



# **Deep Learning-Based Dual-Modal Bird Species Identification Using Audio and Images**

## **IEEE BASE PAPER TITLE:**

### **Automated Bird Detection Using Snapshot Ensemble Of Deep Learning Models**

#### **IEEE BASE PAPER ABSTRACT:**

A diverse variety of bird species are found in the earth. There are about 10,000 unique species of birds discovered so far. Identification and documentation of these birds, their migrating seasons, breeding seasons, food style, etc. requires effective monitoring. Identification of these birds is literally an impossible task for human without technological intervention. Hence deep learning models have been used which helps us in identification of the individual species of birds. In this scenario, it would be important to use an automated method to identify the species of birds. Several deep learning-based models for classifying and identifying bird species are evaluated in this study. These models are trained and tested on publicly available dataset. Snapshot images captured by the cameras are used as the major source of input data. The capability of deep learning models have been used to effectively identify the bird species individually. This study aims in developing a deep learning model using ensemble learning techniques that is capable of identification of individual bird species from a massive collection of input data images.



## **OUR PROPOSED PROJECT ABSTRACT:**

Birds play a critical role in ecosystems, serving as indicators of environmental health, pollinators, and agents of seed dispersal. Identifying bird species accurately is essential for biodiversity conservation, ecological research, and monitoring environmental changes. This project introduces a dual-modal bird species identification system using deep learning, leveraging both audio and image data for robust and accurate classification.

This project explores a dual-modal approach to bird species identification utilizing deep learning techniques, integrating audio and image data for accurate classification. The system is developed using Python, with HTML, CSS, and JavaScript for the frontend, and Flask serving as the web framework. Two distinct models are implemented to process the two modalities: an Artificial Neural Network (ANN) for bird audio classification and the Xception architecture for bird image classification. For audio-based identification, an Artificial Neural Network (ANN) is employed, achieving a training accuracy of 100% and a validation accuracy of 97%. The audio dataset includes 9,107 recordings in .wav format, representing five bird species: American Robin, Bewick's Wren, Northern Cardinal, Northern Mockingbird, and Song Sparrow.

For image-based identification, the Xception architecture is utilized, achieving a training accuracy of 99% and a validation accuracy of 96%. The image dataset comprises 87,050 images across 510 bird species, ranging from Abbott's Babbler to Crested Shriketit.

This dual-modal system highlights the synergy of combining auditory and visual modalities for robust species identification. The project demonstrates the potential of deep learning to advance automated ecological monitoring and biodiversity



conservation efforts. By integrating audio and visual modalities, this system enhances identification accuracy and robustness.

## **SYSTEM REQUIREMENTS:**

### **HARDWARE REQUIREMENTS:**

- System : Pentium i3 Processor.
- Hard Disk : 500 GB.
- Monitor : 15” LED.
- Input Devices : Keyboard, Mouse.
- Ram : 8 GB.

### **SOFTWARE REQUIREMENTS:**

- Operating System : Windows 10 / 11.
- Coding Language : Python 3.10.9.
- Web Framework : Flask.
- Frontend : HTML, CSS, JavaScript.

## **REFERENCE:**

Fazeelath Jahan Shaik, Ganesan V, “AUTOMATED BIRD DETECTION USING USING SNAPSHOT ENSEMBLE OF DEEP LEARNING MODELS”, IEEE Conference, 2024.