International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 31 Issue 04 – AUGUST 2025.

Discovering Bird Species using Neural Models and Sound Signals

Dr. Roopa R[#1], N Basavanagowda[*2]

[#]Assistant Professor, Department of MCA, BIET, Davanagere, Karnataka, India
[*]Student, Dept. of MCA, BIET, Davanagere, Karnataka, India

***Abstract*— Many factors, including human activity, climate change, global warming, forest fires, deforestation, etc., are causing a significant shift in the bird population today. The use of machine learning algorithms for automatic bird species detection has made it feasible to monitor both the population and behavior of birds. This work develops an automatic bird identification system that does not require physical intervention because it takes a lot of time and effort to manually identify different bird species. Convolutional neural networks, as opposed to more conventional classifiers like SVM, Random Forest, and SMACPY, are employed to accomplish this goal. The primary objective is to use the dataset, which includes the voices of various birds, to identify the species of bird. The After pre-processing the input dataset—which includes framing, silence removal, and reconstruction—a spectrogram will be created and fed into a convolutional neural network, which will then be modified, tested, and classified. Birds are categorized based on their characteristics (size, color, species, etc.) after the output is generated and compared to pre-trained data. Monitoring the effects of human activity on the environment is now essential to preventing irreversible damage. One way to keep an eye on these effects is to monitor biodiversity, population dynamics, and animal breeding behavior.**

***Index Terms*— Convolutional neural networks (CNNs), machine learning, pre-processing, environmental monitoring, and bird species identification.**

I. INTRODUCTION

The rapid decline in bird populations worldwide is a pressing environmental concern, driven by various factors such as forest fires, global warming, climate change, human activity, and deforestation. As these challenges intensify, the need for effective monitoring and conservation strategies becomes increasingly critical. Traditional methods of bird species identification, which often rely on manual observation and expertise, are time-consuming and labor-intensive, making them impractical for large-scale applications. Researchers and conservationists can now more effectively monitor bird populations and behaviors thanks to new developments in machine learning and artificial intelligence that have created new opportunities for automatic bird species detection. The goal of this work is to create an automatic system for identifying birds by using Convolutional Neural Networks (CNNs) to analyze their vocalizations. The suggested system seeks to expedite the identification process by using a dataset of bird sounds, lowering the time and effort needed for manual classification. The suggested system improves classification efficiency and accuracy by combining multiple machine learning approaches, such as multilayer perceptrons (MLPs) and artificial neural networks (ANNs), building on current approaches. The architecture of the system is made to preprocess audio data, extract pertinent features, and categorize different bird species according to their distinctive vocal traits. This novel method not only overcomes the drawbacks of conventional systems but also offers a scalable way to track biodiversity and comprehend how human activity affects bird populations.

The deployment of automated bird species detection systems will be essential to conservation efforts as we work to lessen the negative effects of environmental changes because it will enable prompt interventions and well-informed decision-making. This introduction lays the groundwork for a thorough examination of the suggested and current systems, stressing their unique benefits and drawbacks in relation to the identification of bird species..

II. LITERATURE SURVEY

J. Salamon and J. P. Bello IEEE Signal Processing Letters, Deep convolutional neural networks (CNNs) are well suited for environmental sound classification because of their capacity to learn discriminative spectro-temporal patterns. However, the utilization of this family of high-capacity models has been hindered by the relative lack of labeled data. The first of this study's two main contributions is the deep CNN architecture we suggest for classifying environmental sounds. Second, we investigate the impact of various augmentations on the performance of the suggested CNN architecture and suggest using audio data augmentation to get around the issue of data scarcity. The suggested model yields state-of-the-art results for environmental sound classification when combined with data augmentation. We demonstrate that combining a deep, high-capacity model with an augmented training set results in better performance than either the suggested CNN without augmentation or a "shallow" dictionary learning model with augmentation. Lastly, we look at how each augmentation affects the model's classification accuracy for each class. We find that each augmentation has a different effect on the accuracy for each class, which suggests that class-conditional data augmentation could further enhance the model's performance.

V. Bisot, R. Serizel, S. Essid, and G. Richard. In order to train deep neural networks for environmental sound classification applications, this paper presents the use of representations based on nonnegative matrix factorization (NMF). Typically, deep learning systems for sound classification use networks to extract meaningful representations from hand-crafted features or spectrograms. Rather, prior to training deep networks, we present an NMF-based feature learning stage, the value of which is emphasized in this paper, particularly for multi-source acoustic environments like sound scenes. To learn better input representations for deep neural networks, we use two well-established unsupervised and supervised NMF techniques. This will enable us, to achieve competitive performance with more intricate systems, like convolutional networks for the classification of acoustic scenes, using straightforward architectures. The suggested systems perform better than the top systems from the 2016 DCASE challenge and neural networks trained on time-frequency representations on two acoustic scene classification datasets.

M Ashraful Amin, Amin Ahsan Ali, Partha Protim Roy, and Sujoy Debnath "Identifying Bird Species through Their Voices," 5th International Conference on Informatics, Electronics, One of the enigmatic phenomena that we are learning more about over time is biodiversity. The BIOTOPE society has launched an initiative to help people understand nature. It is an open platform that brings researchers together to identify bird species based on their calls or singing. This work's primary concept is to recognize different bird species based on their calls or songs. This article's work investigates the automated technique for identifying bird species from their calls and songs. This will facilitate the archiving process. Sand is used to identify new species or subspecies. Additionally, it will support ecological and behavioral research on birds as well as population surveillance. In this study, we put into practice a system that can automatically recognize different species of birds from sound recordings. According to the ROC performance analysis of the proposed system, we can obtain an area under the curve of 0.96 by employing Random Decision Tree.

Hua-An-Zhao, Yihenew Wondie Marye, and Rong Sun, "FFT Based Automatic Species Identification Improvement with 4-layer Neural Network," 13th International Symposium on Communications and Information Technologies (ISCIT), This paper presents the development of an automatic species identification system. An automatic process was used to segment, process, remove features, and identify the recoded data. Using a 4-layer neural network, a feature quantity method based on FFT with derivative of frequency band power is suggested. The identification of wild bird species based on sound data has shown promise when compared to the results of a 4-layer neural network.

S. D. H. Permana, G. Saputra, B. Arifitama, W. Caesarendra, and R. Rahim, "Classification of bird sounds as an early warning method of forest fires using convolutional neural network (cnn) algorithm," King Saud University-Computer and Information Sciences Journal, Forest fires occur annually in the tropical nation of Indonesia. An extended summer season is the cause of forest fires. People who live near forests may suffer from respiratory illnesses and visual impairments as a result of the thick smoke from forest fires, which are a common occurrence in Indonesia. Neighboring nations also experience the effects of forest fire fumes when the fire spreads over a larger land area, including Brunei Darussalam, Singapore, and Malaysia. Birds are among the many animals that call forests home. Birds can use sound to communicate with their colonies. Bird sounds can be used to communicate with the group through calls, marriage invitations, and forest fire warnings. This study uses the Convolutional Neural Network (CNN) method, one of the Deep Learning (DL) algorithms, to classify bird sounds.

T. Shaw, R. Hedes, A. Sandstrom, A. Ruete, M. Hiron, M. Hedblom, S. Eggers, and G. Mikusinski, “Hybrid bioacoustic and ecoacoustic analyses provide new links between bird assemblages and habitat quality in a winter boreal forest,” Environmental and Sustainability Indicators, vol. 11, p. 100141, 2021. Particularly in multipurpose forests that are also used for timber production, resident birds in boreal forests are frequently species of conservation interest and can act as indicators of habitat quality. We must first identify the structural elements that offer resident birds habitats in order to make well-informed decisions about forest management. This is especially true during the winter, which is a crucial and little-studied season for their survival. In order to draw conclusions about which stand structural features support and enhance the likelihood of winter survival, the study set out to develop trustworthy techniques for tracking bird presence and activity throughout the winter. We investigated whether acoustic recordings could reveal connections between bird diversity using a hybrid bioacoustic and ecoacoustic method. and structural complexity components, and contrasted these findings with the conventional point count approach. Between December 2019 and February 2020, we gathered acoustic recordings, point count surveys, and vegetation surveys at 19 locations in a Swedish boreal forest.

R. Mohanty, B. K. Mallik, and S. S. Solanki, "Species recognition using spike model with bird dataset," Data in Brief, It has long been acknowledged that birds are the first to notice changes in the climate in our surroundings. Identification of bird species has become crucial, not only for assessing bird survival but also as a precursor to the planet's deteriorating climate and health. Previous studies have established the ability to identify different bird species based on region-specific sounds from online repositories. The spike-based bird species recognition model, which addresses the process of identifying the bird species based on their vocalization or call, was introduced in this article. The dataset includes 14 vocalizations of bird species. These recordings were made in their natural habitat. The Central Poultry Development Organization (CPDO), Eastern Region, Bhubaneswar, India, used a digital recorder and a unidirectional microphone to capture the calls.

M. Munk, and L. Juranek. The project of Automatic Bird Species Recognition Based on Bird Vocalization is the subject of this paper. Analysis was done on eighteen bird species from six different families. Initially, specific recordings were used to calculate human factor cepstral coefficients that represented the given signal. Subsequently, segments of bird vocalizations were identified using the voice activity detection system. From these, individual hidden Markov models were used to calculate a likelihood rate, which indicates which model the given code value corresponds to. Only one hidden Markov model was trained for each species of bird. The 81.2% interspecific success rate has arrived. The success rate for family classification is 90.45%.

III. Existing System

For the purpose of crowd counting, Vivekanandam et al. have employed the Lightweight CNN (LWCNN) architecture with VGG. VGG-16 has three max-pooling layers and ten convolution layers in the front end. With a dilation factor of two and a compressed convolution depth of six, it is possible to accurately count the number of people in a crowd.
Extreme Learning Machine (ELM) algorithms, including Evolutionary ELM, Voting based ELM, Ordinal ELM, Fully complex ELM, Symmetric ELM, and others, have been proposed by Manoharan et al. as a way to address the slow computation of Feed-Forward Neural Networks. The ELM-based classification algorithm has a 94.10% accuracy rate. Four bird sound recordings—the cuckoo, sparrow, crow, and laughing dove each with an input space of 100 recordings make up the dataset that Chandu B used. The recordings of bird sounds were gathered from xeno-canto.com, a website dedicated to the exchange of bird sounds from around the globe. To maintain diversity and prevent overfitting, each clip, which ranges in length from 5 to 20 seconds, is converted to a fixed sampling frequency of 44100Hz or 48000Hz. The Google Recording and LibriSpeech ASR datasets provided the data for these examples.

Limitations of Existing System: The current system is time-constrained; analyzing bird sounds takes a long time. Typically, the current system needs a lot more data. A large amount of training data is needed. Additionally, the computational cost of the current system is higher. Training the dataset from scratch using the current system can take weeks. Additionally, more processing power is required.

IV. Proposed System

Data collection from the dataset, which is acquired from Kaggle, is the first stage of implementation. This resource contains the bird audio recordings in.wav format. Audio recordings of the birds in.wav format are included in this dataset. Users can upload their own recordings to Kaggle, an open dataset website. Since the dataset contains a large number of defined features, class definitions (such as genus and species) and bird classifications are based on the combination of these features. One common technique for examining and identifying bioacoustics signals is the classification algorithm of an Artificial Neural Network (ANN). The multilayer perceptron (MLP) is employed as a classification model. The MLP generates a distinct result for each bird species to be identified based on a set of preset attributes. This identifying process consists of two steps: training and testing. The multilayer perceptron was trained using syllables of specific bird sounds, which caused the correct MLP output to be triggered during the training process. The network is repeatedly exposed to known sounds during the training process, and the weighting of the network is then iteratively adjusted. This training's objective is to reduce the overall error between the provided and anticipated results until a predetermined.

The error requirement has been met. The output allows the user to analyze the bird species using a Graphical User Interface, or GUI. The user can upload the dataset, process it, and view the results with the aid of the GUI.

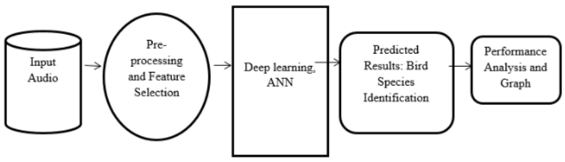

Fig: Architecture Digram

Advantages

The suggested system analyzes bird sounds in a shorter amount of time. Additionally, the suggested system can manage fewer training data points. Comparing the suggested system to the current system models, there is no increase in cost. Compared to the current system model, the suggested system might only require a few hours to train the dataset entirely from scratch. Only a smaller amount of processing power is needed by the suggested system. The suggested system can withstand faults well. Additionally, the suggested has a good distributed memory.

V. Implementation

Login Module: The project features a secure user login system that enables users to access personalized functionalities. After logging in, users can upload audio recordings of bird calls or songs through an intuitive interface. The audio files are then processed by the system to automatically predict and identify the bird species present in the recordings. This login mechanism ensures that only authorized users can utilize the audio upload and bird identification features, providing a tailored and protected environment for monitoring and research activities.

Performance Analysis: The project includes a comprehensive performance analysis module that evaluates the effectiveness of the bird species identification system. During training and testing phases, accuracy metrics are computed to evaluate the model's performance learns from the dataset and performs on unseen data. The analysis provides insights into the model's reliability and robustness by comparing training accuracy against test accuracy. These performance indicators help in fine-tuning the model, validating its predictive power, and demonstrating the system's capability to accurately classify various bird species according to their vocalizations.

VI. RESULT

In the results section of the project, a visual depiction of the performance analysis is presented, showcasing both training and test accuracy metrics. The graph illustrates the model's performance over multiple epochs, highlighting how the accuracy evolves as the training progresses. The training accuracy curve indicates the model's ability to learn from the training dataset, reflecting improvements as it adjusts its parameters to minimize error. Conversely, the test accuracy curve demonstrates the model's performance on unseen data, providing a crucial measure of its generalization capability. By comparing these two curves, we can assess whether the model is overfitting or underfitting. Ideally, both training and test accuracies should converge, indicating that the model is effectively learning the underlying patterns in the data without memorizing specific examples. The results not only confirm the suggested bird species identification system's efficacy while also providing insights into areas for potential improvement, guiding future iterations of the model for enhanced accuracy and reliability.

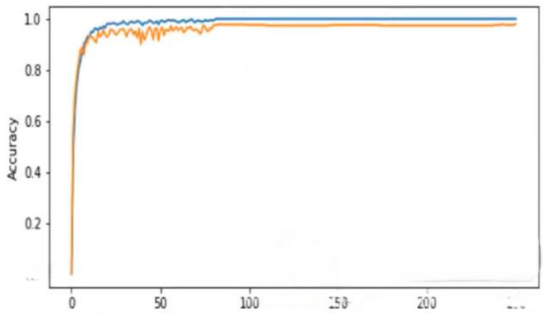

Fig: Resultant graph

VII. CONCLUSION

In summary, the creation of a system for automatically identifying bird species using machine learning techniques—specifically, Convolutional Neural Networks (CNNs)—represents a substantial breakthrough in the field of biodiversity monitoring and conservation. This project addresses the pressing need for efficient and scalable methods to track bird populations and behaviors in the face of environmental challenges such as climate change, habitat loss, and human intervention. By leveraging audio recordings of bird calls and songs, the proposed system not only streamlines the identification process but also enhances the precision and dependability of species classification. Combining pre-processing methods with the application of Artificial Neural Networks (ANNs) further optimize the model's performance, allowing it to operate effectively with less training data and reduced computational requirements compared to existing systems. The results of the performance analysis demonstrate the system's capability to achieve high accuracy in both training and testing phases, validating its potential as a valuable tool for researchers and conservationists. As we continue to face the consequences of environmental degradation, the implementation of such automated systems will be crucial in facilitating timely interventions and informed decision-making to protect avian biodiversity. Overall, this work not only contributes to the sympathetic of bird species identification but also emphasizes the position of technological innovation in addressing ecological challenges. Future research can build upon this foundation, exploring further enhancements and applications of machine learning in wildlife conservation and environmental monitoring.

VIII. REFERENCES

[1] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," IEEE Signal Processing Letters, vol. 24, pp. 279–283, 2021.

[2] "Using deep neural networks with nonnegative representations for enhanced environmental sound classification," IEEE International Workshop on Machine Learning for Signal Processing MLS, 2019; V. Bisot, R. Serizel, S. Essid, and G. Richard.

[3] Identification of Bird Species from Their Singing by Sujoy Debnath, Partha Protim Roy, Amin Ahsan Ali, and M Ashraful Amin, 5th International Conference on Informatics, Electronics, and Vision (ICIEV), 2019 .

[4] Hua-An-Zhao, Yihenew Wondie Marye, and Rong Sun, "FFT Based Automatic Species Identification Improvement with 4-layer Neural Network," 13th International Symposium on Communications and Information Technologies (ISCIT), 2018.

[5] S. D. H. Permana, G. Saputra, B. Arifitama, W. Caesarendra, and R. Rahim, "Classification of bird sounds as an early warning method of forest fires using convolutional neural network (cnn) algorithm," King Saud University-Computer and Information Sciences, vol. 34, no. 7, pp. 4345–4357, 2022.

[6] T. Shaw, R. Hedes, A. Sandstrom, A. Ruete, M. Hiron, M. Hedblom, S. Eggers, and G. Mikusinski, "New connections between bird populations and habitat quality in a winter boreal forest are established through hybrid bioacoustic and ecoacoustic analyses," Environmental and Sustainability Indicators, vol. 11, p. 100141, 2021.

[7] R. Mohanty, B. K. Mallik, and S. S. Solanki, "Species recognition using spike model with bird dataset," Data in Brief, vol. 29, p. 105301, 2020.

[8] Automatic bird species recognition based on bird vocalization, EURASIP Journal on Audio, Speech, and Music Processing, vol. 2018, no. 1, pp. 1–7, 2018; J. Stastny, M. Munk, and L. Juranek