International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Detection of Bird Species Using Acoustic Signals

Dr. Bhavana R.Maale, Rabiya Anjum

Department of computer science , VTU,CPGS
Kalaburagi, Karnataka, India

Department of computer science ,VTU,CPGS
Kalaburagi, Karnataka, India

sg.bhavana@vtu.ac.in

rabiyaanjum410@gmail.com

Abstract— The task at hand is to use supervised learning to determine which bird species are audible in a given recording. In order to extract valuable ecological data from field recordings, it may be necessary to first develop efficient algorithms for classifying bird species. In this study, we use SVM (Support vector machine) technique to categorize bird chipping sounds into their respective species. Memory management, the availability of high-quality bird species for machine recognition, and a difference in signal-to-noise ratio across sets used for training and testing all posed problems for this research. We utilized the SVM technique to solve these problems and found that it provided satisfactory results. In this case, SVM is the most effective technique for dealing with recognizing problems. simply, then CNN tweaking, testing, and categorization. Birds are categorized based on their attributes (size, color, species, etc.) & outcome is contrasted with pre-trained data to provide output. In addition, we use the KNN, LR, and random forest algorithms. In this project, we provide information on many species, including their native territory, diet, name, maximum age, plumage color, body length, and whether or not they are migratory. Hearing who it is on the other end of the line only by hearing their voices. People with low vision may benefit from this endeavor. The concept involves using bird calls to provide early warnings of impending rain.

Keywords— CNN, KNN, LR, SVM, Bird Species, visually impaired

I. Introduction

There are a number of scientific and environmental reasons why listening for birds is so essential. Some crowdsourcing & remote monitoring programs even do automated sound analysis upon recordings they collect. Many birds are far more easily observable by sound than by sight or other signs, making the audio modality ideal for bird monitoring. Walk through the processes used to identify bird species and the problems that need fixing. Then, explain how you plan to push the state of the art forward by providing a new knowledge challenge using fresh public data sets. However, it is necessary to first describe the contexts in which audio bird detection might be useful. In the most fundamental sense, a detector will return a value of 0 if no instances of target species are present in the input sound clip, and a value of 1 otherwise. The user is responsible for running the Non-Real time program. The article makes use of Kaggle datasets including recordings of bird sounds. The underlying principle is Many people choose support vector machine (SVM) above other methods because it achieves high levels of accuracy while utilising very little processing time. Both classification and regression problems are amenable to SVM's predictive power. However, its primary value lies in its categorization applications. What follows is an explanation of following four sections. In Section 2, we present relevant studies and provide a brief description of bird species recognition problem; in Section 3, we outline database used in popularity experiments along with the initial signal pre-processing, syllable classification, feature extraction procedures, as well as demonstrate classification algorithm; in Section 4, we portray experimental outcomes; then in Section 5, we draw conclusions and point to potential future research directions.

Human interference in their habitats and complete habitat destruction, coupled with environmental disasters like global warming, forest fires, as well as other natural calamities, pose a threat to the incredible behavioral and morphological diversity of bird ecosystem. As of the year 2020, 1,481 avian species, or 13.5% of all species for which there is adequate data, are threatened with extinction owing to their shrinking or extinguished ranges. Keeping an eye on bird populations may help with environmental control and evaluation. Some bird populations have declined due to environmental degradation. Therefore, recognizing bird species may help in early detection and mitigation of ecological threats. Since birds are so attuned to their surroundings, we may use them to help us spot different types of life. The collection and organization of data on bird species, however, requires substantial human labor and is thus prohibitively costly. In this case, a reliable system will not only give a wealth of data for bird enthusiasts, but will also be an essential resource for scientists and government officials. The vast majority of individuals are also unable to recognize many species that we see on a regular basis. The goal of this study is to create a computer program that can recognize bird calls automatically. There are several obstacles to overcome when trying to detect and identify birds using audio signals, since background noise like rain or traffic noises often overlap with bird syllables, making detection process more challenging. An unreliable method that highlights need for automated systems is

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

manual spectrogram examination, that needs a large number of professionals & might lead to mistakes. These devices might have financial value since bird watching has become a popular pastime.

Figure 1 :Great tit (Parus major)

It has a basic, vibrant rhythmic verse that sounds a little bit mechanical, like *"te-ta te-ta te-ta,"* or *"te-te-ta te-te-ta,"* with a distinct stress on the third syllable.

Figure 2: Sparrow

sparrows-73709.mp3

II. Related Work

Shidong Pan; Dehai Zhao; Weishan Zhang 2021, [3], The identification of bird species using their calls recorded by implanted equipment is a difficult but crucial undertaking. BIRD ID is a proposed scheme. In order to identify which species of birds are being called, signals are first transformed in spectral field and then fed into a deep neural network.

Fan yang, ying jiang, and yue xu, [4], 2022, Currently available models for recognizing bird calls have a low degree of generalizability, and feature extraction from bird calls requires a complex approach. To solve these issues, we propose a lightweight bird sound recognition model for building a lightweight recognition and extraction of features network using MobileNetV3 as the backbone. The data set contains 264 different types of birds, which should help improve the model's generalization ability. To further adapt network to scale extraction of spatial data as well as channel information, we develop multi-scale feature fusion structure and add PSA module to it.

Noumida A.; Rajeev Rajan [5] 2021, Since birds play such a crucial part in ecosystems, the identification of bird species from audio recordings is both difficult & of great scientific significance. With the use of CNN, DNN, & transfer learning techniques, we hope to create an efficient method of bird call categorization for single-label recordings. Many deep learning applications make utilization of transfer learning models. There was comparison of acoustic MFCC-DNN technique's efficacy to that of transfer learning models like ResNet50, VGG-16, as well as InceptionResNetV2.

Yo-Ping Huang; [6], 2021, Due to interclass similarities & fine-grained traits, recognizing bird species from pictures is a difficult and time-consuming operation. Because of this, in this research we sought to build a transfer learning-based approach employing Inception-ResNet-v2 for the detection and classification of Taiwan-endemic bird species, as well as the ability to identify these species from other object domains.

B. Zhao, X. Wu, J. Feng[11], Q. Peng, & S. Yan,. 2017., In their next article, they used ResNet to install a Convolutional Neural Network for classification. In order to train neural network upon most relevant data, they have segmented recordings of bird songs in signal and noise categories. Noise segments are leveraged to improve the quality of the training samples, which in turn improves generalization.

Literature [12] X. Zhang, H. Xiong, W. Zhou, W. Lin, & Q. Tian, 2016. suggested a system that use the Convolutional Neural Network technique to identify birds by their characteristics (such as genus, species, subspecies, size, color, etc.). The dataset is mined for audio recordings, and those files are then converted to WAV.

III. Proposed System

This project makes use of the Support Vector Machine (SVM) implementation technique, which is a supervised learning algorithm suitable for either binary classification or regression. It's a positional indicator for separate measurements. choice plane support determines the extent of a choice. The system's foundation is environments, and specifically the use of non-real-time bird calls. We also use LR, KNN, & RF. This project aims to provide information on a variety of species, including where they live, what they eat, any potential environmental threats, potential mates, and potential neighbors.

IV. Methodology

Input audio

There are two stages involved in the construction of an audio categorization system: the training phase & assessment stage. The first stage involves use of training data set to teach system how to recognize classes, while second stage involves use of a test data set to assess how well system performs in comparison.

In order to properly categorize incoming noise, it typically requires some kind of pre-processing.

Using the Python program, this is how to feed an audio file to model:

1. Select a file on gadget to begin. Once again, we need to do preprocessing on the data in which we use MFCC to extract the characteristics of the data itself.
2. Now that we have characteristics of the specific file, we can use built model to forecast class labels. Model then assigns a label based on this.
3. Finally, an inverse transformation of the label is required to get class name that identifies specific bird from the audio.

Pre-processing

Noise cancellation and data transformation are only two examples of the types of operations that could be performed during the pre-processing phase.

Segmentation

During segmentation process, data is either partitioned in parallel segments or extracted as individual units. In this research, syllables from recordings of birdsongs are isolated.

Feature Generation

Since segments can be associated with a number of characteristics or features during the feature creation phase, this stage is commonly referred to as the data reduction phase. When choosing features, it is important to incorporate data that may help distinguish across classes. While data reduction may be a part of the feature-generation process, it still requires feature selection from a bigger pool for categorization purposes. The training of the classifier using the training data occurs during the classifier design phase. At this stage, we establish the criteria by which we will divide our classes. After a classifier has been developed and trained, its effectiveness in performing a particular classification job is assessed during the system assessment phase.

Pseudo Code for Pre- & Feature processing

Step 1: Read from read_csv
Step 2: Read audio file based on file_id
Step 3: Fetch sg, mask, data, audio_mask, sample_rate
Step 4: Determine window size
Step 5: Extract features from each audio file for each audio frame get Species, genus, spec_centr_, chromogram_ update weight file
Step 6: write weight file to external csv file

Prediction Pseudo code:

Step 1: Read Input Audio File
Step 2: Fetch sg, mask, data, SampleRate, Audio_Marks
Step 3: Extract Features
Step 4: Import weight file
Step 5: Compare Model weights with Input audio features
Step 6: Displays Bird Species Name

Process for Bird Voice Recognition into Non-Realtime

Step 1: Start
Step 2: Choose the audio file
Step 3: if button is equal to Naïve Bayes then Naïve Bayes algorithm is used for recognition.
Step 4: Bird Species displayed
Step 5: End

Classification: Support Vector Machine

Support vector machine (SVM) is commonly employed for the purpose of pattern classification. There exist various types of patterns, namely linear and non-linear. Linear patterns are characterized by their distinguishable nature and separability in low dimensions. On the other hand, nonlinear patterns exhibit a lack of distinguishability and separability, requiring additional manipulation to achieve ease of separation. The fundamental concept underlying SVM is creation of an optimal hyperplane that enables classification of linearly separable patterns. The optimal hyperplane refers to a hyperplane chosen from a collection of hyperplanes used for classifying patterns. It is selected based on its ability to maximize margin, that is defined as distance between hyperplane & closest point with every pattern. Primary goal of SVM is to optimize margin in order to accurately classify provided patterns. In other words, a larger margin size corresponds to more accurate classification of patterns.

<u>Algorithm 1:</u>

SVM Algorithm Pseudo Code:

```
Candidate SV={closest apir from opposite classes}
While there are violating points do
   Find a violator
Candidate SV=U candidate SV
     S
Violator
If any α<0 due to addition of c to S then
   Candidate SV=candidate SV\p
  Repeat till all such points are pruned
End if
End while
```

Sensitivity is a measure of how well test can detect a condition when it is present.

Sensitivity = TP / (TP+FN)

Specificity is a measure of how well a test excludes the condition (but does not differentiate it) when it is absent. Along these lines,

Specificity = TN / (TN+FP)

Number of true positives in relation to existence of condition is predictive value.

Predictive worth positive = TP / (TP+FP).

The magnitude of negatives associated with deficit in the condition is predictive value negative. Hence,

Predictive worth negative = TN / (TN+FN).

V. System Architecture

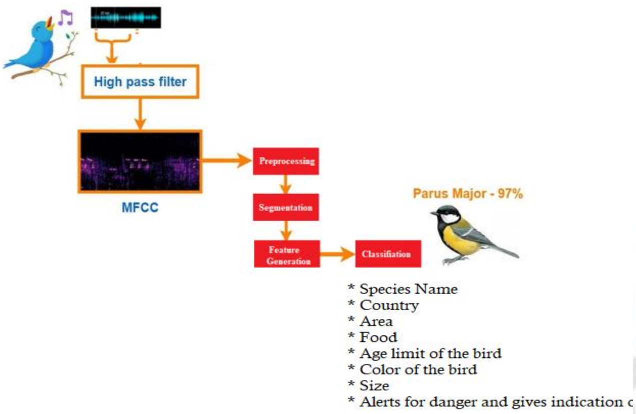

Figure 3: System Architecture

VI. Implementation

Figure 4: Menu

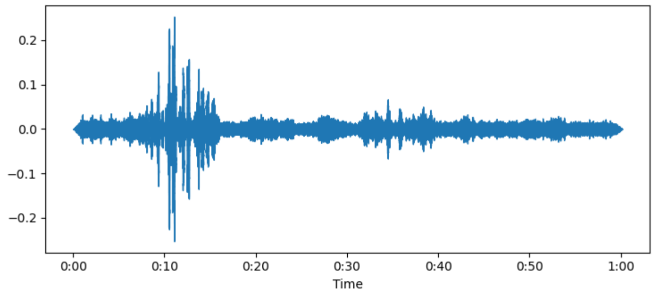

Figure 5: Input voice of Bird

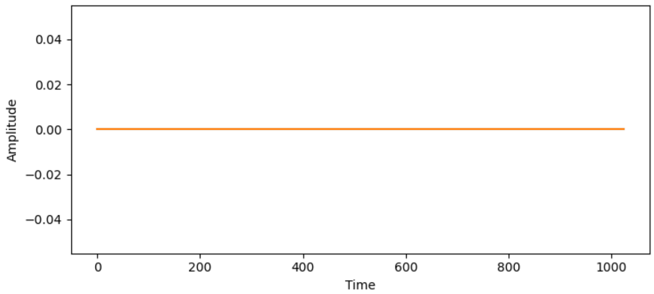

Figure 6: Time Vs Amplitude

VII. Conclusion

The relevance of bird species detection in scientific study is growing. The application builds on prior work to solve the problem of manually identifying non-real-time bird calls. This study showcases a novel SVM approach to avian species recognition in a simulated environment. The results of the studies show that the SVM achieved a non-real-time accuracy of 98% in its classifications. Select an appropriate algorithm and use the audio recording to identify the bird species. In order to further improve accuracy, we want to evaluate and implement additional features in future work that are unrelated to syllable spectral synthesis. It is also possible to incorporate additional characteristics derived from phrases and songs that demonstrate syllable relationships. Furthermore, it is possible to implement in real time. To better serve your users, I recommend developing an Android or iOS app instead of a website. The system's implementation on the cloud, which can store massive amounts of data for analysis, is possible.

References

[1] Hasan Abdullah Jasim, “CLASSIFY BIRD SPECIES AUDIO BY AUGMENT CONVOLUTIONAL NEURAL NETWORK”,2022
[2] Samruddhi Bhor “Automated Bird Species Identification using Audio Signal Processing and Neural Network”, 2022,
[3] Shidong Pan; Dehai Zhao; Weishan Zhang “CNN-based Multi-model Birdcall Identification on Embedded Devices” 2021
[4] FAN YANG, YING JIANG, AND YUE XU, “Design of Bird Sound Recognition Model Based on Lightweight”, 2022

[5] Noumida A.; Rajeev Rajan Deep Learning-based Automatic Bird Species Identification from Isolated Recordings, 2021
[6] Yo-Ping Huang; Haobijam BasantaRecognition of Endemic Bird Species Using Deep Learning Models, 2021
[7] Tianjun Xiao, Yichong Xu, Kuiyuan Yang, Jiaxing Zhang, Yuxin Peng, and Z. Zhang, "The application of two-level attention models in deep convolutional neural network for fine-grained image classification," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2015.
[8] M. Simon and E. Rodner, "Neural Activation Constellations: Unsupervised Part Model Discovery with Convolutional Networks," 2015 IEEE International Conference on Computer Vision (ICCV), Dec. 2015.
[9] Y.; Li, J.; Xiao, H.; Jin, X.; Yan, S.; Feng, J. "Path and Dual Consistency," Constraint Networks, pp. 319–354.
[10] T.-Y. Lin, A. RoyChowdhury, and S. Maji, "Bilinear CNN Models for Fine-Grained Visual Recognition," 2015 IEEE International Conference on Computer Vision (ICCV), Dec. 2015.
[11] B. Zhao, X. Wu, J. Feng, Q. Peng, and S. Yan, "Diversified Visual Attention Networks for Fine-Grained Object Classification," IEEE Transactions on Multimedia, vol. 19, no. 6, pp. 1245–1256, Jun. 2017.
[12] X. Zhang, H. Xiong, W. Zhou, W. Lin, and Q. Tian, "Picking Deep Filter Responses for Fine-Grained Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016.
[13] Y. Zhang, X.-S. Wei, J. Wu, J. Cai, J. Lu, V.-A. Nguyen, and M. N. Do, "Weakly Supervised Fine-Grained Categorization With PartBased Image Representation," IEEE Transactions on Image Processing, vol. 25, no. 4, pp. 1713–1725, Apr. 2016.
[14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016.
[15] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jul. 2017.
[16] W. Zeng, Y. Wang, and R. Jiang, "Integrating distal and proximal information to predict gene expression via a densely connected convolutional neural network," Jun. 2018.
[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016.
[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2016.
[19] Y. Bengio "Biophysical Mechanisms of Invariance: Unsupervised Learning, Tuning, and Pooling," Visual Cortex and Deep Networks, 2016
[20] T.-Y. Lin, A. RoyChowdhury, and S. Maji, "Bilinear CNN Models for Fine-Grained Visual Recognition," 2015 IEEE International Conference on Computer Vision (ICCV), Dec. 2015.
[21] G. Zheng, M. Tan, J. Yu, Q. Wu, and J. Fan, "Fine-grained image recognition via weakly supervised click data guided bilinear CNN model," 2017 IEEE International Conference on Multimedia and Expo (ICME), Jul. 2017.
[22] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a Similarity Metric Discriminatively, with Application to Face Verification," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05).
[23] S. R. K, B. C. K, O. ooruchintala, D. B. Mandru, and kallam suresh, "Deep Convolution Neural Networks Learned Image Classification for Early Cancer Detection Using Lightweight," Oct. 2021.
[24] D. Zhai, X. Liu, H. Chang, Y. Zhen, X. Chen, M. Guo, and W. Gao, "Parametric local multiview hamming distance metric learning," Pattern Recognition, vol. 75, pp. 250–262, Mar. 2018.
[25] E. Hoffer and N. Ailon, "Deep Metric Learning Using Triplet Network," Lecture Notes in Computer Science, pp. 84–92, 2015.
[26] R. Lu, K. Wu, Z. Duan, and C. Zhang, "Deep ranking: Triplet MatchNet for music metric learning," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mar. 2017.
[27] J. M. Hancock, "Bayesian Network (Belief Network, Causal Network, Knowledge Map, Probabilistic Network)," Dictionary of Bioinformatics and Computational Biology, Oct. 2004.