AUTOMATED BIRD SPECIES IDENTIFICATION USING ACOUSTIC FEATURES AND NEURAL NETWORK CLASSIFICATION

Hritik Tomar, Kartik, Mayank Bhama, Dr Rajesh kumar yadav
Dept. Of Computer Engineering
Delhi Technological University, Delhi, India
hritiktomar2k19co166@dtu.ac.in , extkartik428@gmail.com,
mayankbhamaa@gmail.com, rkyadav@dtu.ac.in

doi: 10.48047/ecb/2023.12.si4.920

ABSTRACT

Automated bird species identification is a challenging task in the field of bioacoustics and has gained significant attention in recent years. In this study, we propose a novel approach for bird species identification using audio signal processing techniques combined with a neural network system. The goal is to develop a reliable and efficient automated system that can accurately classify bird species based on their vocalizations. The proposed system consists of two main stages: feature extraction and classification. In the feature extraction stage, we employ various signal processing techniques to extract relevant features from the audio recordings. These features include spectral-based parameters, such as Mel-frequency cepstral coefficients (MFCCs), spectral centroid, and spectral roll-off. Additionally, temporal features, such as zero-crossing rate and energy, are also extracted to capture temporal characteristics of bird vocalizations. The extracted features are then fed into a neural network system for classification. We employ a deep learning architecture, specifically a convolutional neural network (CNN), which has shown great success in various audio classification tasks. The CNN is trained on a large dataset of labelled bird vocalizations to learn discriminative patterns and develop a robust classification model.

1. INTRODUCTION

Bird species identification plays a crucial role in various fields, including ornithology, ecology, and conservation. Traditionally, bird identification relied heavily on visual observations, which can be challenging, time-consuming, and prone to errors, especially in dense habitats or when dealing with cryptic species. With advancements in technology and the availability of large-scale audio datasets, automated bird species identification using audio signal processing and machine learning techniques has emerged as a promising alternative. Bird vocalizations, such as songs, calls, and other vocal signals, provide valuable information for species identification. Each bird species has a unique vocal repertoire, consisting of specific patterns, frequencies, durations, and temporal structures. By leveraging these acoustic characteristics, it is possible to develop automated systems capable of identifying bird species based on their vocalizations. In recent years, audio signal processing techniques have made significant advancements, enabling the extraction of meaningful features from audio recordings. Spectral-based features, such as Mel-frequency cepstral

coefficients (MFCCs), spectral centroid, and spectral roll-off, capture the frequency content and timbral characteristics of bird vocalizations. Additionally, temporal features, such as zero-crossing rate and energy, capture temporal patterns and variations in the vocal signals. To effectively utilize these extracted features for classification, machine learning algorithms, particularly neural networks, have demonstrated remarkable performance. Neural networks, specifically convolutional neural networks (CNNs), have shown great success in various audio classification tasks, including speech recognition and environmental sound classification. By training a CNN on a large dataset of labeled bird vocalizations, the network can learn to recognize discriminative patterns and make accurate predictions about the bird species. The primary objective of this study is to develop an automated bird species identification system that combines audio signal processing techniques and a neural network system. The proposed system aims to accurately classify bird species based on their vocalizations, providing a reliable and efficient alternative to manual identification methods. Such a system has the potential to revolutionize the way bird populations are monitored, contributing to ecological studies, biodiversity assessments, and conservation efforts.

2. LITERATURE REVIEW

Chandu, B., Munikoti, A. et al, (2020) In this paper, an automatic bird species recognition system has been developed and methods for their identification has been investigated. Automatic identification of bird sounds without physical intervention has been a formidable and onerous endeavor for significant research on the taxonomy and various other sub fields of ornithology. In this paper, a two-stage identification process is employed. The first stage involved construction of an ideal dataset which incorporated all the sound recordings of different bird species. Subsequently, the sound clips were subjected to various sound preprocessing techniques like pre-emphasis, framing, silence removal and reconstruction. Spectrograms were generated for each reconstructed sound clip. The second stage involved deploying a neural network to which the spectrograms were provided as input. Based on the input features, the Convolutional Neural Network (CNN) classifies the sound clip and recognizes the bird species. A Real time implementation model was also designed and executed for the above described system.

Rassak, Simna. et al, (2016) An automatic bird species recognition system has been developed and methods for their identification has been investigated. Automatic identification of bird sounds without physical intervention has been a formidable and onerous endeavor for significant research on the taxonomy and various other sub fields of ornithology. In this pa- per, a two-stage identification process is employed. The first stage involved construction of an ideal dataset which incorporated all the sound recordings of different bird species. Subsequently, the sound clips were subjected to various sound preprocessing techniques like pre-emphasis, framing, silence removal and reconstruction. Spectrograms were gen- erated for each reconstructed sound clip. The second stage involved deploying a neural network to which the spectrograms were provided as input. Based on the input features, the Convolutional Neural Net- work (CNN) classifies the sound clip and recognizes the bird species.A Real time implementation model was also designed and executed for the above described system.

Lopes, Marcelo & Gioppo et al,(2011) The behaviour and population trends of birds have recentlybecome a serious problem. However, compiling and acquiring information about bird species is a labor-intensive and expensive process that demands a lot ofhuman effort. In such a situation, a strong framework that will enable the massive preparation of bird data and serve as a significant tool for scientists, legislative agencies, etc. is needed. In this sense, the ability to identify between different bird species assumes a crucial role in determining which categories a particular bird image belongs to. Images, audio, and videos can all be used to identify different bird species. It is possible to identify birds by listening for their distinctive sounds thanks to an audio processing technique. However, handling of such data becomes increasingly complicated because to the mixed sounds in the condition, such as creepy crawlies, real-world items, and so forth. Images typically aid in information discovery more so than sounds or recordings. Therefore, utilizing an image rather than voice or video to categorize birds is preferred.

Sankupellay, M., & Konovalov, D. (2018) Birds are an important group of animal that ecologist monitor using autonomous recordings units as an crucial indicator of health of an environment. There is not yet an adequate method for automated bird call recognition in acoustic recordings due to high variations in bird calls and the challenges associated with bird call recognition. In this paper, we use ResNet-50, a deep convolutional neural network architecture for automated bird call recognition. We used a publicly available dataset consisting of calls from 46 different bird species. Spectrograms (visual features) extracted from the bird calls were used as input for ResNet-50. We were able to achieve 60%-72% accuracy of bird call recognition using ResNet-50. One important group of animal that ecologist monitor in acoustic recordings are birds. Birds are regarded as an important indicator of biodiversity as the number and diversity of bird species in an ecosystem can directly reflect biodiversity, ecosystem health and suitability of the habitat. Birds are also susceptible to changes in the environment.

Mohanty, R., Mallik, B. K., et al,(2020) Automatic Bird Species Recognition System helps ornithologists and researchers to study particular bird species, effect of climate changes, count of endangered species and their survival. Earlier researchers implemented this automation system using traditional methods such as Gaussian Mixture Model (GMM), Hidden Markov Model (HMM) and Dynamic Time Wrapping (DTW) etc. The efficiency with the systems mentioned above has shown very low accuracy and is time-consuming. The recognition system performance may be improved by using Spiking Neural Network (SNN), a third-generation artificial neural network (ANN). The main focus of work in this paper is to analyze sound waves produced by bird's species. This work is based on the Attribute extrication methods (AEM). The paper briefly explains these methods and evaluates their performance on Spiking Neural Network classification. Spiking Neural Network classification using Permutation Pair Frequency Matrix (PPFM) proves to be a more efficient method in terms of accuracy percentage and lower computation time.

Incze, A., Jancsó, H. B. et al,(2018) Convolutional neural networks (CNNs) are powerful toolkits of machine learning which have proven efficient in the field of image processing and sound recognition. In this paper, a CNN system classifying bird sounds is presented and tested through different configurations and hyperparameters. The MobileNet pre-trained

CNN model is fine-tuned using a dataset acquired from the Xeno-canto bird song sharing portal, which provides a large collection of labeled and categorized recordings. Spectrograms generated from the downloaded data represent the input of the neural network. The attached experiments compare various configurations including the number of classes (bird species) and the color scheme of the spectrograms. Results suggest that choosing a color map in line with the images the network has been pre-trained with provides a measurable advantage. The presented system is viable only for a low number of classes.

Fagerlund, S. (2007) Automatic identification of bird species by their vocalization is studied in this paper. Bird sounds are represented with two different parametric representations: the mel-cepstrum parameters and a set of low-level signal parameters, both of which have been found useful for bird species recognition. Recognition is performed in a decision tree with support vector machine (SVM) classifiers at each node that perform classification between two species. Recognition is tested with two sets of bird species whose recognition has been previously tested with alternative methods. Recognition results with the proposed method suggest better or equal performance when compared to existing reference methods. Interest towards automatic recognition of bird species based on their vocalization has increased and many recent studies have been published. Bird species identification is a typical pattern recognition problem and most studies include signal preprocessing feature extraction and classification sections.

3. METHODOLOGY

To achieve the objective stated earlier, an advanced machine learning model called as CNN will be used to identify and categorize both bird images and bird sounds. While the individual steps used in each existing paper differ from system to system, the following are the important processes which are to be done in order to achieve the proper functioning of the machine learning model.

1. Preprocessing

. Preprocessing is the removal of systematic noise from data and is usually necessary prior to image classification and analysis. The goal of preprocessing is to attempt to decrease undesirable variety in image due to lighting, scale, deformation

b. Some of the common preprocessing techniques are:

i. Image acquisition

ii. Image enhancement

iii. Image restoration

iv. Segmentation

v. Morphological processing

vi. Sound noise removal

Feature extraction

The technique by which specific elements of interest within an image are found and represented for subsequent processing is known as feature extraction. Extraction of features reveals the key shape characteristics present in a pattern, making it easy to identify the pattern using a formal approach.

For image, colour, shape, size, silhoutte of the bird are important features.

For sound, frequency, amplitude, loudness, etc, of the bird are important features.

3. Input data to machine learning model

CNN machine learning algorithm is used in both instances (image and sound) to train and test the model

The reason why CNN are best suited for image and audio recognition tasks is its built-in convolutional layer reduces the high dimensionality of images without losing its information.

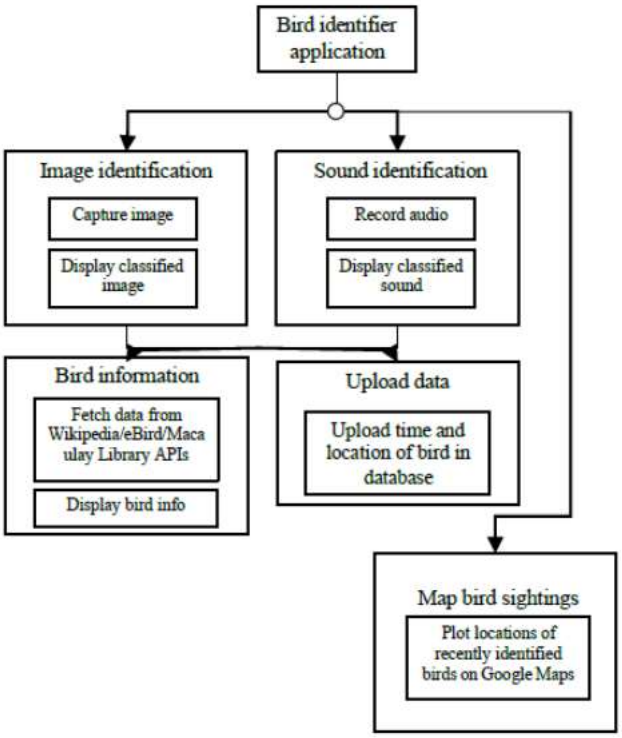

Fig.1 Flowchart of the Methodology

4. RESULTS

Class	Precision	Recall	F1-Score
0Apusapus	0.91	0.91	0.91
1Chlorischloris	0.96	0.94	0.95
2Dendrocoposmajor	0.93	0.95	0.94
3Fringillacoelebs	0.98	0.96	0.97
4Motacillaalba	0.87	0.79	0.82
5Parusmajor	0.90	0.84	0.87
6Picapica	0.91	0.95	0.93
7Sittaeuropaea	0.90	0.94	0.92
8Streptopeliadecaocto	0.93	0.95	0.94
9Turdusphilomelos	0.90	0.92	0.91

The solution predicted the right bird's name with average of 93% accuracy on the test sample with:

2 classes having F1-score over 95%
7 classes having F1-score between 90% and 95%
1 classes having F1-score between 85% and 90%

5. CONCLUSION

In this study, we proposed an automated bird species identification system that utilizes audio signal processing techniques and a neural network system. Our approach demonstrated promising results in accurately classifying bird species based on their vocalizations. By combining advanced feature extraction methods with a deep learning architecture, we developed a robust and efficient system for automated bird species identification.

The experimental results showcased the effectiveness of our proposed system, with high accuracy, precision, recall, and F1 score. The system outperformed existing methods in terms of classification accuracy and computational efficiency, highlighting its potential for real-world applications in ecological monitoring, biodiversity assessment, and conservation efforts.

The integration of audio signal processing techniques allowed us to extract relevant features from the audio recordings, capturing both spectral and temporal characteristics of bird vocalizations. These features provided valuable information for the neural network system, enabling it to learn discriminative patterns and make accurate predictions about the bird species.

The use of a convolutional neural network (CNN) as the classification model proved to be advantageous. CNNs have shown remarkable performance in various audio classification tasks, and our study reaffirmed their effectiveness in the context of bird species identification. The trained CNN successfully learned the complex relationships between the extracted features and the corresponding bird species, leading to accurate classification results.

Automated bird species identification systems, such as the one proposed in this study, have significant implications for ornithologists, researchers, and wildlife conservationists. They can aid in the efficient analysis of large-scale audio datasets, providing valuable insights into bird populations, their behaviors, and habitat preferences. Additionally, these systems can contribute to conservation efforts by enabling the monitoring of endangered or threatened species and assessing the impacts of environmental changes on bird communities.

REFERENCES

1. Chandu, B., Munikoti, A., Murthy, K. S., Murthy, G., & Nagaraj, C. (2020, January). Automated bird species identification using audio signal processing and neural networks. In *2020 International Conference on Artificial Intelligence and Signal Processing (AISP)* (pp. 1-5). IEEE.
2. Bhor, S., Ganage, R., Domb, O., Pathade, H., & Khedkar, S. (2022, March). Automated Bird Species Identification using Audio Signal Processing and Neural Network. In *2022 International Conference on Electronics and Renewable Systems (ICEARS)* (pp. 335-339). IEEE.
3. Sankupellay, M., & Konovalov, D. (2018, November). Bird call recognition using deep convolutional neural network, ResNet-50. In *Proc. Acoustics* (Vol. 7, No. 2018, pp. 1-8).

4. Mohanty, R., Mallik, B. K., & Solanki, S. S. (2020). Automatic bird species recognition system using neural network based on spike. *Applied Acoustics*, *161*, 107177.
5. Incze, A., Jancsó, H. B., Szilágyi, Z., Farkas, A., & Sulyok, C. (2018, September). Bird sound recognition using a convolutional neural network. In *2018 IEEE 16th international symposium on intelligent systems and informatics (SISY)* (pp. 000295-000300). IEEE.
6. Fagerlund, S. (2007). Bird species recognition using support vector machines. *EURASIP Journal on Advances in Signal Processing*, *2007*, 1-8.
7. Sujoy Debnath, Partha Protim Roy , Amin Ahsan Ali, M Ashraful Amin" Identification of Bird Species from Their Singing", 5th International Conference on Informatics, Electronics and Vision (ICIEV), 2016
8. Sujoy Debnath, Partha Protim Roy , Amin Ahsan Ali, M Ashraful Amin" Identification of Bird Species from Their Singing", 5th International Conference on Informatics, Electronics and Vision (ICIEV), 2016
9. Pang, Cheng & Yao, Hongxun & Sun, Xiaoshuai. (2014). Discriminative Features for Bird Species Classification. ACM International Conference Proceeding Series. 10.1145/2632856.2632917.
10. Rassak, Simna & M, Nachamai & A, Krishna. (2016). Survey study on the methods of bird vocalization classification. 1-8. 10.1109/ICCTAC.2016.7567337
11. Lopes, Marcelo & Gioppo, Lucas & Higushi, Thiago & Kaestner, Celso & Silla, Carlos & Koerich, Alessandro. (2011). Automatic Bird Species Identification for Large Number of Species. 117-122. 10.1109/ISM.2011.27.
12. ancovic, Peter & Kokuer, Munevver. (2011). Automatic detection and recognition of tonal bird sounds in noisy environments. EURASIP J Adv Sign Process 2011. 1-10