Whale Sound Detection and Noise Classification

Introduction:

The study of whale sounds is important in marine ecology and conservation efforts. Whale vocalizations give insights into communication, migration, and environmental adaptation. The problem is in differentiating whale sounds from noise and classifying them according to species or behavior.

Leveraging AI and ML techniques promises to be a promising approach to automate the detection and classification of whale sounds. These techniques can process vast amounts of acoustic data, thereby enhancing the efficiency and accuracy of bioacoustic research. This project aims to:

- 1. Distinguish whale sounds from noise.
- 2. Identify specific whale species.
- 3. Classify types of noise.

Dataset Preparation:

Dataset Types

The datasets used in this project include:

- Whale Species:* Sperm whale, humpback whale, bowhead whale, pilot whale, and finback whale.
- Noise Types:* Anthropogenic noise, e.g. ship engines; submarine noise, and marine noise, e.g. wave patterns.

Labelling:

Precise labeling of spectrograms was vital in maintaining data quality. For every recorded sound, an image spectrogram was generated and labeled, relying either on the annotation by the expert or validation using some pre-existing database.

Data Amount and Preprocessing:

Dataset Size: About 5,000 labeled spectrograms for noise detection, 4,000 for species classification.

Data Preprocessing Operations:

Noise reduction using filtering methods

Normalizing the audio inputs

Resize the images uniformly to 224x224 pixels.

Applying augmentation technique such as pitch shift and time-stretch the dataset to further diversify

Noise Detection Model:

A CNN was designed for binary classification. The discriminated acoustic event was between whale sounds and noise.

CNN Structure:

The CNN structure is comprised of the following:

- Input Layer: Accepts 224x224 spectrogram images
- Convolutional Layers: Extract features with ReLU activation.
- Pooling Layers: Reduce the spatial dimensions.
- Fully Connected Layers: Final classification with softmax activation.

Training Process:

The model was trained using a balanced dataset of 2,500 samples for whale sounds and 2,500 for noise. Robustness through data augmentation has been applied to enhance the system. Training showed 91.53% accuracy while validation is 90.17%, meaning there is robust generalization.

Classification of Whale Species:

Following the implementation of the noise detection model, species classification uses a multi-class CNN, to classify which type of whales have been sounded.

Dataset and Training:

The dataset included 4,000 spectrograms across five species. Training resulted in a model accuracy of 88.9%, with confusion matrices highlighting species-specific performance.

Noise Type Classification:

The CNN model was extended to classify noise into anthropogenic, submarine, or marine categories.

Challenges:

Subtle differences in spectrogram features made classification difficult. Advanced feature extraction and augmented datasets improved model accuracy to 85.7%.

Whale Call Type Classification:

This feature of the project is currently not implemented as the datasets classifying whale calls as mating, food, or SOS calls are unavailable. In the future, it may be implemented if such datasets are available.

Model Ensemble:

All the models were put into one pipeline

- 1.Noise Detection: Filter out unnecessary noise.
- 2. Species Classification: Identifies the type of whale species.
- 3. Noise Type Classification: Identifies the noise types.

System Architecture:

The system processes raw audio data by way of spectrogram generation, sequentially applying the integrated models for comprehensive analysis.

Results and Evaluation:

Performance Metrics

- - Noise Detection: 91.53% accuracy.
- - Species Classification: 88.9% accuracy.
- - Noise Type Classification: 85.7% accuracy.

Visualizations:

Confusion matrices, accuracy trends, and classification outputs were used to evaluate and improve model performance.

Challenges and Future Directions:

Challenges:

- - Data Limitations: Insufficient labeled data for specific whale calls.
- - Preprocessing Demands: High computational costs.

- - Subtle Feature Variations: Difficulty distinguishing similar spectrograms.

Future Enhancements:

- - Incorporate larger datasets with diverse environments.
- - Experiment with advanced architectures like transformers.
- - Implement real-time analysis for immediate conservation efforts.

Conclusion:

The "Whale Sound Detection and Noise Classification" project is very important because the development of a good system through AI/ML shows potential towards better conservation techniques, advancing scientific knowledge of the marine ecosystem as a whole, and deep insight into the under-world.