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:05/May-2024 Impact Factor- 7.868 www.irjmets.com

"BIRD SOUND PREDICTION SYSTEM"

Neha Bhalerao*1, Pallavi Kasture*2, Renuka Dhule*3, Saurabh Chaunde*4, Prof. Jaya, A. Zade*5

*1,2,3,4Department of Computer Engineering, Government College of Engineering, Dr. Babasaheb Ambedkar Technological University Lonere Yavatmal, Maharashtra, India.

*5Assistant Professor, Department of Computer Engineering, Government College of Engineering, Dr. Babasaheb Ambedkar Technological University Lonere Yavatmal, Maharashtra, India.

DOI : https://www.doi.org/10.56726/IRJMETS56666

ABSTRACT

In this paper, we describes how we created and tested a system to identify bird species based on their sounds. Using advanced machine learning techniques like Artificial neural networks (ANNs) and recurrent neural networks (RNNs), our system can analyze and classify bird calls from large collections of audio recordings. Key parts of our system include methods to clean up the audio and extract important features, and a new model that combines different aspects of the bird sounds for better accuracy. We built a large dataset, carefully designed our models, and adjusted various parameters to improve performance. We measured the system's success using accuracy, precision, recall, and F1-score, and found it to be significantly better than existing methods. We also explore how this system can be used in real-world situations, such as monitoring bird populations and helping with conservation efforts. The results show that our bird sound prediction system is a powerful tool for ornithologists and conservationists, making bird species identification faster, more accurate, and more efficient.

Keywords: Artificial Neural Network, multilayer perceptron, RecurrentNeural Network, MFCC, Spectrograms.

I. INTRODUCTION

Bird sound prediction is a fascinating and valuable field at the intersection of ornithology, technology, and environmental conservation. It entails using advanced audio processing techniques and machine learning algorithms to anticipate and identify the vocalizations of various bird species. This innovative approach to understanding and interacting with the avian world has gained significance due to its potential to address critical ecological and scientific challenges.[1]

The ability to predict and recognize bird sounds is not only a testament to the everevolving capabilities of technology but also a promising tool for ornithologists, conservationists, and bird enthusiasts. By deciphering the intricate language of birds, we can gain insights into their behaviors, migratory patterns, and breeding habits. This, in turn, contributes to a deeper comprehension of avian biodiversity and ecosystem dynamics, ultimately aiding in the conservation and protection of our feathered coinhabitants on this planet. In the vast tapestry of natural sounds, avian vocalizations form a rich and diverse melody, offering valuable insights into the world of birds. The predictive analysis of bird sounds stands as a pioneering endeavor at the intersection of ornithology and machine learning. This report delves into the innovative methodologies and technologies employed to decipher the intricate language of birdsong. By leveraging advanced algorithms and signal processing techniques, this study aims to not only identify and classify bird species through their unique vocalizations but also to contribute significantly to ecological research, conservation efforts, and our understanding of avian communication and behavior. [2]

Existing System

Vivekanandam et. al. have used Light weight CNN (LWCNN) 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.

Manoharan et. al. 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%.

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:05/May-2024 Impact Factor- 7.868 www.irjmets.com

Chandu B 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 44100Hz or 48000Hz in order to preserve diversity and avoid overfitting. The data for these examples comes from the Google Recording and LibriSpeech ASR datasets.[3]

Disadvantages Of Existing System

- The existing system is a time constraint; it takes a long time to analyze the sound of the birds.
- The existing system usually requires much more data. Lots of training data is required
- The existing system is also more computationally expensive.
- The existing system can take several weeks to train the dataset completely from scratch.
- The amount of computational power needed is also more.

Proposed System

The first step of implementation is gathering data from dataset which is obtained from Kaggle. The audio recordings of the birds in .wav format are included in this resource. This dataset contains audio recordings of the birds in .wav format. Kaggle are open websites dedicated for dataset where users upload their own recordings. 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.

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.

For the output, user can use GUI i.e., Graphical User Interface to analyse the species of the bird. With the help of GUI user can upload the dataset, process and show the outcome.[4]

Advantages Of Proposed System

- The proposed system requires only lesser time to analyze the sound of the birds.
- The proposed system can handle with less number of training data also.
- The proposed system is not more expensive when compared to the existing system models.
- The proposed system requires only less amount of computational power.
- The proposed system has good fault tolerance.
- The proposed is also has a good distributed memory.
- The proposed system may take only few hours train the dataset completely from scratch which is much lesser than the existing system model.

Software/ Hardware Requirement

Hardware Requirements

Disk: Capacity >500MB for smooth performance

RAM: 2GB (minimum), 4GB(recommended) and above.

Processor: Intel Core 2 Quad or Intel i3-5th Gen and above (recommended)

Software Requirements

OS: Windows 7 and above or linux

Front End Technology: HTML5 CSS3, JavaScript, Bootstrap 4

Backend Technology: PHP & MySql

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:05/May-2024 Impact Factor- 7.868 www.irjmets.com

Web Framework: Flask

Code Editor: Visual Studio Code, Notepad,Anaconda or Other Code Editor

Web Browser: Google Chrome, Mozilla Firefox, Opera or Other Web Browser.

II. PROPOSED METHODOLOGY

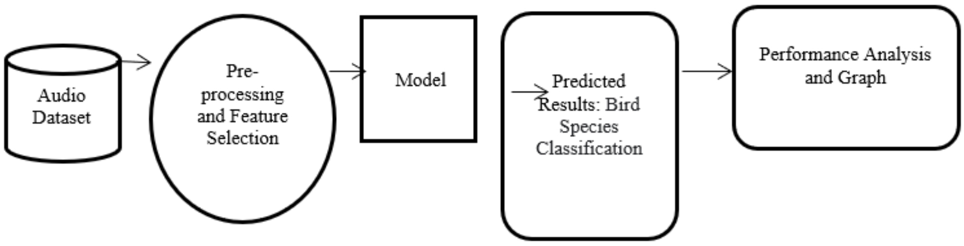

Figure -1: System Architecture

The system will require users to create an account, which will be subject to admin approval before login credentials are provided. Once approved, users can log in and perform file sharing actions with other users registered on the portal. The system will have a "Send File" page, where the user can upload files to share. Upon file upload, a unique key will be generated and stored in the database. To download a file, the receiver will need to request the key from the sender. The system will notify the sender of the key request, and the sender will have the option to approve or deny the request. If approved, the receiver can then download the file using the provided key.[5]

The proposed methodology for developing this system will involve the following steps:

- Requirement Analysis: Identifying the requirements and specifications of the system based on user needs and feedback.
- System Design: Developing the system architecture and designing the user interface for the application.
- Database Design: Designing the database schema to store user information and file data securely.
- Implementation: Developing the system using appropriate programming languages, frameworks, and tools.
- Testing: Conducting various tests to ensure that the system is functional, secure, and meets the required specifications.
- Deployment: Deploying the system on a suitable hosting platform for real-time usage.

Activity diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

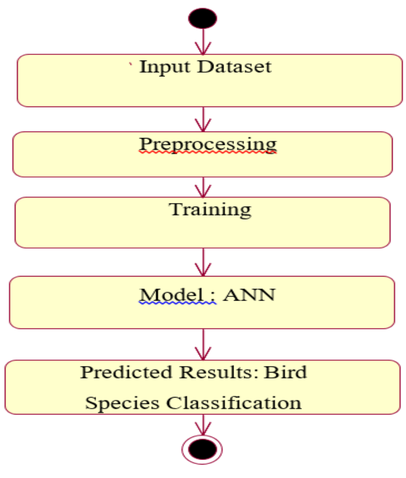

Figure 2: Activity Diagram for Bird Sound Prediction System

IRJMETS

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:05/May-2024 Impact Factor- 7.868 www.irjmets.com

III. IMPLEMENTION

Dataset:

In the first module, we developed the system to get the input dataset for the training and testing purpose.

The dataset consists of 5,422 Bird Sound Classification Using Deep Learning and can practice it on Kaggle itself.

Importing the necessary libraries:

The very important and great library that supports audio and music analysis is Librosa. Simply use the Pip command to install the library. It provides building blocks that are required to construct an information retrieval model from music. Another great library we will use is for deep learning modeling purposes is TensorFlow, and I hope everyone has already installed TensorFlow.

```
pip install pandas
pip install Pillow
pip install --upgrade pandas==1.3.4
pip install Keras==2.8.0
pip install tensorflow==2.8.0
pip install Flask
pip install librosa==0.9.1
```

Exploratory Data Analysis of Audio data

We have 5 different folders under the urban dataset folder. Before applying any preprocessing, we will try to understand how to load audio files and how to visualize them in form of the waveform. If you want to load the audio file and listen to it, then you can use the IPython library and directly give it an audio file path. We have taken the first audio file in the fold 1 folder that belongs to the dog bark category.

Now we will use Librosa to load audio data. So when we load any audio file with Librosa, it gives us 2 things. One is sample rate, and the other is a two-dimensional array. Let us load the above audio file with Librosa and plot the waveform using Librosa.

Sample rate – It represents how many samples are recorded per second. The default sampling rate with which librosa reads the file is 22050. The sample rate differs by the library you choose.

2-D Array – The first axis represents recorded samples of amplitude. And the second axis represents the number of channels. There are different types of channels – Monophonic(audio that has one channel) and stereo(audio that has two channels).

We load the data with librosa, then it normalizes the entire data and tries to give it in a single sample rate. The same we can achieve using scipy python library also. It will also give us two pieces of information – one is sample rate, and the other is data.[6]

When you print the sample rate using scipy-it is different than librosa. Now let us visualize the wave audio data. One important thing to understand between both is- when we print the data retrieved from librosa, it can be normalized, but when we try to read an audio file using scipy, it can't be normalized. Librosa is now getting popular for audio signal processing because of the following three reasons.

1. It tries to converge the signal into mono(one channel).
2. It can represent the audio signal between -1 to +1(in normalized form), so a regular pattern is observed.
3. It is also able to see the sample rate, and by default, it converts it to 22 kHz, while in the case of other libraries, we see it according to a different value.

Imbalance Dataset check

Now we know about the audio files and how to visualize them in audio format. Moving format to data exploration we will load the CSV data file provided for each audio file and check how many records we have for each class.

IRJMETS

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:05/May-2024 Impact Factor- 7.868 www.irjmets.com

The data we have is a filename and where it is present so let us explore 1st file, so it is present in fold 5 with category as a dog bark. Now use the value counts function to check records of each class. When you see the output so data is not imbalanced, and most of the classes have an approximately equal number of records.[7]

Data Preprocessing:

Some audios are getting recorded at a different rate-like 44KHz or 22KHz. Using librosa, it will be at 22KHz, and then, we can see the data in a normalized pattern. Now, our task is to extract some important information, and keep our data in the form of independent(Extracted features from the audio signal) and dependent features(class labels). We will use Mel Frequency Cepstral coefficients to extract independent features from audio signals.

MFCCs – The MFCC summarizes the frequency distribution across the window size. So, it is possible to analyze both the frequency and time characteristics of the sound. This audio representation will allow us to identify features for classification. So, it will try to convert audio into some kind of features based on time and frequency characteristics that will help us to do classification.[8]

To demonstrate how we apply MFCC in practice, first, we will apply it on a single audio file that we are already using.

Now, we have to extract features from all the audio files and prepare the dataframe. So, we will create a function that takes the filename(file path where it is present). It loads the file using librosa, where we get 2 information. First, we'll find MFCC for the audio data, And to find out scaled features, we'll find the mean of the transpose of an array. Now, to extract all the features for each audio file, we have to use a loop over each row in the dataframe. We also use the TQDM python library to track the progress. Inside the loop, we'll prepare a customized file path for each file and call the function to extract MFCC features and append features and corresponding labels in a newly formed dataframe.[9]

Splitting the dataset

Split the dataset into train and test. 80% train data and 20% test data.

Audio Classification Model Creation

We have extracted features from the audio sample and splitter in the train and test set. Now we will implement an ANN model using Keras sequential API. The number of classes is 10, which is our output shape(number of classes), and we will create ANN with 3 dense layers and architecture is explained below.

1. The first layer has 100 neurons. Input shape is 40 according to the number of features with activation function as Relu, and to avoid any overfitting, we'll use the Dropout layer at a rate of 0.5.
2. The second layer has 200 neurons with activation function as Relu and the drop out at a rate of 0.5.
3. The third layer again has 100 neurons with activation as Relu and the drop out at a rate of 0.5.

Compile the Model:

To compile the model we need to define loss function which is categorical cross-entropy, accuracy metrics which is accuracy score, and an optimizer which is Adam.

Train the Model:

We will train the model and save the model in HDF5 format. We will train a model for 250 epochs and batch size as 32. We'll use callback, which is a checkpoint to know how much time it took to train over data.

Check the Test Accuracy:

Now we will evaluate the model on test data. We got near about 97 percent accuracy on the training dataset and 100 percent on test data.

Deployement:

We used Flask, a simple web framework, to make our bird sound prediction system easy to use. Here's how it works: we set up a Flask server to host our machine learning model. Users can upload audio files through a web page. The server processes these files, cleans them up, and extracts important features. Then, it runs the processed audio through our trained neural network to identify the bird species. The results are shown on the web page. This setup makes the system easy to access and use, even for people without technical expertise, helping bird enthusiasts, researchers, and conservationists identify birds by their sounds.

IRJMETS
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:05/May-2024 Impact Factor- 7.868 www.irjmets.com

Real time Oputput:

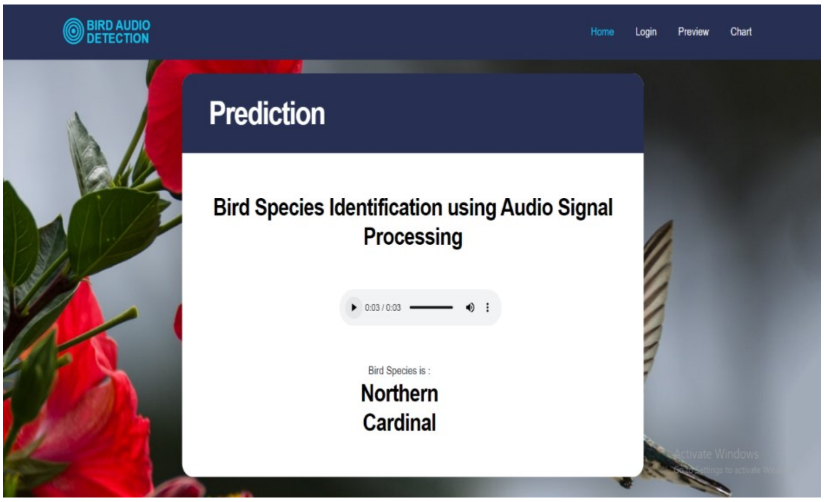

Figure3 : Real Time Output of the Bird Sound Prediction System

The real-time output of the bird sound prediction system's user interface shows a clear and interactive dashboard where users can immediately see which bird species have been identified. It includes an audio waveform visualization to display the sound's structure, and a results panel that shows the bird species name, a confidence score, and the time of the recording. Users can play back the bird call using audio controls, and real-time alerts inform them of specific bird detections or unusual sounds. The interface also features trend graphs to track bird species over time and a map showing where each bird was detected. Additionally, users can export data and provide feedback, making the system easy to use for bird monitoring and analysis.

IV. CONCLUSION

A model for automatic bird species identification using Artificial Neural Networks model (ANN model) is proposed in this paper. Due to effect of climate changes and count of endangered animal, many researchers proposed animal species recognition system to help them for specific study. In this project we have developed the system to identify bird sound identification using Artificial Neural Network (ANN). Each bird has a different tone of sounds.

ANN is applied to classify and recognise the bird sounds using Python. Firstly, all required data in term of power spectral density of bird is used in order to obtain data for each bird types. The next process is to train ANN to identify species of birds. Only one bird can identify in one time. Lastly, the graphical user interface (GUI) of bird sound identification have been developed that required the user to fed audio input of bird sound in order to recognise bird species. This project is done successfully and can be used to identify bird species.[10]

V. 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.[11]

VI. 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 Identificat ion using Audio. 2020 International Conference on Artificial Intelligence and Signal Processing (AISP).

IRJMETS
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:05/May-2024 Impact Factor- 7.868 www.irjmets.com

[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 and Development (SCOReD), pp. 342-345, 2020.

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

[6] 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.

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

[8] 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 of Signal Processing.

[9] Rai, P., Golchha, V., Srivastava, A., Vyas, G., & Mishra, S. (2016). An automatic classificat ion of bird species using audio feature ext ract ion and support vector machines. Internat ional Conference on Inventive Computation Technologies, 1-5.

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

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