e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

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

Vidyashree KP*1

*1Post-Graduation Student, Department Of MCA PES Institute Of Technology And Management, Shivamogga Karnataka, India.

ABSTRACT

This research presents a new approach to identifying bird species using audio data processing and convolutional neural networks (CNN). Following their collection, recordings of birdsong were processed to extract crucial elements known as mel-frequency ceptral coefficient(MFCC). These characteristics were then applieds to train a CNN model designed to accurately classify different species of birds. Across a variety of datasets, our methodology greatly increased classification accuracy when compared to traditional methods. The results illustrate the high degree of coherence between processing audio signal and deep learning for bioacoustic applications, providing new opportunities for automated wildlife monitoring and conservation efforts.

Keywords: Bioacoustic, Conventional, Auditory Signal, MFCC.

I. INTRODUCTION

In bioacoustics, automatically identifying bird species identified by their calls is a difficult but important issue. The complex and species-specific vocalizations used by birds to communicate make their auditory signals a valuable resource for ecological monitoring and species identification. It takes a lot of effort, specialized knowledge, and is frequently not feasible for large-scale investigations to use manual identification methods. On the other hand, deep learning and automated methods that take use of recent developments in audio signal processing present viable answers to this problem. In recent years, there have been substantial breakthroughs in the application of CNNs or convolutional neural network for sound classification challenges, owing to their success in picture identification, to solve audio classification issues. These developments have made it possible to investigate CNNs in the field of bioacoustics, especially for the identification of species from sound recordings. Robust and reliable classification has been made possible by techniques like the extraction of Mel-frequency cepstral coefficients (MFCC), which have been essential raw audio data transformed into Feature depictions those are compatible with CNNs. The main goals of current research on automated identification of bird species are to increase classification accuracy, cover a wider range of species, and create models that can withstand noise and variability found in real-world environments. Research has shown that it is possible to use CNNs to identify between bird species with high accuracy based only on their vocalizations, which represents a paradigm shift in the monitoring and conservation of wildlife.

With processing audio signal and CNN, this study presents a thorough methodology for automated identification of bird species, hoping to make a contribution to this rapidly developing subject. Our technique aims to enhance the effectiveness and precision of species identification by utilizing developments in machine learning and signal processing. This will make it possible to apply ecological studies and conservation biology more broadly.

II. LITERATURE REVIEW

Automated Identification of bird species using Neural Networks and Audio Signal Processing Authors: Chaitra Nagara, Kartik S. Murthy, Chandu B., Akash Munikoti, and Ganesh Murthy V. Their research examines bird identification techniques and develops an automated system that can recognize various bird species. Undertaking in-depth study on taxonomy and other subfields of ornithology to enable identification of bird sounds without human assistance has proven to be a difficult and challenging undertaking. A two-stage identification approach is used in this research. The first step was to create an ideal dataset that included sound recordings of every species of bird. Following that, the sound clips were subjected to several sound preprocessing procedures, including pre-emphasis, silence removal, reconstruction, and framing.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

Cheng et al.'s approach use discriminative features in conjunction with SVM and a Normal Bayes classifier to categorize bird species based on their anatomical traits. A different researcher, Marini et al. suggested a method to construct normalized color histograms to be able to extract feature vectors for classification and remove background objects using color segmentation.

Using deep neural networks and dynamic kernel-based support vector machines, Paawan Mukker, Deep Chakraborty, Padmanabhan Rajan, and A. D. Dileep identify bird calls. For this investigation, we using voice and audio processing methods for avian vocalization analysis and categorize species found in the regions of lower Himalayans. MFCC are taken from every recording. These recordings are displayed as feature vector that having different lengths as a result, these patterns of varied lengths formed from voice signals are commonly classified using deep neural networks (DNNs) and dynamic kernel-based SVMs. In this piece of writing, we explores to classify feature vector sets representing sounds of birds using dynamic kernel-based SVMs and DNNs.

III. SYSTEM ARCHITECTURE

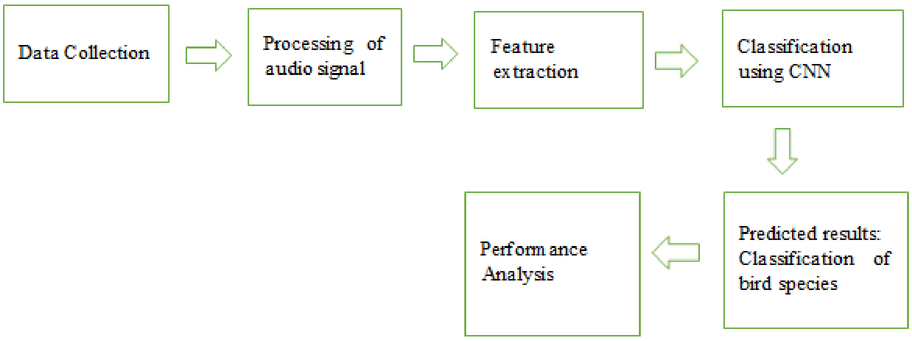

Fig 1: System Architecture

IV. METHODOLOGY

To construct a diverse dataset, by collecting all the recordings of the bird sounds from multiple online archives and field recordings. After that, these recordings undergo preprocessing to guarantee consistency, which includes splitting lengthy recordings into segments and translating them to a standard format. After that, we use a method called Mel-frequency cepstral coefficients (MFCC) extraction, which converts the audio signals into a collection of feature vectors to be able to extract the key elements of bird vocalizations. To obtain the MFCCs, this process entails pre-emphasis filtering, framing, windowing, Fourier Transform, and mapping the power spectrum to the Mel scale. Finally, the Discrete Cosine Transform (DCT) is applied.

A CNN developed specially for this use receives various attributes that have been removed. The CNN design is made up of fully connected layers, pooling layers to lower dimensionality, numerous ReLU-activated convolutional layers with a softmax output layer for classification. To improve the robustness of the model, the dataset is divided into training, validation, and test sets.

To supplement the data, methods including pitch shifting, noise adding, and temporal stretching are used. With early halting based on validation performance to avoid overfitting, the model have been trained by using the Adam optimizer and categorical cross-entropy loss function. Accuracy,F1-score,recall,precision and a confusion matrix are metrics used to evaluate performance when evaluating classification across species. The trained model's enhanced performance in identifying bird species is then demonstrated by comparing it with traditional machine learning models.

V. CONCLUSION

The study shows that accuracy in bird species identification is significantly better when audio signal processing is paired with CNNs or convolutional neural networks. The system outperformed conventional techniques with excellent classification accuracy because to the use Mel-frequency cepstral coefficients (MFCC) are utilized, and a sturdy CNN architecture. Data augmentation increased robustness and ensured good performance under a range of circumstances. Large-scale ecological research are made possible by this method's scalable solution for wildlife protection and monitoring. To enhance automatically identification of bird species and broaden the applications of bioacoustic classification, upcoming investigation includes augmenting the dataset and improving the model.

VI. FUTURE SCOPE

The model's generalization can be improved by expanding the collection of data in future studies to include other bird species and a wider variety of vocalizations. Accuracy and resilience might be increased even further by honing the CNN architecture and implementing cutting-edge audio processing methods. In-depth solutions for wildlife conservation might also be obtained by investigating real-time deployment for field deployment and combining the system with additional ecological monitoring instruments. Research and management of the environment and biodiversity could be greatly expanded by applying this technology to additional bioacoustic applications, such as monitoring the noises of the surrounding environment or other wildlife species.

VII. REFERENCES

[1] "Automated identification of bird species using audio signal processing and neural networks," B. Chandu and Akash Munikoti IEEE 2020 AISP (artificial intelligence and signal processing) sinternational conference, DOI: 10.1109/AISP48273.2020.9073584.

[2] Samruddhi Bhor, Ruthuja Ganage, "Identification of Bird species using audio processing" IEEE 2022 International Conference on Electronics and Renewable Systems (ICEARS) 16-18 March.

[3] Narasimhan, R., Plant, X. Z., and Raich, R. (2017). Synchronous Division And Arrangement Of Bird Melody Utilizing Cnn. IEEE Gathering, 146-150.

[4] P. Somervuo, A. Harma, IEEE Transactions on Audio, Speech, and Language / S. Fagerlund, Parametric representations of bird noises for automatic species recognition Processing, vol. 14, no. 6, pp. 22522263, 2006.

[5] Aarti Madhavi, R. P. (2018). Profound Learning Based Sound Classifier for Bird Species. IJSDR.