing decision support systems and automated responses to mitigate the impact of threats on marine ecosystems.

Algorithm 2 MultiSensOCN - Variable AutoEncoder Network

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

    b) Decoder Network:
        - Decode the latent space representation to recover the acoustic signals.
        - Decode the latent space representation to reconstruct the Sonar data.
        - Decode the latent space representation to reconstruct the visual data.
4) **Output:**
    - Reconstructed Acoustic signals, Sonar data, and Visual data from the joint latent space.

Module 3: Report Generation & Model Evaluation

Overview:

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. This module plays a crucial role in informing decision-making processes in marine conservation efforts.

Components:

- **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, habitat assessments, and conservation recommendations into accessible and actionable documents.
- **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, and F1-score to

quantify the model's performance. Additionally, it tracks model performance over time, facilitating continuous learning and improvement.

- **Continuous Learning and Involvement**: This component ensures ongoing refinement of the detection and classification models. It involves monitoring model performance in real-world scenarios, gathering feedback from stakeholders, and integrating new data to enhance model accuracy and adaptability to changing environmental conditions.

Algorithm 3 Data Fusion Techniques and Anomaly Detection

Input:
- Extracted multi-modal features from the latent vector of VAE
Output:
- Anomaly probability score
Process:
- Collect data from separate encoder networks.
- Apply supervised deep learning algorithm like CNN for classifying the detected fishes.
- Apply late fusion techniques gathered from different encoders to report the anomaly.
Algorithm:
Step 1: Collect Data from Separate Encoder Networks:
- **Input:** None
- **State:**
  - Features extracted from the sonar encoder network: X_{sonar}
  - Features extracted from the acoustic encoder network: X_{acoustic}
  - Features extracted from the optical encoder network: X_{optical}

Step 2: Apply Supervised Deep Learning Algorithm for Fish Classification:
- **Input:** Multi-modal features: $X_{\text{multi-modal}}$
- **State:**
  - Predicted fish classes: Y_{fish}

Step 3: Apply Late Fusion Techniques for Anomaly Detection:
- **Input:**
  - Features from all modalities: $X_{\text{sonar}}, X_{\text{acoustic}}, X_{\text{optical}}$
- **State:**
  - Fused feature vector: X_{fused}
  - Anomaly probability scores: P_{anomaly}

=0

Design:

- The report generation component synthesizes findings from species detection and classification algorithms into comprehensive reports, providing actionable insights for conservation efforts.
- Model performance evaluation assesses the effectiveness of detection and classification algorithms by calculating performance metrics and tracking performance over time.
- Continuous learning and involvement ensure ongoing refinement of the detection and classification models, incorporating feedback from stakeholders and new data to improve accuracy and adaptability.

EVALUATION METRICS:

Evaluation Metric	Description
Accuracy	Measures the overall correctness of the system's predictions.
Precision	Measures the proportion of true positive predictions among all positive predictions.

Recall	Measures the proportion of true positive predictions among all actual positive instances.
F1-score	The harmonic mean of precision and recall, providing a balance between the two metrics.
Mean Average Precision (mAP)	The average precision across all classes, commonly used in object detection tasks.
Sensitivity (*True Positive Rate*)	Measures the system's ability to detect true positives under different environmental conditions.

TEST CASES:

Test Case	**Scenario**	**Input**	**Expected Output**
1	Acoustic signal of a known invasive species detected	Spectrogram of the acoustic signal containing the signature of the invasive species	System correctly identifies and classifies the invasive species with high precision and recall.
2	Underwater video containing plastic debris	Frames extracted from the underwater video	YOLO or Faster R-CNN accurately detects and classifies plastic debris in the frames with high Intersection over Union (IoU) scores.
3	Real-time monitoring detects low luminous levels and triggers acoustic signal gathering	Real-time luminous level data below a predefined threshold	System activates acoustic data collection and processes it for underwater species detection and identification.
4	Deployment of the fused model in a diverse underwater environment	Underwater data collected from various locations with different environmental conditions	The fused model demonstrates robustness and generalization, accurately detecting and classifying threats across different scenarios with high accuracy and reliability.
5	Single Marine Species	Input with a single marine species	OSN detects and classifies the single marine species with high confidence.
6	Multiple Native Species	Input with diverse native species	OSN accurately identifies and classifies multiple native species present.

7	Mixed Marine Environment	Mix of native and invasive species	OSN distinguishes between native and invasive species, classifying each accurately.
8	Background Noise	Presence of non-species objects (e.g., bubbles)	OSN filters out background noise and focuses on detecting and classifying species only.
9	Underwater Vegetation	Detection and classification of underwater plants	OSN detects and classifies underwater vegetation accurately, distinguishing it from species.
10	Long-Distance Detection	Detection of marine species from a distance	OSN accurately detects and classifies species even at long distances with high confidence.
11	The input image does not contain any objects of interest.	Image containing no object	OSN correctly identifies that no objects are present in the image.
12	Input image are partially obscured by other objects or environmental factor	Image with partially occluded objects	OSN accurately identifies and classifies partially occluded objects with precision.

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.