IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Automated Bird Species Identification For Conservation

Dr. Srinivas B V
*Information Science and Engineering*
*BMS Institute of Technology*
Bangalore, India

Mr. Akshay kumar
*Information Science and Engineering*
*BMS Institute of Technology*
Bangalore, India

Mr. Anchith G Raj
*Information Science and Engineering BMS Institute of Technology* Bangalore, India

Mr. Abhishek Raj
*Information Science and Engineering*
*BMS Institute of Technology*
Bangalore, India

Mr. Dhruv K
*Information Science and Engineering*
*BMS Institute of Technology*
Bangalore, India

***Abstract* — Nowadays bird population is changing drastically because lots of reasons such as human intervention, climate change, global warming, forest fires or deforestation, etc., With the help of automatic bird species detection using machine learning algorithms, it is now possible to keep a watch on the population of birds as well as their behavior. Because manual identification of different bird species takes a lot of time and effort, an automatic bird identification system that does not require physical intervention is developed in this work. To achieve this objective, Convolutional Neural Network is used as compared to traditionally used classifiers such as SVM, Random Forest, SMACPY. The foremost goal is to identify the bird species using the dataset including vocals of the different birds. The input dataset will be pre-processed, which will comprise framing, silence removal, reconstruction, and then a spectrogram will be constructed, which will be sent to a convolutional neural network as an input, followed by CNN modification, testing, and classification. The result is compared with pre-trained data and output is generated and birds are classified according to their features (size, color, species, etc.)**

***Keywords*— *Machine learning, Automatic identification, Convolutional neural network, SVM, Random Forest, SMACPY Pre-process, Spectrogram, Classification.***

## I. INTRODUCTION

The bird ecosystem has incredible diversity in terms of behavior and size and shape, but this biological diversity is under potential danger due to human intrusion into their habitats and total habitat destruction, which is accompanied by environmental disasters such as global warming, forest fires, and other natural disasters. As of 2020, 1,481 bird species, or 13.5 percent of total data sufficient species, are globally endangered with extinction due to their small and diminishing ranges.

Controlling and assessing the environment is one of the most important reasons for bird monitoring. Pollution in the air and water affects some bird species. As a result, identifying bird species can be used to spot and avert environmental problems. Birds can also assist us in detecting other life forms in the environment since they respond fast to changes in the environment. However, collecting and collating information about bird species necessitates a significant amount of human effort and is also more expensive, making it unaffordable.

In this situation, a trustworthy system will provide a wealth of information about birds while also serving as a vital tool for scientific researchers and government authorities.

Most people are also unable to identify a large number of species that are routinely observed in our daily lives. The purpose of this research is to develop an automated system that can identify birds from their audio recordings . Detecting and recognizing birds through the audio signals is quite difficult and comes with many challenges as it includes noise such as rain or traffic sounds often overlap with bird syllable making the recognition task more puzzling. Manual spectrogram inspection may result in errors because it needs many specialists, which is unreliable and points to the necessity for automated systems. Bird watching is a hobby these days so such systems can be commercially useful.

For conservation planning and management, biodiversity monitoring can give information on environmental health, migration pathways, and population status. Image, voice, and video can all be used to identify someone. Birds can be identified using an audio processing technology that captures their auditory signals. It used to be necessary to have direct touch with the bird to be identified or classified. However, with the introduction of computerized bird classification systems, it is now possible to easily identify them.
The following steps are involved in this process:

1) Outdoor recording of bird song.

2)Audio pre-processing methods are employed to improve signal quality because these recordings are often taken in noisy environments.

2) Extracting elements from the audio input and

4)Combining these qualities into machine learning algorithms to develop a judgment mechanism capable of determining bird species.

## II. PROBLEM STATEMENT

Birds exchange a variety of information through audio expression. Birds use their sounds to send various warnings and threats about impending danger, to identify individual birds or insects in a flock, and to delimit and demarcate territory. The call specialty indicates that they are rather instant and more efficient vocal expressions.

The difficulty of distinguishing diverse bird species from their recorded songs can be defined as the problem of identifying birds using an automated system with the use of bird sounds. Bird songs, according to experts, are more musical and better for species identification than bird sounds, which are employed here. To find the most relevant segment of the signal and extract features, the entire signal is pre-processed.

## III. LITERATURE SURVEY

During the previous several years, a lot of research has been done on automated bird species recognition, and several research papers have been published on the issue. Some of them have successfully classified some species, and each has its own set of advantages and disadvantages.

In literature [1], they have used Lightweight CNN (LW- CNN) architecture with VGG for crowd counting purpose. In the front end, VGG-16 has given 10 convolution layers and 3 max- pooling layers. It is feasible to reliably count the number of individuals present in a crowd using a compressed convolution depth of 6 and a dilation factor of 2.

From the study of [2] they have suggested the use of Extreme Learning Machine (ELM) algorithm to overcome the drawback of Feed-Forward Neural Network such as slow computation with different methodologies such as Evolutionary ELM, Voting based ELM, Ordinal ELM, fully complex ELM, Symmetric ELM, etc. Accuracy of ELM based classification algorithm is 94.10%.

Authors of paper [3] have used dataset having 400 samples of bird sound recordings in total, with recordings of four birds: cuckoo, sparrow, crow, and laughing dove each having an input space of 100 recordings. Bird sound recordings were collected from xeno-canto.com, a site devoted to the sharing of bird sounds from throughout the world. Each clip is between 5 and 20 seconds long and is transformed to a fixed sampling frequency of 44100 Hz or 48000 Hz in order to preserve diversity and avoid overfitting. Because the sound clip of bird voices may contain surrounding noise and human voice, they added 100 samples of human voice clips and surrounding environmental noise. The data for these examples comes from the Google Recording and LibriSpeech ASR datasets.

In the paper [4], they have used the dataset having Zebra Dove, Crested Bulbul, Magpie bird and Blue-crowned Hanging-Parrot these four birds vocals. Then the data is preprocessed which afterwards is trained using neural network.

According to researcher [5], deployed a Convolution Neural Network for classification with ResNet in their following paper. They have separated bird song recordings into two different sound classes such as signal and noise so that neural network can get train on most relevant data. To increase generalization, the noise segments are used to boost the training samples.

Literature [6] proposed a system which uses Convolutional Neural Network algorithm to predict the bird species according to their features (e.g.: genus, species, subspecies, size, color, etc.) Audio files are collected from dataset which are converted into WAV format. To clean the data, data/sound pre-processing techniques such as framing, silence removal, reconstruction, and so on are used. These WAV files are used to generate spectrogram which act as input to our Convolutional Neural Network-trained networks like Mobile Net is used After the training the network, generated output is compared with pre-trained neural network and bird species is predicted.

Article [7] have used a deep learning framework to simultaneously segment and classify the bird syllables. They used a Convolutional Neural Network in conjunction with an encoder-decoder architecture that was created for visual segmentation. The method is performed on the in-situ recordings from the forests which have shown significant improvement when compared to cutting-edge approaches.

[8] in their research article, the parameter-free clustering techniques, which are based on compression distances, were employed in this research. The following study is being done to test the performance of the NCD in this issue: The distances between groups of their audio data are computed when a small subset of bird species is chosen. Different clustering methods are applied to each distance matrix in the second stage. Finally, the quality of their clustering is used to assess their abilities. With this method, one can acquire the required results of species separability without having to analyze the data beforehand.

From literature [9], the tweeting sounds of birds usually seen in India were acquired from the environment and a few from a standard dataset for this project for e.g., xeno-canto. For each recording, MFCC dependent features are extracted. The data is then trained and tested using the SVM algorithm. The combination of MFCCs with SVMs classification is the key contribution of this approach. This technique was used to identify the following four kinds of bird species: duck, blackbird, parrot, and house crow. Mel cepstral coefficients were employed as extracted features in a multi SVM approach for classification of test samples. The overall accuracy was found to be 64%, with an 89 percent highest accuracy.

The researchers of [10] proposed a design in which they have collected five audio analysis inputs and saved them as wave files in a system. These audio files were imported into MATLAB using the wavread tool. MSE-based technique, Correlation based on frequency shifting and stability property, Wiener Filter theory, and MFCC-based strategy were all used to get the results, which were all implemented in MATLAB environment. The final graph, as well as tables, are generated to show the results. This research focuses on the technology of bird frequency analysis for species identification. It has pre-recorded bird noises that are analyzed utilizing the above approaches. The disadvantage of this system is the time constraint; it takes a long time to analyze the sound since it compares the results of four ways at once. The outcome is highly reliable in terms of determining the species.

According to article [11] , study employs the techniques of 'Segmentation and Estimation of Frequency Tracks' and 'HMM-based Modelling of Frequency Tracks' to develop an automatic system for identifying distinct bird species in audio files. They used sounds of bird species in utterances of 1, 2, and 3 seconds to experiment and achieve results for both single and multiple species recognition at the same time. Using utterances/slots of one, two, and three seconds

of the received signal by the system, this method obtained bird species recognition correctness of up to 92.0 percent, 88.8 percent, and 83.3 percent for single bird species.

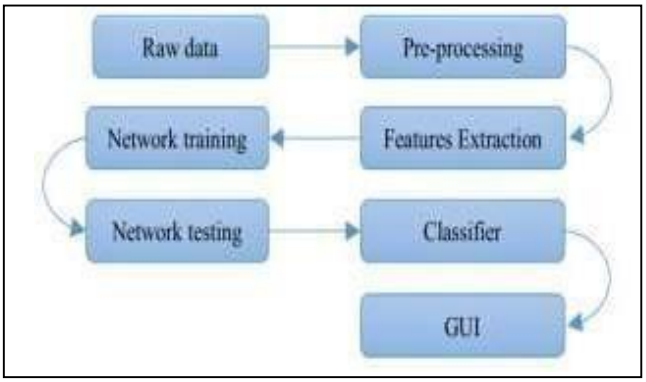
In the literature of authors [12], focused on the segmentation of vocal signals is the subject of this research in the challenge of automatic bird diversity analysis. It's based on research from a database of bird songs from about 70-75 different species located along South America's Southern Atlantic Coast.

"Three situations were considered: the very first consisted of bird song recordings, whereas the second technique relied on manual audio signal segmentation. The third possibility considers automatic audio segmentation using traditional signal processing measures computed on bird song." The automatic segmentation strategy, on the other hand, improves the classification accuracy of around 43 out of 75 bird species, including 20 bird species for which the manual segmentation approach fails (giving accuracy equal to 0 percent in each case) [12].

From research paper [13], the Hidden Markov Model (HMMM) approach is utilized to identify the specific bird species in this work. The system's primary goal is to detect bird species in field recordings automatically. A sinusoidal detection approach is used to partition the auditory scene into discrete Spectro-temporal segments. Each segment is represented by a frequency track, which is a time sequence of the observed sinusoidal frequencies. During the investigation, HMM's collection was used to depict each bird species. Unsupervised learning is used to train these HMM sets. To make the detection, the target model's probability ratio was compared to the background model's likelihood ratio.

In literature [14], the note models were trained on bird calls using data from various CDs in this study. Many note models are taught and stored in the system, which has a variety of note numbers. The acoustic features that were employed in the GMM modelling to train the note models are the same ones that were used in this study. Mel frequency cepstral coefficient (MFCC) is a characteristic computed on various bird vocalization waveforms, and the GMM system is used to model it. The GMM-UBM model paradigm is used in this system.

Bird noises are proposed as an input in an automated bird identification system in the article [15]. The audio might be anything from male birds singing songs to entice females to low/high amplitude alerts. The system's primary goal is to keep track of the problems that birds encounter near airports. The problem is solved using a conventional machine learning scenario after characteristics are extracted from recorded bird songs using a specialized audio treatment. The decision technique is built using the same information that was used to predict the species of a new birdsong. The findings of the experiments compare performance in a wide range of scenarios, including complete audio signals acquired in the field and brief audio fragments (pulses) retrieved from the signals using a split technique.




## IV. PROPOSED METHODOLOGY

The discussion's main goal is to anticipate bird species based on their voice/audio. The suggested framework contains five major phases, as depicted in Figure 1:

Fig. 1. Flowchart of methodology

1. Data is collected from respected dataset.
2. On the dataset, data pre-processing techniques such as framing and noise removal are used.
3. The data is processed using the Convolutional Neural Network approach.

### *A. Dataset*

The first step of implementation is gathering data from dataset which is obtained from XENO-canto/Kaggle. The audio recordings of the birds in MP3 format are included in this resource. This dataset contains audio recordings of the birds in MP3 format. XENO-canto/Kaggle are open websites dedicated for dataset where users upload their own recordings. In case of our survey, bird audio related dataset is required. Genus, species, subspecies, locality, type, color, size, and bird sound quality are all labeled in this dataset (from A to E, Where A is the best quality). Since many features are defined in dataset, combination of them are used to define class (like genus and species, etc.) and classify birds according to them.

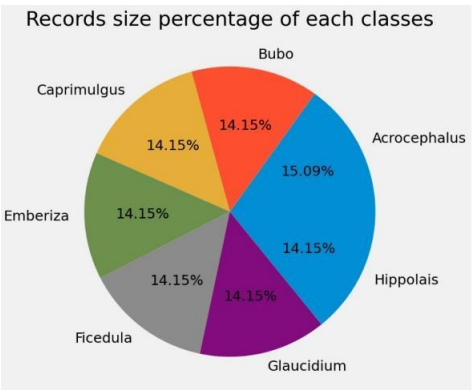



Fig.2.Visualization

### *B. Data/Sound pre-processing*

Following data collecting, sound recordings are pre-

processed. The WAV format is used to convert the MP3 files obtained from the dataset. These Wav files are normalized after being separated into equal-length segments. Threshold filtering is used to create chunks/segments with a high amplitude and no noise or disruption. The waveform was created using power spectral density (PSD). The PSD was a metric that measured the amount of power per unit of frequency.

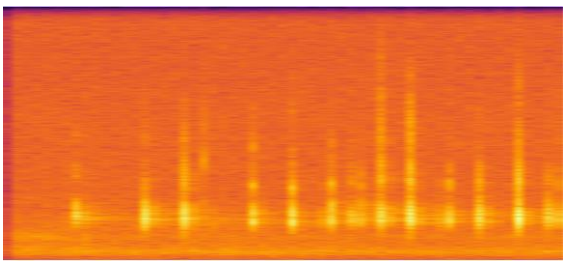


Fig. 3. Spectrogram

### C. *Classification with Neural Network*

An Artificial Neural Network (ANN) classification algorithm is a popular method for analyzing and recognizing bioacoustics signals. As a classification model, the multilayer

perceptron (MLP) is used. The MLP takes a set of predetermined attributes as input and produces a unique outcome for each bird species to be identified. Training and testing are the two steps in this identifying procedure. In the training process, syllables of specified bird sounds were utilized to train the multilayer perceptron, resulting in the right MLP output being triggered. The training process is carried out by repeatedly delivering known sounds to the network and then iteratively adjusting the network's weighting. The goal of this training is to lower the total error between the supplied and expected results till a predefined error requirement is accomplished.

### D. *Output*

For the output, user can use GUI i.e., Graphical User Interface to analyses the species of the bird. With the help of GUI user can record, process and show the outcome.

## V. RESULT

The waveform of a bird's voice can be obtained using MATLAB software. Dataset can be individually captured in mp3 format, which must be converted to a.wav file, and it is evident that each bird has a distinctive voice. Through a process of training, an ANN is tuned for a specific purpose, such as pattern classification.

Following table shows the output

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| Acrocephalus | 0.92 | 0.72 | 0.81 | 200 |
| Bubo | 0.70 | 0.94 | 0.80 | 200 |
| Caprimulgus | 1.00 | 0.88 | 0.94 | 200 |
| Emberiza | 0.89 | 0.87 | 0.88 | 200 |
| Ficedula | 0.76 | 0.76 | 0.76 | 200 |
| Glaucidium | 0.97 | 0.90 | 0.93 | 200 |
| Hippolais | 0.67 | 0.72 | 0.70 | 200 |
| accuracy | | | 0.83 | 1400 |
| macro avg | 0.84 | 0.83 | 0.83 | 1400 |
| weighted avg | 0.84 | 0.83 | 0.83 | 1400 |

Systematic recordings of outdoor noises are now possible thanks to automated audio recorders, which have recently opened up new opportunities for environmental conservation and restoration. Due to the fact that many bird species have extraordinarily high vocal frequency, audio recordings have become one of the most effective techniques to do study on them. Biological sounds of birds can give detailed and standardized data on the dynamics and distribution of wildlife habitats. Audio research and surveys are a good tool to analyze the species' density, abundance, and occupancy because many bird species produce distinct and consistent sounds.

Furthermore, picturized monitoring is problematic for many small and sensitive birds, enigmatic species, and species living in environments that environmentalists find difficult to reach. Bird audio tracking is also useful for other environmental operations, such as assessing the impact of wildfires and determining the extent of forest regeneration.

## VI. LIMITATIONS

Audio signals provide much more information about a bird, as it can be further classified into songs, calls, and sounds so noise in these various types of audios can be an issue. Having this kind of extra set of properties and classifications, also makes identification of birds little difficult.

## VII. FEASIBILITY STUDY

The Convolutional Neural Network (CNN) is a type of artificial intelligence that has been used to categorize bird sounds. Other academics have experimented with a variety of strategies for recognizing bird songs. Following points suggests the ways of using automated systems :-

- Automatic bird sound recognition from continuous environmental recordings would be a significant contribution to ornithology and biology research. Because manually identifying bird species takes a lot of time and money, this technology will be useful to government agencies and investigators.
- Birding is a favorite activity in many countries; thus, such systems offer a lot of potential for profit. The CNN can be run on hardware such as Raspberry Pi. These technology installations can assist environmental parks, preservation parks, and wildlife sanctuaries.
- An android application for a range of mobile devices can be designed and published, allowing users to use their cell phones as gadgets for bird sound detection and evaluation.
- The data gathered can be saved on a local hard drive or in the cloud. The data gathered will be extremely useful in studies of bird migration routes, total population,

## VIII. CONCLUSION

A model for automatic bird species identification using Convolutional Neural Networks, TF-Slim Tensor files, and spectrograms is proposed in this paper. The dataset's audios were discovered on the xeno-canto and Kaggle websites. The first step is to pre-process the bird vocals before creating a spectrogram of them. These were used to build the CNN based categorization algorithm. Bird species can be categorized depending on the spectrogram image created from their voices using this method.

## IX. FUTURE SCOPE

This technique allows for a greater number of classes to be worked on when identifying and classifying bird species, resulting in more accurate findings. Successfully using this software as a product can be extremely valuable. as a useful tool for estimating bird population size, recognizing natural habitats, and following a wide range of other species Environmentalists and wildlife admirers might also benefit from a user-friendly programmed. Also, because RNN has internal storage to remember its input, using it for categorization can improve accuracy.

## REFERENCES


[1] Vivekanandam, B. "Speedy Image Crowd Counting by Light Weight Convolutional Neural Network." Journal of Innovative Image Processing 3, no. 3 (2021): 208-222.

[2] Manoharan, J. Samuel. "Study of Variants of Extreme Learning Machine (ELM) Brands and its Performance Measure on Classification Algorithm." Journal of Soft Computing Paradigm (JSCP) 3, no. 02 (2021): 83-95

[3] Chandu B, A. M. (2020). Automated Bird Species Identification using Audio. 2020 International Conference on Artificial Intelligence and Signal Processing (AISP).

[4] M. M. M. Sukri, U. Fadlilah, S. Saon, A. K. Mahamad, M. M. Som and A. Sidek, "Bird Sound Identification based on Artificial Neural," IEEE 2020 IEEE Student Conference on Research andDevelopment (SCOReD), pp. 342-345, 2020.

[5] Incze, A., Jancso, H.-B., Szilagyi, Z., Farkas, A., & Sulyok, C. (2018). Bird Sound Recognition Using Convolutional Neural Network. IEEE 16th International Symposium on Intelligent Systems and Informatics, 000295–000300.

[6] Narasimhan, R., Fern, X. Z., & Raich, R. (2017). Simultaneous Segmentation And Classification Of Bird Song Using Cnn. IEEE Conference, 146-150.

[7] Guillermo Sarasa, A. G. (2017). An Approach of Algorithmic Clustering Based on String Compression to Identify Bird Songs Species in Xeno-canto Database. 3rd International Conference on Frontiers ofSignal Processing.

[8] C., A. P. (2015). Automatic Recognition of Birds Through Audio Spectral Analysis . Fifth International Conference on Advances in computing and Communication, 395–398.

[9] Jancovic, P., & Kokuer, M. (2015). Acoustic recognition of multiple bird species based on penalized maximum likelihood. IEEE Signal ProcessingLetters, 1-1.

[10] C., A. P. (2015). Automatic Recognition of Birds Through Audio Spectral Analysis . Fifth International Conference on Advances in Computing andCommunication, 395–398.

[11] Jancovic, P., & Kokuer, M. (2015). Acoustic recognition of multiple bird species based on penalised maximum likelihood. IEEE Signal ProcessingLetters, 1-1.

[12] Evangelista, T. L., Priolli, T. M., Jr, C. N., Angelico, B. A., & Kaestner, C. A. (2014). Automatic Segmentation of Audio Signals for Bird Species Identification. IEEE International Symposium on Multimedia,223–228.

[13] Jancovic, P., Kokuer, M., & Russell, M. (2014). Automatic Detection of Bird Species from Audio Field Recordings using HMM-based Modelling of Frequency Tracks. EUSIPCO, 8252– 8256.

[14] Graciarena, M., Delplanche, M., Shriberg, E., & Stolcke, A. (2011). Bird species recognition combining acoustic and sequence modeling. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 341-344.

[15] Lopes, M. T., Gioppo, L. L., Higushi, T. T., Kaestner, C. A., Silla Jr., C. N., & Koerich, A. L. (2011). Automatic Bird Species Identification for Large Number of Species. IEEE International Symposium on Multimedia, 117-122