IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

Machine Learning tool for Bird voice recognition

Sapna Jain[1], M Afshar Alam[2]

Assistant Professor, Department of Computer Science, Jamia Hamdard, New Delhi, India[1]

Professor, Department of Computer Science, Jamia Hamdard, New Delhi, India [2]

Abstract: Bird voice recognition can be used for a variety of purposes, such as bird conservation, research, and monitoring. For example,conservation organizations can use bird voice recognition to monitor populations of endangered or threatened bird species, researchers can use it to study bird behaviour and ecology, and birdwatchers can use it to identify birds in the field. The primary objective of bird voice recognition is to establish a dependable and effective approach to identifying bird species through their vocalizations. This functionality allows users to conveniently recognize birds in their surroundings without relying on visual cues.

.

Keywords: voice recognition, ecology, machine learning, convolutional neural networks.

I. INTRODUCTION

Birds are known for their beautiful and unique vocalizations, which they use for communication and various other purposes such as mating, territory defence, andwarning calls. Birds use a combination of different vocalizations including songs, calls, and other sounds to communicate with each other In current years, there was a growing interest in the use of voice popularity era to select out and feature a observe bird species. This era involves using software program algorithms and machine mastering techniques to analyse and classify chook vocalizations based on their acoustic residences. Voice recognition technology has numerous applications in the field of ornithology, including bird population monitoring, habitat assessment, and biodiversity conservation. Through the interpretation of bird sounds, scientists can gain valuable information about the behaviour, range, and population patterns of various bird species. Universal, the usage of voice popularity generation in bird research has the potential to greatly enhance our know-how of these charming creatures and their position within the atmosphere. Monitoring birds based on their sound is of utmost importance for various environmental and scientific objectives. Numerous crowd-sourcing and remote monitoring initiatives now capture and, to some extent, automatically

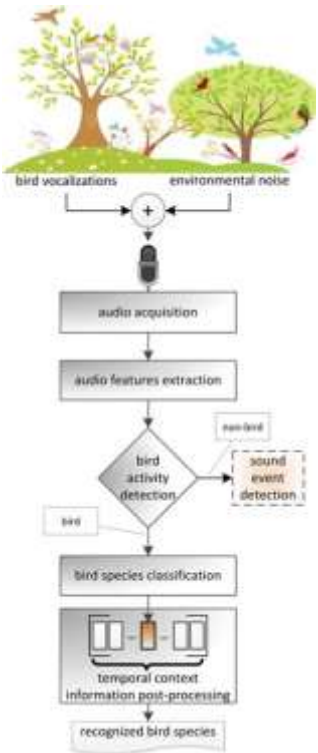

Fig 1. Classification Process

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

analyse these auditory recordings .The audio modality proves to be highly suitable for bird monitoring as many bird species are more easily detectable through sound compared to visual or other indicators. When addressing the challenge of automatic recognition of bird vocalizations, insights derived from research on speech recognition serve as a fundamental basis. Bird vocalization recognition and speech recognition share several common aspects, and both require the resolution of various fundamental problems. This project uses data set that contains bird songs. collected from Kaggle. The data set includes 114 different species of birds with 2161 different sounds in mp3 format .Utilizing Convolutional Neural Networks (CNNs) for bird voice recognition models has proven to be effective. CNNs are well-suited for tasks involving pattern recognition in visual and auditory data, making them a suitable choice for analysing spectrograms or other representations of bird vocalizations .By applying CNNs to bird voice recognition, the model can learn hierarchical features and capture intricate patterns in the spectrograms or audio representations. This enables the model to distinguish between different bird species based on their unique vocal characteristics. CNN architectures commonly consist of convolutional layers to extract local features, followed by pooling layers to down sample the representations and reduce computational complexity. Fully connected layers and softmax activation are typically employed for classification purposes.

Training a CNN for bird voice recognition involves providing labelled data, such as a dataset of bird vocalizations with corresponding species labels. The model learns to map the input spectrograms to their respective hen species during the education technique. It could then be used to are expecting the species of unseen bird vocalizations.
Bird voice recognition using Android Studio involves several steps:

1. Record bird vocalizations: You will need to record bird vocalizations using a microphone or smartphone. You can use an app like "Voice Recorder" or "EasyVoice Recorder" to record the bird songs and calls.
2. Process the recordings: The recordings need to be processed to extract relevantacoustic features that can be used for bird voice recognition. You can use software like Audacity or Raven Pro to process the recordings and generate spectrograms or other visual representations of the bird vocalizations.
3. Train the machine learning model: For training a bird vocalization recognition model, a machine learning algorithm like CNN can be utilized. The data needs to be labeled by identifying the bird species linked to each recording. This labeling process provides the necessary information for the model to learn and differentiate between different bird vocalizations. The CNN gains proficiency in identifying and categorizing bird species based on their vocalizations by being trained on this labelled data.
4. Build the Android app: Using Android Studio, you can build an app that can recognize bird vocalizations in real-time. The app can use the microphone to record bird vocalizations and process them using the machine learning model toidentify the bird species.
5. Test and improve the app: You will need to test the app with various bird vocalizations to ensure its accuracy. You can also continue to train the modelwith additional data to improve its accuracy over time.

Bird vocalizations are an essential means of communication for birds and are used forvarious purposes, such as mate attraction, territorial defence, and identifying membersof the same species. Studying bird vocalizations can provide insights into bird behaviour, ecology, and evolution. However, identifying bird species based on their vocalizations can be challenging, especially for non-experts. This creates a barrier to accurately monitoring bird populations and assessing the impact of environmental changes on bird behaviour. Therefore, the problem is to develop an accurate and efficient bird voice recognition system that can assist birdwatchers, researchers, and conservationists in identifying bird species based on their vocalizations.

Recognizing bird species can be difficult, especially for novice birdwatchers. Navigating through comprehensive field guides can be overwhelming. However, by solely recording bird vocalizations, species can be classified. This process involves capturing bird audio and inputting it into a system. Nonetheless, it demands extensive analysis and classification, requiring significant time and effort. As a result, conducting large-scale bird identification becomes nearly impractical due to the magnitude of the task.

II. LITERATURE REVIEW

Juha Niemi uses both feature extraction and signal classification to identify a picture. They conducted an experimental investigation on datasets with various picture types. However, the background species were not taken into account in their studies. Larger amounts of training data—which might not be available—are needed to identify the background species. Juha T. Tanttu et al. (2018) delivered the CNN algorithm and deep residual neural networks, two deep studying strategies for image type, and recommended a convolutional neural community skilled with John Martinsson et al. (2017).Additionally, it suggested a data augmentation technique that rotates and converts photos to the appropriate colour.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

The final identification is determined by combining the radar's data with the forecasts of the image classifier. An efficient automatic bird species identification method based on picture feature analysis was proposed by Li Jian, Zhang Lei, et al. (2014). used the similarity comparison technique and the usual picture database.

A software program was created by Madhuri A. Tayal, Atharva Magrulkar, et al. (2018) to make the procedure of identifying birds easier. This bird identification program uses a picture as input and outputs the birds identify. For the identification procedure, MATLAB and transfer learning are the technologies utilized.

In an effort to remove background elements and define potential areas where the bird might be present in the image, E. Andreia Marini, Jacques Facon, et al. (2013) presented a novel method based on color features extracted from unconstrained images using a color segmentation algorithm. The histograms' intervals were reduced to a predetermined number of bins via aggregation processing. The CUB-200 dataset was used in this paper's experimentation, and the findings indicate that this method is more accurate.

III. PROPOSED MODEL

Bird voice recognition using Android Studio involves several steps:
1.Record bird vocalizations: You will need to record bird vocalizations using a microphone or smartphone. You can use an app like "Voice Recorder" or "EasyVoice Recorder" to record the bird songs and calls.

2.Process the recordings: The recordings need to be processed to extract relevantacoustic features that can be used for bird voice recognition. You can use software like Audacity or Raven Pro to process the recordings and generate spectrograms or other visual representations of the bird vocalizations.

3.Train the machine learning model: For training a bird vocalization recognition model, a machine learning algorithm like CNN can be utilized. The data needs to be labeled by identifying the bird species linked to each recording. This labeling process provides the necessary information for the model to learn and differentiate between different bird vocalizations. With the aid of training the CNN on this categorised facts, it will become gifted in recognizing and classifying chicken species primarily based on their vocalizations.

4.Build the Android app: Using Android Studio, you can build an app that can recognize bird vocalizations in real-time. The app can use the microphone to record bird vocalizations and process them using the machine learning model to identify the bird species.

5.Test and improve the app: You will need to test the app with various bird vocalizations to ensure its accuracy. You can also continue to train the modelwith additional data to improve its accuracy over time.

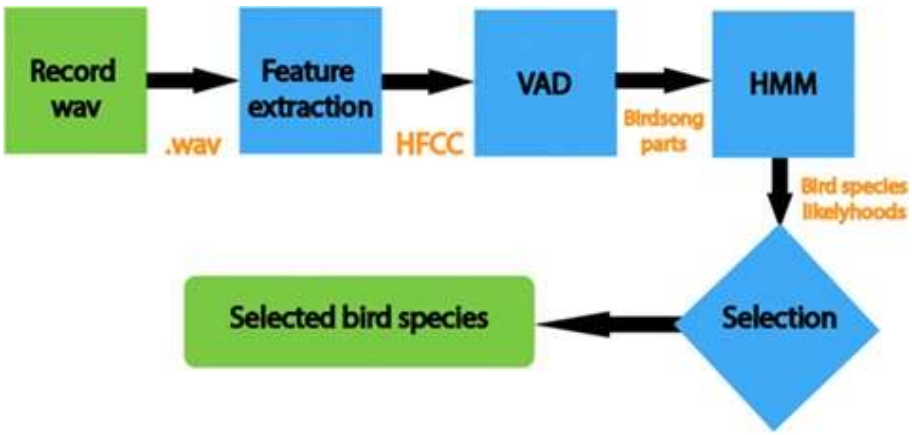

Fig 2. Process of feature extraction

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

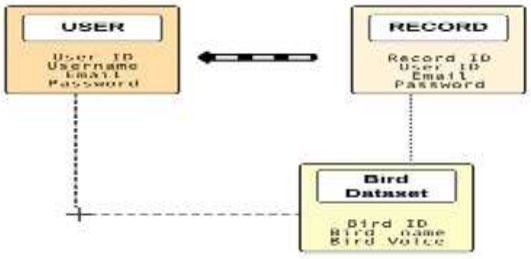

Fig 3 : Steps for using dataset

The dataset containing 114 exceptional species of birds with 2161 specific sounds in mp3 layout is a treasured aid for researchers and researchers interested in studying the vocalizations of birds. The dataset offers a wide range of sounds from a diverse set of bird species, making it useful for a variety of applications, including birdwatching, conservation, and research in fields such as bioacoustics and animal behaviour.

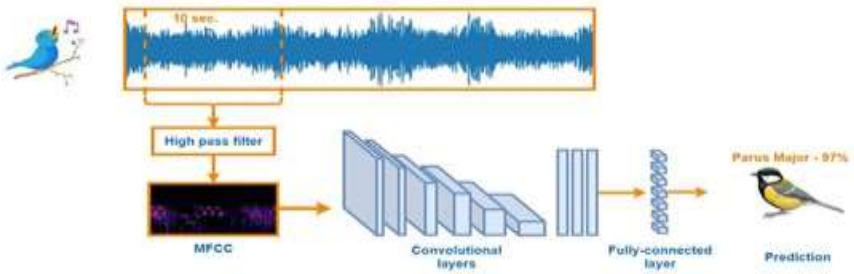

Fig 4 : Filtration Process

One potential use of the dataset is for bird identification and monitoring. By analyzing the acoustic characteristics of different bird vocalizations, researchers and birdwatchers can identify the species of birds present in a given area and monitor their populations over time. These statistics can be used to tell conservation efforts and better apprehend the position of birds in one-of-a-kind ecosystems. Libraries are essential for building a Python model as they provide pre-written code for specific functions, saving time and effort when developing programs. There are numerous libraries available for Python, each with its own set of functionalities. Some of the most commonly used libraries in data analysis and machine learning include NumPy, Pandas, Matplotlib, and Scikit-learn.CSV files are a common format for storing and exchanging data in a table. CSV stands for Comma-Separated Values, and is a simple way to store data in a text file where each line represents a row and each column is separated by a comma.

YAMNet is a pre-skilled neural community that makes use of the MobileNetV1 depthwise-separable convolution structure. This network is able to taking an audio waveform as input and making character predictions for 521 audio events derived from the AudioSet dataset. To process the audio signal, the model internally extracts "frames" and operates on batches of these frames. This approach enables efficient analysis of audio data in a structured manner, allowing the model to make predictions for various audio events.

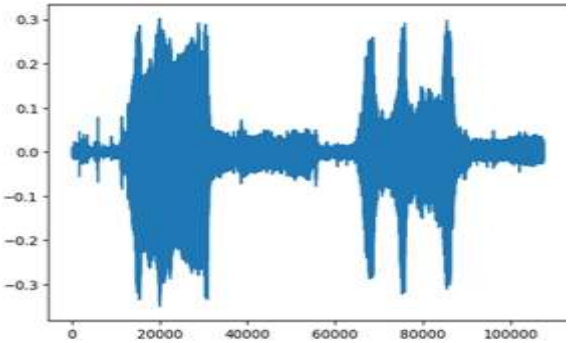

Fig 5 : using the audio dataset

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

Steps:

Step 1: To load a pre-trained YAMNet convolutional neural network and examine its layers and classes, you can use the "yamnet" function provided in the Deep Learning Toolbox. This function loads the pre-trained YAMNet network and returns an output object of type "SeriesNetwork."

Step 2: To train a TensorFlow Lite model using a custom dataset, you can utilize the Model Maker library, which simplifies the process through transfer learning. This approach allows you to retrain a TensorFlow Lite model with your own dataset, reducing the need for extensive training data

Step 3: Explore the data.The audio files have already been pre-processed and organized into separate train and test folders. Within each split folder, there is a dedicated folder for each bird species, identified by their corresponding bird_ code. All the audio files are in mono format and have a sample rate of 16kHz.For instance, the bird_ code 'houspa' is associated with the bird species House Sparrow. This organization of the audio data allows for easy access and management during the training and evaluation process.

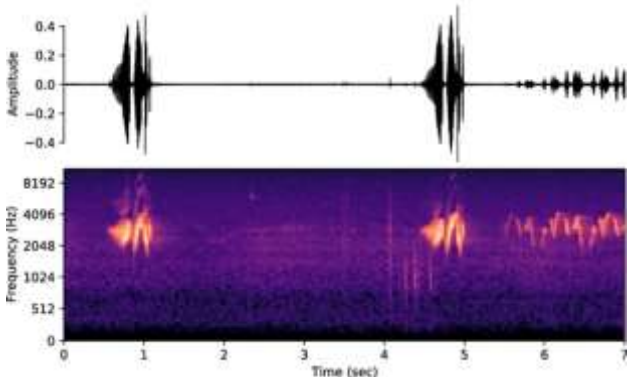

Fig 6: Training the model

Step 4: Training the Model YAM Net is a trained audio event classifier specifically designed to predict audio events based on the Audio Set ontology. The model has been trained using the Audio Set data set. When using the YAM Net in this model, the expected input format for audio data is as follows:

- Sampled at a frequency of 16kHz
- Containing a single audio channel (mono)

By adhering to these specifications, the input audio data can be accurately processed and classified by the YAMNet model.

The Model Maker library, designed for training TensorFlow Lite models, automatically handles the resampling of audio data, relieving the user from manually handling this task.

In the context of training audio models using Model Maker, certain parameters are relevant:

- The frame_ length parameter determines the duration of each training sample. In this case, it is set to EXPECTED_WAVEFORM_LENGTH * 3s.
- The frame_ steps parameter determines the spacing between consecutive training samples. For instance, the ith sample will start at EXPECTED_WAVEFORM_LENGTH * 6s after the (i-1)th sample.

These values address specific limitations found in real-world datasets. By adjusting the frame length and steps, it becomes possible to effectively handle audio data with varying characteristics and event durations.It's worth noting that the specific values for EXPECTED_WAVEFORM_LENGTH, frame_length, and frame_steps may differ based on the YAMNet model implementation and requirements you are working with.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

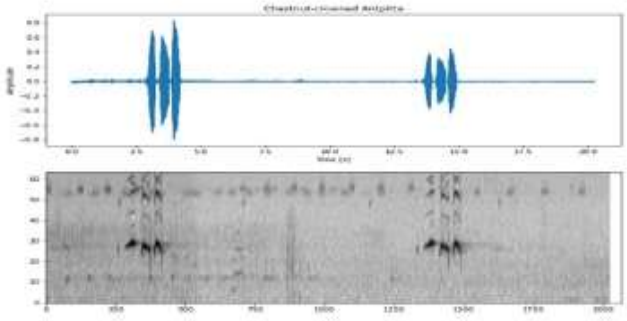

Fig 7: the resultant

Step 5: Trying out the model
try the model on a sample audio from the check dataset just to see the effects.

IV CASE STUDY

Bird recognition is an important skill for birdwatchers and conservationists alike, as it allows them to accurately identify and monitor bird populations in different habitats. However, recognizing bird species based on their vocalizations can be a challenging task, even for experienced birdwatchers. In order to better understand how people recognize bird species based on their vocalizations, a study was conducted at Okhla Bird Sanctuary, a protected area located in the National Capital Territory of Delhi, India.

Methodology:

1. Survey Design: To collect information on participant demographics, experience with bird voice identification, the impact of bird voices on mental health, and views regarding the incorporation of technology into bird voice, a structured questionnaire was created.

2. Data Collection: The survey was conducted on-site at the Okhla Bird Sanctuary. The visitors and local communities were approached, explaining the purpose of the survey and encouraging participation. The questionnaire was administered through face-to-face interviews, ensuring clarity of questions and allowing for additional explanations when needed.

3. Sample Size: A total of 15 participants voluntarily participated in the survey.

4. Data Analysis: The collected data was analyzed using various techniques, including descriptive statistics and frequency distributions. Descriptive statistics were utilized to summarize the responses and provide an overview of the data. This involved calculating measures such as means, medians, standard deviations, and percentages to gain insights into the central tendencies and variations within the data. Additionally, frequency distributions were employed to examine the occurrence and distribution of different variables or categories within the dataset. This allowed for a clearer understanding of the distribution patterns and the relative frequencies of specific responses or themes.

Survey Findings:

1) How confident are you in your ability to recognize bird species based on their vocalizations?
In this study, the recruited participants from a variety of backgrounds, including birdwatchers, students, and nature enthusiasts, to listen to recordings of bird songs and calls from different species at Okhla Bird Sanctuary. Participants were asked to identify the species of bird based on their vocalizations, and their responses were recorded for analysis. I also collected additional information from participants, such as their level of experience with birdwatching and their familiarity with the bird species in the recordings as depicted in F.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

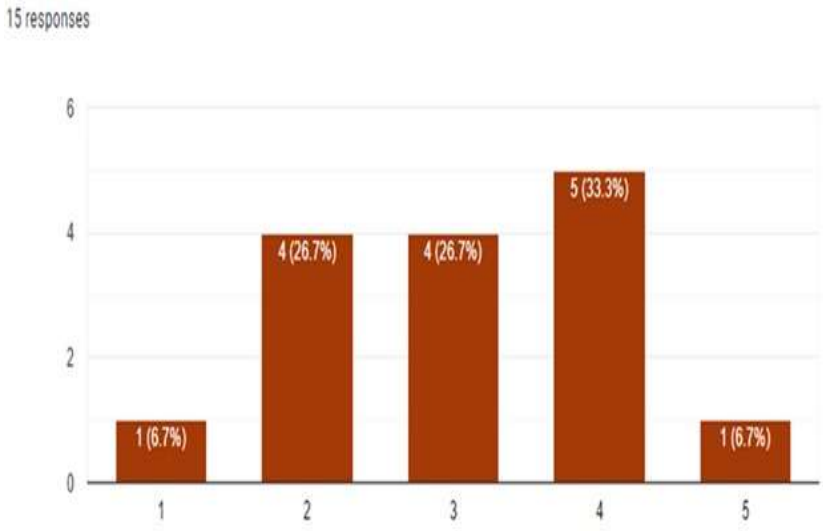

Fig 8 : Ability to recognise bird species

2) Have you ever tried to identify a bird solely based on its vocalization survey as depicted in Fig 9 .?

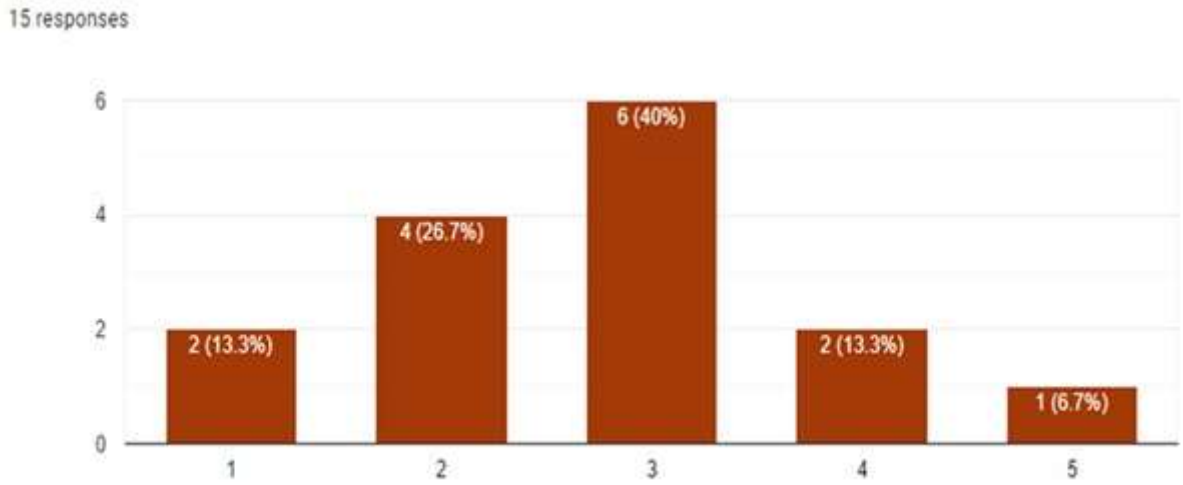

Fig 9 : Bird identification based on vocalization

3) How often do you listen to bird sounds or calls in natural environments Fig 10?

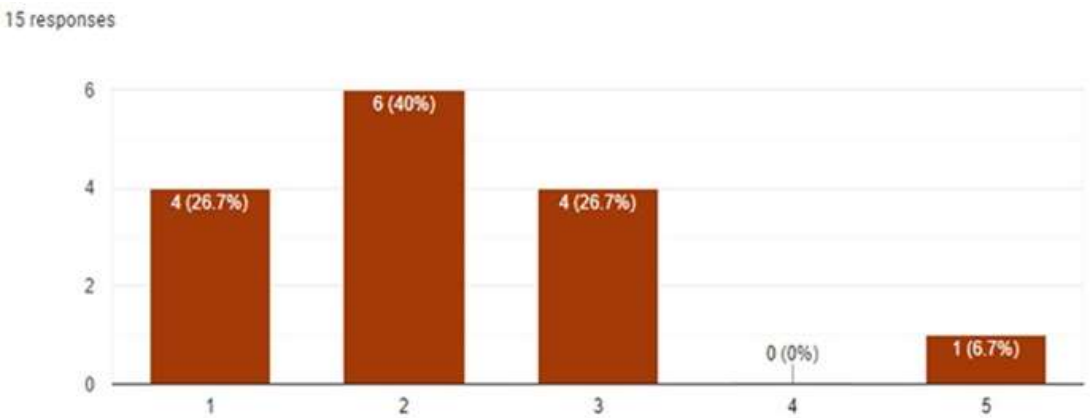

Fig 10 : Listrning birds sounds

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

4)How does listening to bird voices affect your mental health?

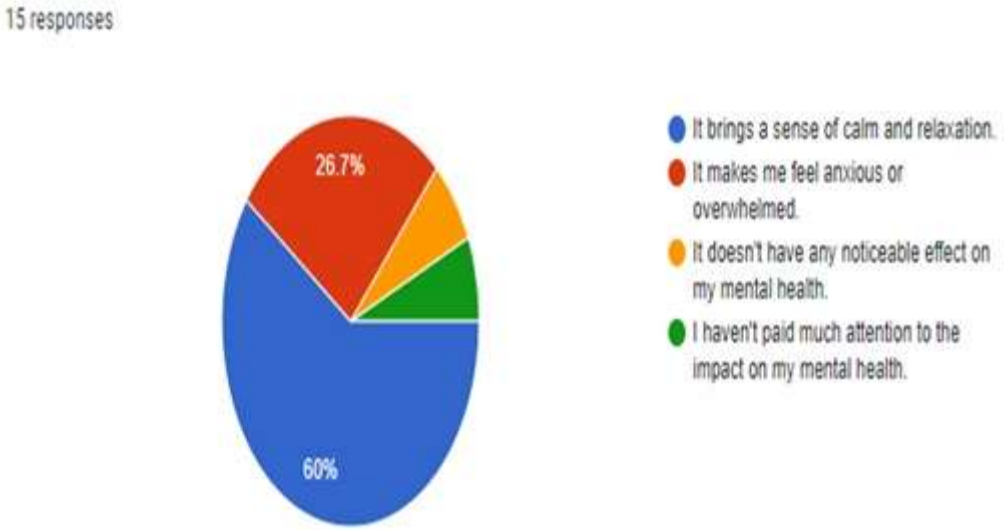

Fig 11 : Effect on mental health

Fig 11 depicts the calming and soothing effects of bird songs have been attributed to their ability to reduce stress and anxiety, promote relaxation, and improve mood. However, it is worth noting that some participants in our survey reported feeling anxious or overwhelmed by the sounds of birds. This may be due to individual differences in sensitivity to sound, as well as the specific characteristics of the bird vocalizations that were played during the survey.

5)How important do you consider bird vocalizations as a means of identifying bird species?

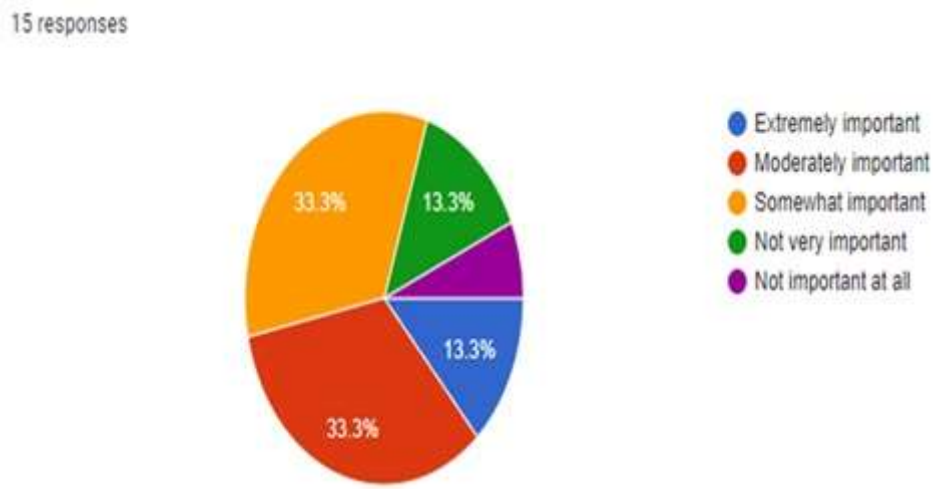

Fig 12 : Importance of Bird volcalisation

Fig 12 shows the survey results which provides the valuable insights into the challenges and opportunities of bird recognition based on vocalizations, and highlights the importance of continuing to develop new tools and techniques for birdwatchers and conservationists. By understanding how people recognize bird species based on their vocalizations, we can better monitor and protect bird populations in different habitats, and contribute to our understanding of the complex relationships between birds and their environment.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

Application Screens made for the implementation of the above case study are as follows:-

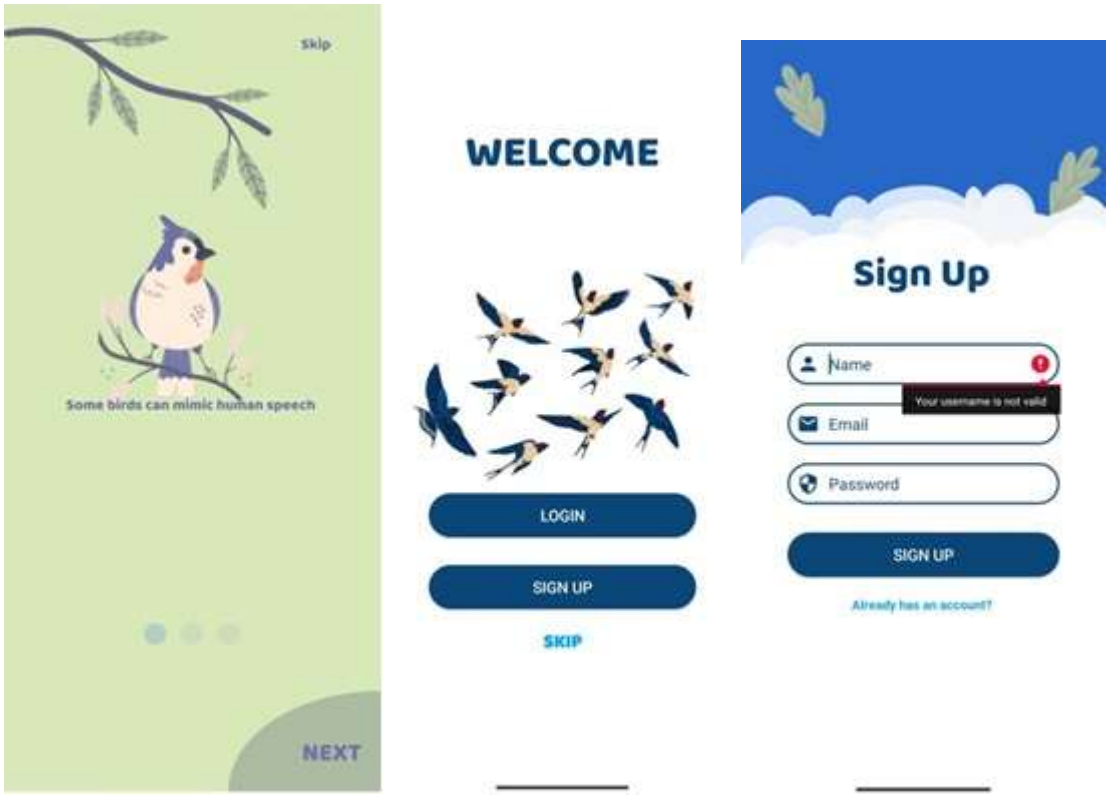

Fig 13: Main screen,, Welcome screen and Signup screen

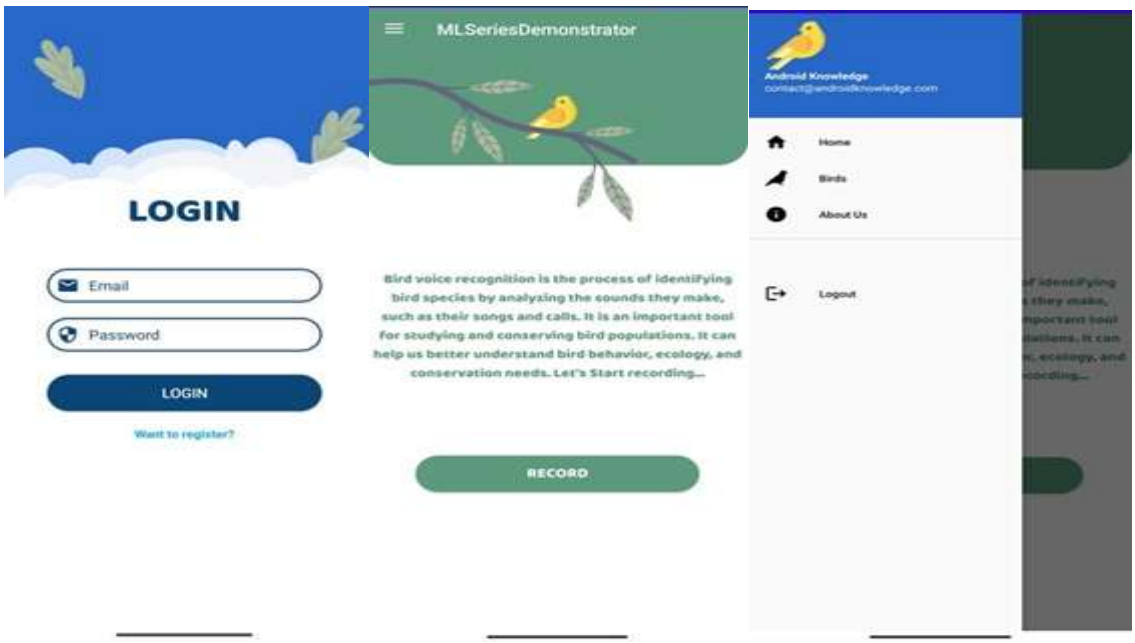

Fig 14 Login, recording screen

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12134

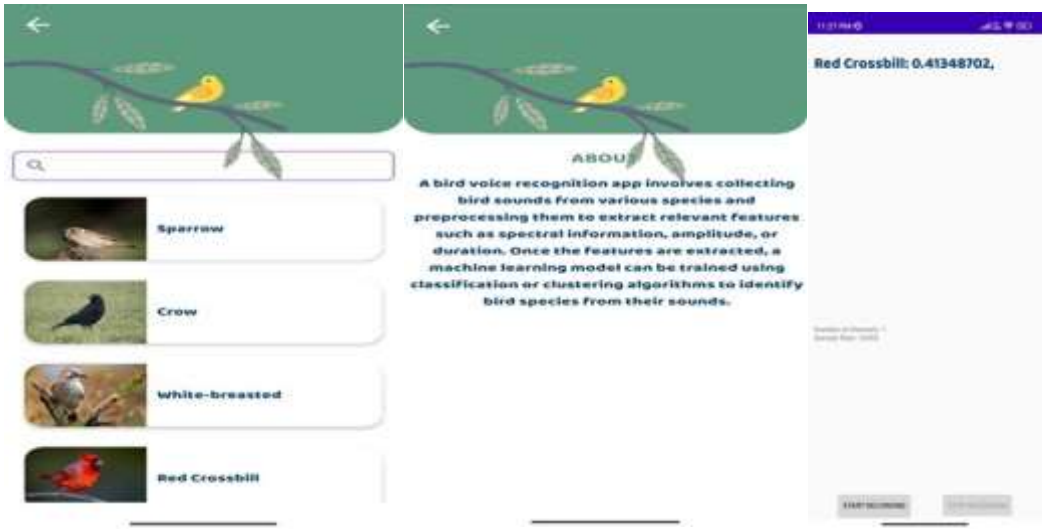

Fig 15 Bird voice recognition steps screen

V CONCLUSION

Bird voice popularity generation has the potential to revolutionize the manner we study and conserve hen populations. By way of presenting a greater accurate and green approach for figuring out and reading bird vocalizations, we are able to advantage valuable insights into bird behaviour, ecology, and evolution. This technology can alsoaid in monitoring bird populations, tracking migration patterns, and studying the effects of environmental changes on bird behaviour. Although there are still some limitations and challenges to overcome, such as the variability of bird vocalizations and the need for more robust and adaptable machine learning algorithms, ongoing research and development in this field show promise for the future of bird voice recognition. With continued innovation and collaboration among researchers, conservationists, and technology experts, we can unlock the full potential of bird voice recognition and contribute to the conservation of these magnificent creatures.

ACKNOWLEDGMENT

The authors conducted the experiments in the FIST and PURSE lab of Department of Computer Science, Jamia Hamdard and thank for the support rendered in the implementation of the proposed method in the paper.

REFERENCES

[1]. Fagerlund, Seppo. "Bird species recognition using support vector machines." EURASIP Journal on Advances in Signal Processing 2007, no. 1 (2007): 038637
[2]. Marini, AndrÃ©ia, Jacques Facon, and Alessandro L. Koerich. "Bird species classification based on color features." In 2013 IEEE International Conference on Systems, Man, and Cybernetics, pp. 4336- 4341. IEEE, 2013.
[3]. Barar, Andrei Petru, Victor Neagoe and Nicu Sebe. "Image Recognition with Deep Learning Techniques." Recent Advances in Image, Audio and Signal Processing: Budapest, Hungary, December 10-2 (2013).
[4]. Qiao, Baowen, Zuofeng Zhou, Hongtao Yang, and Jianzhong Cao. "Bird species recognition based on SVM classifier and decision tree." First International Conference on Electronics Instrumentation & Information Systems (EIIS), pp. 1-4, 2017.
[5]. Branson, Steve, Grant Van Horn, Serge Belongie, and Pietro Perona. "Bird species categorization using pose normalized deep convolutional nets." arXiv preprint arXiv: 1406.2952 (2014).
[6]. Madhuri A. Tayal, Atharva Mangrulkar, Purvashree Waldey and Chitra Dangra. Bird Identification by Image Recognition. Helix Vol. 8(6): 4349- 4352
[7]. Atanbori, John, Wenting Duan, John Murray, Kofi Appiah, and Patrick Dickinson. "Automatic classification of flying bird species using computer vision techniques." Pattern Recognition Letters (2016): 53-62.
[8]. https://asmpeurasipjournals.springeropen.com/articles/10.1186/s13636-018-0143-7
[9]. https://www.ijraset.com/research-paper/bird-species-identifier-using-convolutional-neural-network#introduction
[10]. https://www.tensorflow.org/lite/models/modify/model_maker/audio_classification
[11]. https://www.kaggle.com/code/jamilahfoucher/classifying-sounds