International Journal of Combined Research & Development (IJCRD)
eISSN:2321-225X; pISSN:2321-2241 Volume: 14; Issue: 8; August- 2025

AUTOMATED BIRD SPECIES IDENTIFICATION USING AUDIO SIGNAL PROCESSING AND IMAGES USING NEURAL NETWORKS

Varshini D G
PG, Student
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
varshinidgmca2025@gmail.com

Sowmya J
Assistant Professor
Dept. of MCA
The Oxford College of Engineering,
Bommanahalli, Bengaluru- 560068
sowmyaj@theoxford.edu

ABSTRACT

The project's goal is to create a system that uses visual and aural data to identify the species of birds perfectly. To accomplish this objective, the deep learning approaches are deployed in the system whereby species identification based on images will be fixed using VGG16 whereas audio identification will be fixed using Multi-Layer Perceptron (MLP) model. The two modalities used together enable the system to be able to determine whether a bird species is seen or heard. The model makes a close guess of a bird species using a user-friendly Flask-based web application that enables an individual to upload a picture or an audio file. The VGG16 model which was originally trained on ImageNet has now been refined in order to classify bird photos. This assists in differentiating the birds of different sound types. This is displayed on the web at the Flask setup, where the users can view and name the bird types online. This publication will improve bird watching and nature studies because it will be a convenient method of identifying birds.

KEYWORDS: *Acy position system will be developed to determine what process you have taken a picture of birds from the photos and sounds. The bird type with deep instrument learning techniques can be identified by the system and VGG16 can identify the bird type based on its physical look, MTLP can identify the bird type based on its sound. The app is web based, built on Flask frame, is simple to use. Here users can add the sound image to recognise the bird online. This project marries the two approaches to assist people spotting birds more clearly and assist in nature conservation by enabling people to become bird able which will facilitate future work to keep them safe.*

INTRODUCTION

A state-of-the-art tool that uses machine learning to recognize the automated bird species recognizes different bird species by

their unique sounds. The system offers precise species identification through the use of a Multi-Layer Perceptron (MLP) classifier and Mel-frequency cepstral coefficients (MFCCs) feature extraction. Field research, bird watching, and education can all benefit greatly from this user-friendly web application, which is based on Flask and allows users to upload recordings of bird sounds and receive species predictions. This cutting-edge system could completely change how we recognize and research bird species, supporting conservation initiatives and deepening our knowledge of the natural world.

LITERATURE SURVEY

1.Images-Based on Deep Learning-Based Species Identification (2018)-(Zhao et al.)

This study investigates the application of deep learning methods, particularly Convolutional Neural Networks (CNN),to categorize bird species according to images. That research demonstrates their CNN, especially architectures likes VGG16 and ResNet can be achieve high accuracy in identifying bird species from their larges datasets of bird images. The paper highlights the challenges of dealing with varying images qualities, lighting conditions, and backgrounds. It concludes that using data augmentation techniques such as random rotations, filp, and cropping improves the models ability to the generalize and identify species under different real-world condition.

Key Findings:

- VGG16 and ResNet model are highly effective for birds species classification.
- Data augmentation help overcome challenge with image quality.
- The model can classify birds species with over 90% accuracy in controlled environment.

2.Birdsong Recognition for Automatic Bird Identification (2019) – (Stowell et al.)

Stowell at al focus on the automatic classification of bird species based on bird sound recordings. The study emphasizes how difficult it is to differentiate between species using their vocalizations due to similarity in call, environmental noise, and the overlapping calls from multiples bird. The authors propose using machine learning model, including Support Vector Machines(SVMs) and Recurrent Neural Networks(RNNs), for feature extraction and classification. These demonstrate that spectrogram analysis combined with advanced audio features can enhance the system accuracy.

Key Findings:

- Identifying bird species from audio requires robust features extraction techniques.
- RNNs and SVMs show good results in classifying birds call.
- Combining acoustic features and environmental noise reduction techniques improves accuracy.

3. A Multi-Modal How to Determine the Species of Birds (2020) – (Chen at al.)

The paper introduce a multi-modal approach combining both images and audio information about birds species identification. There are the researcher train deep learning models for both image-based classification (using CNNs) and audio classification (using MLPs). The study finds the combining these two data sources result in more than robust and accurate predications, especially when one modality (e.g., images) is unavailable or compromised. These research demonstrates the potential of hybrid model in enhancing identification performance under various environmental condition.

Key Findings:

- Multi-modal systems combining both images and audio information provide better accuracy.
- Hybrid deep learning model outperform single-modal system.
- These approach is effective for real-time birds species identification.

4.Birds Classification Using Deep Neural Networks and Audio Features (2021)– (Xie at al.)

Xie at al present as the study on bird classification based on audio features using deep neural networks (DNNs). This study investigates the effectiveness of the Mel-frequency cepstral coefficients (MFCCs) and other spectral features in identifying bird species from their call. The authors highlights that although audio features can significantly aid in classification challenges persist in noise filtering and distinguishing similar-soundings species. The study also compares DNNs to the traditional machine learning techniques, finding that deep learning models consistently outperform older method.

Key Findings:

- MFCCs are effective for audio-based bird species classification.
- DNNs offer better performance than traditional methods like SVMs and decision trees.
- The accuracy of audio classification depends heavily on noise filtering techniques.

5. A Comprehensive Dataset of Bird Sounds for Classification and Recognition (2022) - (Saha at al.)

The research focuses on the creation of a larges bird audio dataset designed specifically for the classification of bird species. These dataset includes recordings from various habitats and environmental condition. The authors emphasize the importance of diverse and high-quality audio data for improving machine learning model performance. They also explore the use of advanced signal processing techniques to enhance the quality of bird

vocalization recognition such as noise cancellation and signal enhancement method.

Key Findings:

- A large diverse audio dataset is crucial for training accurate bird species classification models.
- Advanced signal processing techniques likes noise reduction improve classification accuracy.
- The dataset provides a valuable resource for training deep learning model.

6. Combining Image and Audio for Bird Identification in Field Conditions (2021) - (Tao at al.)

Tao et al propose ais the field application where both images and bird calls are used for real-time birds identification. The System combines deep learning techniques for image classification with audio-based recognition using CNNs for image processing and RNNs for audio signal processing. This multi-input system is the designed to work in field conditions such as forests and open areas where light conditions may vary and birds may be heard but not easily visible. The paper demonstrates that such hybrid systems significantly improve real-time bird monitoring accuracy.

Key Findings:

- Combining images and audio data improves real-time birds species identification.
- The system are suitable for use in field conditions where traditional image-only model struggle.
- Multi input model offer robustness in challenging environments such as dense forests or noisy urban areas

EXISTING WORK

In this existing system bird species identification systems traditional methods often rely heavily on image based classification modelal one. use deep learning model likes Convolutional Neural Networks (CNNs) such as ResNet or VGG16 for recognizing from images. However they face limitation in scenarios where birds are not visible or the images are unclear due to environmental factors such as dense foliage or low lighting. Additionally some system relay on extensive databases of bird images and detailed metadata which may not be comprehensive or diverse enough to cover all bird species especially those from remote or under studied regions. Some older system may also attempt to identify birds species from audio recording using specialized sound recognition model. However these model often suffer from limitations in accurately distinguishing between similar birds species or environmental noise such as wind or human activity which can interfere with the clarity of bird call. These system might also be constrained by the quality and variety of the audio data used for training often limiting their

reliability and accuracy in real world application.

PROPOSED SYSTEM

The proposed system aims to overcome the limitations of existing approaches by integrating both image and audio classification method into one unified system. The system utilizes VGG16 a the pre trained Convolutional Neural Network (CNN) for accurate image based bird species identification and a Multi-Layer Perceptron (MLP) model for classifying bird species based on audio recording of bird call. By combining these two modalities the proposed system improves the precision of bird species identification especially in challenging situation where only one type of data (either image or sound) is available. The new system provides an the intuitive web based interface developed using Flask that enables users to upload both images and audio files of birds for classification. This dual input approach enhances the system versatility and reliability making it more robust across various real world scenarios such as identifying birds in dense forests (where sound recordings might be more reliable than images) or urban environments (where visual data might be more readily available).

The proposed solution ensures that the system performs well in diverse conditions and offers higher accuracy by leveraging both image and sound classification techniques which have been trained using a broad diverse dataset. The proposed system is the intelligent web based application capable of recognizing the types of birds from the vocalizations using system machine learning techniques. These system utilizes an Multi-Layer Perceptron (MLP) classifier trained on a carefully selected dataset of bird species sounds to predict the species based on extract audio features. The applications aims to simplify the process of bird identification by providing an accessible and automated platform for users including researchers birdwatches and conservationists.

The system workflow is divided into following key components:

1. Audio Input Interface:

Users can upload audio files(.wav or .mp3) containing birds vocalizations through a simple Flask web interface.

2. Preprocessing and Features Extraction:

These uploaded audio is the preprocess to remove the noise and standardize format. Features such as Mel-frequency cepstral coefficients (MFCCs) are extracted using library which effectively captures the spectral characteristics of bird calls.

3. MLP Classifier for Prediction:

There are extracted features are fed into a pre-trained MLP model that has learned to recognize pattern in birds vocalization and classify them into specific species.

4. Result Display:

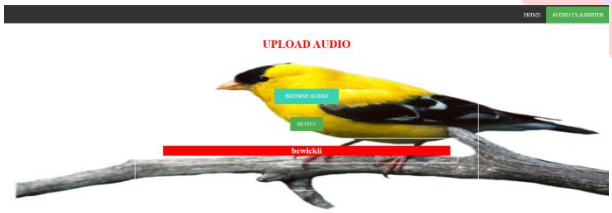
The predicated birds species name is displayed on the web interface for the user, along with

optional logging or saving of results for further analysis.

5. Extensibility and Future Integration:

The system's modular architecture enables future integration with real-time audio streams, larger datasets, and more complex deep learning models (like CNNs or LSTMs) to improve accuracy and handle a noisy environment.

EXPERIMENTAL RESULTS



CONCLUSION

The Bird Species ID from Images and Audio combines image and sound techniques for pinpointing bird classes. It rides the Convolutional Neural Networks (CNNs) with the pictures coming from the VGG16 speccie, and the sounds from a Multilayer Perceptron (MLP). Basis this plan shows a good and certain way to recognize many species of birds. Using pre-built models and learning from other data it can reach good results even having no data. Similarly, a web app designed with Flask permits the user-sized uploads of images and uploading/playing sound files to identify the species of the bird.

The app design handles both images and sounds, therefore is excellent and dependable in various situations. It's as useful to bird enthusiasts as much as it is to those who study (birding) and save (bird conservation) and teach (ornithologists) birds. It allows bird numbers a real-time tracking. The equipment also promptly identifies birds on its own spending less time for folks to look over and write down their discoveries by hand.

REFERENCES

1. Hu, B., and Zhang, W. (2019). A Survey on Multi-modal Machine Learning Techniques for Audio-Visual Applications. Journal of Artificial Intelligence Research, 61, 233-269.

2. Xeno-canto. (2021). Xeno-canto: A Repository of Bird Sounds. Retrieved from https://www.xeno-canto.org/

3. BirdNET (2021). Bird NET: A Deep Learning-based Bird Sound Identification System. Retrieved from https://www.birdnet.de/

4. Librosa. (2021). Librosa: A Python Package for Music and Audio Analysis. Retrieved from https://www.birdnet.de/

5. Flask Documentation.(2021).Flask Web Framework Documentation. Retrieved from https://flask.palletsprojects.com/