Bird Sound Recognition

Varad Ingale Vithika Pungliya Atharva Purohit

[varad.ingale20@vit.edu](http://varad.ingale20@vit.edu) [vithika.pungliya20@edu](about:blank) [atharva.purohit20@vit.edu](http://atharva.purohit20@vit.edu)

D D D

Mentor: Ridya Ghosh

**Introduction -**

As the “extinction capital of the world,” Hawai'i has lost 68% of its bird species, the consequences of which can harm entire food chains. Researchers use population monitoring to understand how native birds react to changes in the environment and conservation efforts. But many of the remaining birds across the islands are isolated in difficult-to-access, high-elevation habitats. With physical monitoring difficult, scientists have turned to sound recordings. Known as bioacoustics monitoring, this approach could provide a passive, low labor, and cost-effective strategy for studying endangered bird populations.

Current methods for processing large bioacoustics datasets involve manual annotation of each recording. This requires specialized training and prohibitively large amounts of time. Thankfully, recent advances in machine learning have made it possible to automatically identify bird songs for common species with ample training data. However, it remains challenging to develop such tools for rare and endangered species, such as those in Hawai'i.

The Cornell Lab of Ornithology's K. Lisa Yang Center for Conservation Bioacoustics (KLY-CCB) develops and applies innovative conservation technologies across multiple ecological scales to inspire and inform the conservation of wildlife and habitats. KLY-CCB does this by collecting and interpreting sounds in nature and they've joined forces with Google Bioacoustics Group, LifeCLEF, Listening Observatory for Hawaiian Ecosystems (LOHE) Bioacoustics Lab at the University of Hawai'i at Hilo, and Xeno-Canto for this competition.

In this competition, we’ll use your machine learning skills to identify bird species by sound. Specifically, you'll develop a model that can process continuous audio data and then acoustically recognize the species.

This project will help advance the science of bioacoustics and support ongoing research to protect endangered Hawaiian birds.. It will be easier for researchers and conservation practitioners to accurately survey population trends. They'll be able to regularly and more effectively evaluate threats and adjust their conservation actions.

**Literature Survey -**

| **Project Title** | **Author** | **Methodology** | **Limitations** |
| --- | --- | --- | --- |
| Overview of BirdCLEF 2021: Bird call identification in soundscape recordings | Stefan Kahl1 , Tom Denton2 , Holger Klinck1 , Hervé Glotin3 , Hervé Goëau4 , Willem-Pier Vellinga5 , Robert Planqué5 and Alexis Joly6 | This paper describes how the various algorithms were evaluated and synthesizes the results and lessons learned. Deep convolutional neural networks were the go-to tool in this competition.  In many cases, participants chose to use off-the-shelve architectures pre-trained on ImageNet (like EfficientNet , DenseNet , or ResNet ).Additionally, in 2021 Competition , off-the-shelve CNN backbones pre-trained on ImageNet provided strong results without the need to investigate the design of domain-specific architectures further. |  |
| Bird Call Recognition using Deep Convolutional Neural Network,  ResNet-50 | Mangalam Sankupellay and Dmitry Konovalov | In this paper, we use ResNet-50, a deep convolutional neural network architecture for automated bird call recognition.  They used a publicly available dataset consisting of calls from 46 different bird species. Spectrograms (visual  features) extracted from the bird calls were used as input for ResNet-50. They were able to achieve 60%-72% accuracy of bird call recognition using ResNet-50. |  |
| Bird Sound Recognition Using a Convolutional  Neural Network | A´ gnes Incze, Henrietta-Bernadett Jancso´, Zolta´n Szila´gyiy, Attila Farkasy and Csaba Sulyok | In this paper, a CNN system classifying bird sounds is presented and tested through different configurations and hyperparameters.The MobileNet pre-trained CNN model is fine-tuned  using a dataset acquired from the Xeno-canto bird song sharing  portal. | The presented system is  viable only for a low number of classes. |

**Methodology -**

1. The first step in our project includes Exploratory Data analysis (EDA).
2. Through our EDA, we found the labels of the birds and the audio tapes under each label. After plotting the numbers, the graph was very skewed, hence only 16 classes according to their top rating of audio quality (rating greater than 4 out of 5) were selected.
3. The birds that were selected are, 'brnowl', 'cangoo', 'comsan', 'dunlin', 'eurwig',

'gnwtea', 'houfin', 'houspa', 'mallar3', 'norcar', 'normoc', 'rinphe', 'rorpar', 'skylar', 'wesmea'.

1. The next step is Data Pre-processing, which involves converting the audio files to

Image spectrograms and storing those images under the same labels.

1. The total number of audio files used were 3919, and 9539 spectrograms were

generated.

1. Now the dataset is split in the ratio of 80/20, for training and validation purposes.
2. There are 7637 training images and 1902 validation images.
3. The images were resized to 224 by 224. The number of epochs was set to 10. Batch size was set to 64 and learning rate was 0.001.
4. Three models were implemented which were Convolutional Neural Networks,

Resnet50 and EfficientnetB0.

1. Convolution Neural Network consist of 4 Convolutional Layers, 4 MaxPooling Layers,2 dense layers and 1 flattened layer. The activation function in hidden layers is RELU and the optimiser used is Adam. The activation function of the output layer is

sigmoid.

1. The result obtained was not very good as the validation accuracy was 14.72% in the

last epoch. So, this model was discarded.

1. Next model was Resnet50, which is a transfer learning model, which has 48

convolutional layers, 1 MaxPool, 1 Average Pool layer and 1 dense layer on top of this base model.

of Resnet50, a GlobalAveragePooling2D layer, flatten layer and Dense layer were

fitted.

1. The accuracy came out to be as 83.77 % and validation accuracy was 31.09 %, as

This model was majorly overfitting; it was discarded.

1. The last model trained was the EfficientnetB0 as the base model and using the 1

MaxPool, 1 Average Pool layer and 1 dense layer on top of this base model.

1. The accuracy obtained was 91.27% and validation accuracy was 73.98%.

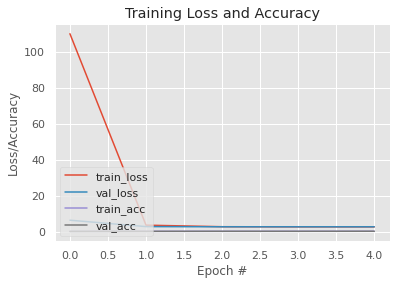
**Additional Information - (Section heading to be renamed to the specific technology / model)**

The dataset is very skewed, therefore only some few classes have to be selected.

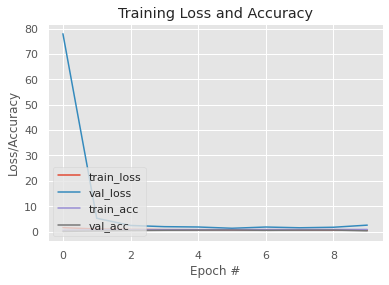
**Results –**

| Model | Accuracy | Validation Accuracy |
| --- | --- | --- |
| CNN | 13.88 | 14.72 |
| Resnet50 | 83.77 | 31.09 |
| EfficientnetB0 | 91.27 | 73.98 |

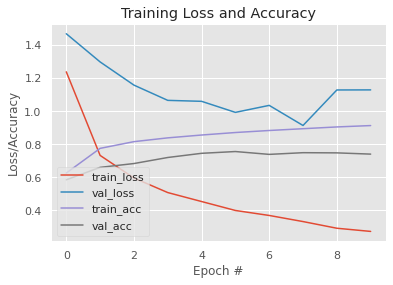
Loss Curve of CNN:



Loss Curve of Resnet50:



Loss Curve of EfficientnetB0:



**Conclusion –**

The project aims to process continuous audio data and then acoustically recognize the species.

In this project we understood the handling of audio datasets. We converted the audio files to image spectrograms. These images were later trained on 3 models CNN, Resnet50 and EfficientNetB0 out of which efficient net gave us the highest accuracy of 91.27% and validation accuracy of 73.98%. This project will help advance the science of bioacoustics and support ongoing research to protect endangered Hawaiian birds. It will be easier for researchers and conservation practitioners to accurately survey population trends.

**References –**

1. Bird Sound Recognition Using a   Convolutional Neural Network- <https://sci-hub.hkvisa.net/10.1109/sisy.2018.8524677>

1. Bird Call Recognition using Deep Convolutional Neural Network, ResNet-50- <https://www.researchgate.net/profile/Dmitry-Konovalov-2/publication/328418948_Bird_Call_Recognition_using_Deep_Convolutional_Neural_Network_ResNet-50/links/5bcd62dc458515f7d9d02755/Bird-Call-Recognition-using-Deep-Convolutional-Neural-Network-ResNet-50.pdf>

1. [SOUND-BASED BIRD CLASSIFICATION. How group of Polish women used deep… | by Magdalena Kortas | Towards Data Science](https://towardsdatascience.com/sound-based-bird-classification-965d0ecacb2b)

1. <http://ceur-ws.org/Vol-2936/paper-123.pdf>