International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04157
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Bird Sound Prediction

Rajlaxmi Kasture, Sujata Patil

Rajlaxmi S. Kasture MCA & Trinity Academy of Engineering, Pune
Sujata R. Patil MCA & Trinity Academy of Engineering, Pune

ABSTRACT - This paper presents a machine learning-based system designed for recognizing bird species from their vocalizations. Leveraging Artificial Neural Networks (ANNs), specifically a Multilayer Perceptron (MLP), the system processes pre-recorded bird calls, extracts Mel Frequency Cepstral Coefficients (MFCCs), and classifies the sound using a trained ANN model. The proposed system is computationally efficient and accurate, making it a valuable tool for ecological research and conservation. The system is deployed via a user-friendly Flask web interface for real-time usage.

Keywords: Bird Sound Recognition, Deep Learning, MFCC, Librosa , TensorFlow, Flask

KeyWords: optics, photonics, light, lasers, stencils, journals

1.INTRODUCTON - Bird vocalizations offer crucial insights into their behavior, breeding patterns, and habitat preferences. Technological advances now allow automatic classification of such sounds using artificial intelligence. This paper proposes a lightweight, high-accuracy ANN-based system to classify bird species based on audio features. Develop a methodology enabling automatic recognition of bird species through sound analysis.

2.PROBLEM STATEMENT - Recent years, the monitoring of bird populations has become increasingly important for ecological research and conservation efforts. Traditional methods such as manual observation and tagging are time-consuming, laborintensive, and often lack scalability.

3.Methodology

Dataset Collection - The dataset was sourced from Kaggle and comprises audio clips in .wav format. Each clip was standardized to a sampling rate of 22050 Hz using the librosa library. Data was categorized by species, and filenames were mapped to corresponding class labels.

4.DATASET AND PREPROCESSNG - The dataset consists of 5,422 labeled audio samples distributed across 10 bird species, hosted on Kaggle. Data were evenly distributed across classes. Each audio sample was normalized and resampled to 22kHz using Librosa, ensuring consistent input for modeling. MFCCs (Mel Frequency Cepstral Coefficients) were extracted to represent time-frequency characteristics of the signals.

4. FEATURE EXTRACTION - Mel Frequency Cepstral Coefficients (MFCCs) were extracted from each audio file. MFCCs give a simple version of how the strength of different sounds in an audio signal changes over time, and they are commonly used to help recognize or classify different types of sounds. Each audio file's MFCCs were converted into a feature vector for training the model.

5. SYSTEM DESIGN -

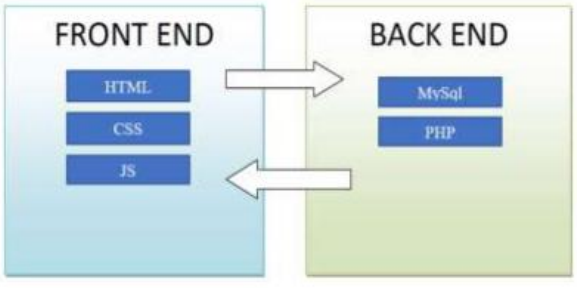


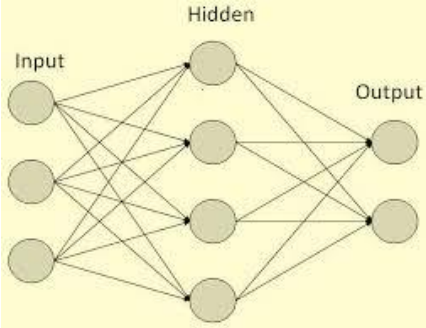
Figure 3-2: System Model

6. MODEL ARCHITECTURE - We implemented an Artificial Neural Network (ANN) using the Keras Sequential API.

• First Hidden Layer: 100 neurons with ReLU activation and a dropout rate of 0.5

• Second Hidden Layer: 200 neurons with ReLU activation and a dropout rate of 0.5.

• Third Hidden Layer: 100 neurons with ReLU activation and a dropout rate of 0.5
•Output Layer: 10 neurons with Softmax activation, corresponding to the ten bird species class.



7. IMPLEMENTATON AND DEPLOYMENT - Training was performed for 250 iterations with a minibatch size of 32. It achieved approximately 97% training accuracy and 100% test accuracy. The system was deployed using Flask, providing a web interface where users can upload audio files and receive instant species identification results, along with a confidence score.

8. DATASET - During the initial module, a framework was established to collect the dataset used for training and evaluating the model. The dataset comprises 5,422 labeled bird sound recordings and is publicly accessible for experimentation on Kaggle. Importing libraries: The important library that supports audio and music analysis is Librosa. Librosa offers essential tools for developing models focused on audio information retrieval. We use TensorFlow, a strong and popular tool, to build deep learning models.
• pip install pandas
• pip install Pillow
• pip install –upgrade pandas==1.3.4
• pip install Keras==2.8.0
• pip install tensorflow==2.8.0
• pip install Flask
• pip install librosa==0.9

9. FEATURE SCOPE
Future enhancements may include:
• Expanding the system to support more bird species.
• Incorporating Recurrent Neural Networks (RNNs) for time-series modeling.
• Changing How We Watch Nature.

10. CONCLUSIONS
Its high accuracy, low computational requirements, and web-based interface make it suitable for real-world applications in environmental monitoring and ornithology.

11. REFERENCES

1. Vivekanandam, B. (2021). Speedy Image Crowd Counting by Light Weight CNN, JIIP 3(3), 208–222.

2. Manoharan, J. S. (2021). Variants of ELM in Classification, JSCP 3(2), 83–95.

3. Chandu, B. et al. (2020). Automated Bird Species Identification, IEEE AISP.

4. Incze, A. et al. (2018). Bird Sound Recognition Using CNN, IEEE SISY, 295–300.

5. Rai, P. et al. (2016). Audio Feature Extraction for Bird Species, ICT, 1-5.