International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

BeakBook: Bird Species Recognition Mobile Applications

Chiranjivi Hole, Sakshi Kakade, Priti Mhatre

Department of Data Science, Usha Mittal Institute of Technology (Santacruz) Shreemati Nathibai Damodar Thackersey Women's University
DOI: https://doi.org/10.55248/gengpi.5.0424.1038

ABSTRACT

Birds, alongside humans, inhabit our world with grace and significance, serving as vital indicators of environ- mental changes. Given the imminent threat of extinction that many bird species face, it is critical to implement effective conservation techniques. With the use of cutting-edge deep learning algorithms, visual categorization based on bird photos becomes a viable method for species identification. The goal of this research is to create an effective model for rapid identification of bird species by utilising deep learning and neural networks. The neural network receives intense training to get excellent classification accuracy by using large datasets that include photos and sounds of birds. For bird voice recognition libraries like librosa are used for claassification

Advanced methodologies of Neural networks are used to extract complex information from bird photos, improving the model's capacity to identify minute species distinctions. This project intends to make a substantial contribution to scientific study and conservation efforts by automating the process of bird species identification. This will enable a deeper knowledge of avian biodiversity and assist in the preservation of our natural environment. In addition to speeding up the identification procedure, this paradigm shift towards automated bird species recognition promotes a greater comprehension of avian biodiver- sity. Effective deep learning models can help focus conservation efforts more precisely and effectively, reducing the threats that endangered bird populations must contend with. Through the democratisation of access to cutting-edge technology, this project enables people all around the world to take part in scientific research and bird conservation projects. The success of these initiatives marks the beginning of a new chapter in animal conservation, one in which technology acts as a catalyst for favourable alterations in the environment. We work together and with creativity to create a future where all bird species coexist peacefully with their natural environments.

***Index Terms*—Neural Networks, Deep Learning, EfficientNet, YamNet, ,Bird species recognition, Birds voice recognition.**

I. Introduction

Many people find satisfaction in being able to recognize and appreciate the beauty of our feathered friends in a world full of fascinating and unique avian species. Our Bird Species Recognition Application is made to get you closer to the fascinating world of birds, whether you are an ornithology enthusiast, a nature lover, or simply inquisitive about the rich tapestry of birdlife that surrounds us.We have utilized the potential of Convolutional Neural Networks (CNNs) by utilizing the effectiveness of the EfficientNet B0 architecture and the power of cutting-edge technology. We can provide you with a seamless and precise tool for recognizing bird species from photographs thanks to this novel approach. For both professionals and bird enthusiasts, our application offers a convenient and accessible platform.

1) Birdwatching is a popular hobby promoting appreciation for nature and birds because there are so many different kinds of birds and their minute visual distinctions, it can be difficult to identify them and for birdwatchers, carrying field guides or reference books is frequently impracticable.
2) Identifying and learning about bird species is a global interest for enthusiasts, scientists, and nature lovers. .
3) Recognizing different bird species can be challenging due to visual variations and the vast number of bird species.
4) The aim is to improve bird species recognition using Neural Network and deep learning.

II. Problem Defination:

Birdwatching is a well-liked hobby that builds a profound respect for the avian world and helps people connect with nature. Whether it's the dazzling plumage of a tropical bird, the amazing flight of a raptor, or the lovely singing of a songbird, these species captivate our hearts and inspire wonder. The identification and study of these species will benefit scientific knowledge and human delight, and bird enthusiasts, scientists, and nature lovers worldwide will be interested.

But distinguishing between various bird species can some- times prove to be a challenging and time-consuming endeav- our, even for seasoned birdwatchers. It's not always practicable to bring field guides or reference books since many bird species have subtle visual changes.

III. Survey Of Literature

In the process of developing this project, an extensive review of various papers and journals were conducted

In [1] research paper four different transfer learning models namely InceptionV3, ResNet152V2, DenseNet201, and MobileNetV2 was implemented on the identified dataset. Out of this ResNet152V2 provide with maximum accuracy 95.45. Thus in these research paper a model was built to identify bird species using Convolutional Neural Network In[2]research paper the project builds an Application for Bird Species detection on varied dataset using Convolutional Neural Network. The accuracy of the built application was 75 In[3]the bird's dataset was built based on Asian bird species. Two different models were proposed which showed that the proposed pretrained ResNet model achieved greater accuracy in comparison to the based model.

Their final model shows 97.98 accuracy in identifying the bird species [4] Reserach paper served as our primary reference throughout the project. In[4] a CNN based architecture was developed and evaluated for 525 different bird species categorization. By combining strong data augmentation tactics with transfer learning approaches, this method produced impressive accuracy rates.

When it came to classifying bird species, the model performed well and showed characteristics that could be applied to other contexts without experiencing overfitting. Interestingly, the test set showed a similar accuracy of 86.7%, but the validation set indicated an accuracy of 87%.

The extension of current databases has been identified as one path for future refinement, with a particular focus on addressing imbalances among underrepresented species and gender groups. To improve model performance, more research into different approaches to data augmentation is necessary. The work In[5]investigates the use of the Na¨ıve Bayes algorithm to the recognition of bird species using acoustic characteristics, with a 91.58% accuracy rate.

The research highlights how different the vocal tracts of different birds produce sounds, as well as the difficulties that come with memory control and signal-to-noise ratio optimisation.

An iterative procedure for non-real-time bird speech recognition is outlined, including steps like choosing an audio clip and classifying the data using the Na¨ıve Bayes method. In[6]the paper details an experiment that used Human Factor Cepstral Coefficients (HFCCs) and Hidden Markov Models (HMMs) to automatically classify the vocalisations of eighteen different bird species.

An interspecific success rate of 81.2% and a classification success rate of 90.45% for families were attained in the experiment; data from other models might be taken into account for possible improvement.

The quality and processing of the input data are critical to the experiment's success; possible enhancements include more precise recording segmentation and the use of deep neural networks for the identification of bird sounds.

TABLE I

Features extracted for Bird recognition

NO	Species	Color	Shape Features
1	Fairy Tern	White, Black, Grey	Small, Slim body, Long wings
2	Zebra Dove	Tan,White,Black	Small, Plump body, Short beak
3	Harpy Eagle	Dark brown, Gray	Large, Robust body, Huge wings.
4	Canary	Yellow,Gray	Small, Slender body, Short beak

IV. Established Operational Infrastructure.

1. Merlin Bird ID:

The Merlin Bird ID from the Cornell Lab of Ornithology is renowned for its extensive database and user-friendly interface. Using machine learning algorithms to identify bird species based on user-submitted images or descriptions, it assists birdwatchers in rapidly and correctly identifying different bird species.

Key Features: Offers a large database of bird species, simple photo-based identification, the capacity to identify bird cries, and comprehensive species information, making it a priceless tool for ornithologists and enthusiasts alike.

2. eBird:

Recording bird sightings as part of a citizen science pro- gramme is the goal of eBird, a robust platform. Users can contribute photos, audio recordings, or descriptions of their observations of birds to a worldwide dataset used for research on bird populations and conservation efforts.

Key Features: Provides tools for species identification, allows users to submit photos and audio recordings, encourages community participation in birdwatching and conservation, and provides insights on bird populations.

3. iBird:

For bird species, iBird is a comprehensive digital field guide. It has a vast collection of bird information, including detailed descriptions of characteristics, habitats, and behaviours along with images and sounds.

V. Algorithmic framework design:

A. Birds Image Recognition

EfficientNet, a convolutional neural network architecture has emerged as a prominent framework for image classifi- cation tasks owing to its balance between performance and computational efficiency. The EfficientNet-B0 variant, chosen as the foundation for our bird species classification endeavor, is characterized by a compound scaling method that simulta- neously adjusts network depth, width, and resolution. Let X represent the input image with dimensions $W \times H \times C$, where W is width, H is height, and C denotes the number of channels. The architecture comprises the following key elements: **Input Stem:** Initial processing layers, including convolutional and pooling operations, extract fundamental features from the input image.

Efficient Blocks: These blocks, iteratively applied throughout the network, feature convolutional layers with varying kernel sizes and expansion ratios. Each block's output undergoes recalibration via a squeeze-and-excitation (SE) mechanism to enhance feature responsiveness.

Global Average Pooling: Following the last block, global average pooling aggregates feature maps into a fixed-size representation.

Fully Connected Layer: A fully connected layer, followed by a softmax activation function, facilitates classification.

B. Birds Voice Recognition

YAMNet is a convolutional neural network (CNN) that pro- cesses audio spectrograms. It uses a spectrogram as input and uses convolutional filters to extract features from the signal. The convolutional layers extract abstract representations of audio features, which are then downsampled to lower their dimensionality while keeping crucial data intact. This hierar- chical processing allows YAMNet to learn hierarchical features of audio events. The network's top layers are connected and followed by softmax activation for classification. The output classes, representing various audio events or categories, are mapped to the learned features. The softmax function normal- izes the output scores across all classes to create a probability distribution. YAMNet uses gradient descent optimisation and backpropagation to modify the weights of its layers, aiming to minimize the discrepancy between input audio sample true labels and predicted probabilities. This training process on the AudioSet dataset makes YAMNet proficient in classifying various audio events, making it useful for various audio classification tasks.

VI. Experiments

A. Methodology for Bird Image Recognition Model

1) Data Source:

: The subsequent stages for project involes downloading and organizing two datasets, ”525-western-bird-species” and ”25- indian-bird-species,” sourced from Kaggle. The script unzips the datasets and organizes them into separate directories, sim- plifying data management and setting the stage for data pro- cessing. The script also merges datasets, transferring images from the ”indian-bird-species” dataset into the ”western-bird- species” dataset, expanding the scope of species the model can classify, enhancing the project's classification capabilities.

The preview of dataset used for project:

Fig. 1. A preview of birds in dataset

2) Data Preprocessing

: A. Data Augmentation:

Leveraging TensorFlow’s ’ImageDataGenerator,’ we define a set of strategies to augment our training data. This augmen- tation process generates an array of diverse variations of our training images, which is instrumental in improving our model’s capacity to generalize across a range of data scenarios. This approach enhances the model’s robustness and its ability to adapt to variations in real-world data.

B. Creating Data Generators:

To facilitate the data flow into our model during training, we thoughtfully create two pivotal data generators using TensorFlow’s ’ImageDataGenerator.’ These data generators are meticulously designed to load and preprocess images from our training and validation directories. This crucial step streamlines the data supply process, ensuring that the model is fed with appropriately processed data during the training phase..

3) *Training::* The training phase involves constructing a model using the ’fit’ method, which provides training and validation data through data generators. The model’s progress is tracked in the ’history’ variable. Class names are saved in a JSON file for interpreting predictions. The saved model configuration is then loaded from the checkpoint file, encapsulating the training efforts and ready for predictions on new data. Class names are then retrieved from the JSON file to provide human-readable labels to the model’s outputs. This process ensures meaningful associations between numerical predictions and the actual bird species they represent.

4) *Results::* The process involves loading, preprocessing, and passing new images through a model for classification. If the model’s prediction confidence exceeds a predefined threshold, the class is printed, and if it falls below, the model is explicitly indicated, ensuring transparency and precision in the results.

Fig. 2. Predictions for Bird Recognition through Image

B. Methodolgy For Bird Voice Recognition :

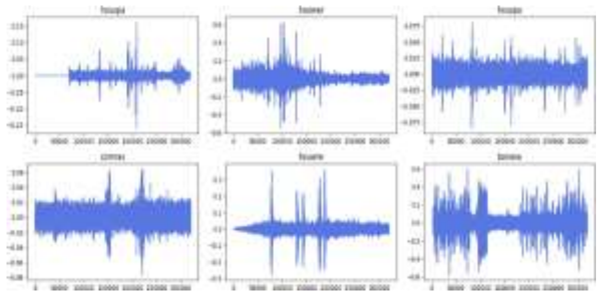
Fig. 3. Visualizing bird audio data

1) Data Source:

: Explore the data, noting that the audio files are divided into train and test folders, with each split containing folders named after bird codes. All audios are mono and have a sample rate of 16kHz

2) Data Preprocessing:

: Import Model Maker, TensorFlow, and any additional li- braries required for manipulating audio, and creating visuali- sations.

3) Training :

: To train the model using Model Maker for audio, begin with a model spec, which serves as the base model for extracting information and learning about the new classes. It also de- termines how the dataset will be transformed to comply with the model’s specifications, such as sample rate and number of channels. YAMNet, an audio event classifier trained on the AudioSet dataset, is used for predicting audio events from the AudioSet ontology. The create method in the audio_classifier module builds and starts the model’s training. With a batch size of 128 and 100 epochs, training is conducted.The saved model configuration is then loaded from the checkpoint file, encapsulating the training efforts and ready for predictions on new data. Class names are

then retrieved from the JSON file to provide human-readable labels to the model's outputs. This process ensures meaningful associations between numerical predictions and the actual bird species they represent.

C. Methodolgy for Application Development

1) *Deep Learning Model Conversion and Data Preparation:* Converted the models (Images and Sound) to Tensor- Flow Lite format (tflite) for efficient execution on resource- constrained devices. Created a JSON file to store the names of all bird classes used for classification.

2) *Android App Development Using Android Studio:* Choosed Android Studio for the project, providing tools and resources for building Android applications efficiently. Created a new Android project in Android Studio, configuring the target Android version, defining the project's structure, and setting up essential dependencies.

3) *Programming Language and Framework:* Java was se- lected as the primary programming language for Android app development. TensorFlow Lite, a framework developed by Google, was integrated into the Android app.

4) *Integration of TFLite Models and JSON Files::* The TensorFlow Lite models (.tflite) was integrated into the Android project. The JSON files containing birds labels from image and sound recognition models were integrated into the Android project.

5) *App Testing and Debugging:* Extensive testing on vari- ous Android devices and emulators was conducted to ensure functionality and performance. Android Studio's built-in de- bugging tools were utilized to identify and fix issues.

6) *User Interface Design and User Experience:* XML layouts were used to design an intuitive and user-friendly interface. User experience (UX) principles were taken into account during UI design.

7) *Building the Android App::* The Android Studio build system compiled the source code, integrated resources, and bundled them into an Android application package (APK) file. The final APK file was generated, containing the app, the integrated TensorFlow Lite model, and the bird class names JSON file.

8) *Deployment and Distribution:* The APK file was de- ployed onto Android devices for further testing, validation, and user feedback.

Fig. 4. Bird Species Recognition by Image

Fig. 5. Bird Species Recognition by Bird's audio

VII. Conclusion

The journey of this project reflects a symphony of inter- disciplinary collaboration and innovation where data science, deep learning, and mobile app development work in perfect harmony to create a technological marvel. Every stage of its development, painstakingly planned and carried out, is evi- dence of the effectiveness of cooperative integration between various fields. The meticulous selection of library imports and the deliberate blending of unrelated datasets during the project's inception set the stage for its later successes. A powerful machine learning model is developed by means of thorough experimentation and iterative refinement. It is then expertly trained and meticulously optimised after training. The foundation of the project's success is this model, which is a monument to the unrelenting pursuit of excellence and embodies the marriage of state-of-the-art algorithms with meticulous craftsmanship.

Yet, the project's journey extends far beyond the confines of the laboratory. With the seamless integration of TensorFlow Lite into the Android application The project is a ray of hope that points the way to a future in which technology and people live in harmony with the natural world as we look towards the horizon of possibility. Its legacy, which is intertwined with environmental awareness and scientific advancement, is proof of the persistent value of innovation when applied for the greater good. And as we set out on this path of inquiry and learning, driven by the values of compassion and curiosity, we welcome the hope of a better future in which the wonders of nature are appreciated and conserved for future generations.

VIII. Future Scope

In considering future enhancements, it is clear that the framework's audio model has shown a great deal of room for development in order to further increase its accuracy. As a result, there's a great chance for improvement that will boost its effectiveness and performance.

This proactive optimisation step shows a dedication to enhanc- ing the overall robustness and reliability of the application, which in turn strengthens its capacity to correctly identify bird species using both visual and auditory modalities.

Furthermore, investigating the incorporation of environmental sensing technology shows potential for anticipating the exis- tence of particular bird species through environmental cues. By giving users information about the kinds of birds that are likely to live in a particular area, this creative method could completely change birdwatching while also enhancing the experience and promoting a closer bond with nature.

Acknowledgment

We would like to express our sincere appreciation to Prof Merrin Solomon from UMIT for their invaluable contributions to this research. Their expertise, guidance, and support have been instrumental in the completion of this work.

References

[1] Bird Image Classification using CNN Transfer Learning Architectures by Asmita Manna, Nilam Upasani, Shubham Jadhav, Ruturaj Mane, Rutuja Chaudhari, Vishal Chatre January 2023 International Journal of Advanced Computer Science and Applications 14(3)

[2] Building a bird recognition app and large-scale dataset with citizen sci- entists by :Grant Van Horn, Jessie Barry, Steve Branson, Panos Ipeirotis, Ryan Farrell, Pietro Perona, Scott Haber, Serge Belongie June 2015 ,Conference: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

[3] .PakhiChini: Automatic Bird Species Identification Using Deep Learning byKazi Md Ragib, Raisa Taraman Shithi, Shihab Ali Haq, Md Hasan, Kazi Mohammed Sakib, Tanjila Farah. July 2020 ,Conference: 2020 Fourth World Conference on Smart Trends in Systems Security and Sustainablity (WorldS4)

[4] Automating bird species classification: A deep learning approach with CNNs and EfficientNet by Renjun Cai Nanjing Forestry University, Nanjing, China

[5] Bird Species Detection From Voice Features by Rachana B, Kavya Hegde, Navya Bhat Department of Computer Science, Srinivas Institute of Technology, Man- galore, Karnataka, India

[6] Automatic bird species recognition based on birds vocalization by Jiri Stastny1,2 , Michal Munk3 and Lubos Juranek1