Engenia Solutions Pty Ltd.

Internship Aptitude Test: Bird Call Identification Using Deep Learning

Introduction

Globally, there are more than 10,000 different species of birds. Each species has a unique set of calls and songs. Birds are great indicators of pollution and declining environmental quality. Monitoring biodiversity trends and conditions in ecosystems is essential. Accurate sound identification and categorization, enabled by machine learning, allows scientists to improve their capacity to support international conservation initiatives.

In this test, you will identify three different types of bird species using deep learning:

1. Brown Tinamou
2. Great Tinamou
3. Cinereous Tinamou

You can download the data from: https://xeno-canto.org/collection/species/all

In audio signal classification, the following features are commonly used:

- Spectrograms
- Mel Spectrograms
- Mel-Frequency Cepstral Coefficients (MFCC)

You can use these features for your classification task. However, you are not restricted to only using these; you can use any audio feature you wish. Additionally, you can employ any deep-learning algorithm, pre-trained model, or transfer-learning method for this task. Python is preferred, but you may use any programming language.

Task Definition

We need an accurate deep-learning model to identify Brown Tinamou, Great Tinamou, and Cinereous Tinamou bird species by their sound. This model should be suitable for identifying these species from long-duration (1 – 24 hours) sound recordings.

Your model should:

- Identify the above bird species with timestamps if they are present in any given soundtrack.
- Not assign any other bird species to training species classes (Brown Tinamou, Great Tinamou, and Cinereous Tinamou) if a soundtrack does not include these species but includes any other species. This requirement reduces false positives.

Steps

Here are the suggested steps you can follow:

1. Create your dataset (Do not download more than 50 soundtracks for a single class).
2. Apply data preprocessing/augmentation methods if needed.
3. Convert raw audio files into features.
4. Split the dataset into train, validation, and test sets.
5. Train your model.
6. Evaluate your model.
7. Optimize your model if needed.
8. Create a method for obtaining predictions from user-input sound files.

If you can deploy your model on an AWS server as a web application, you will earn additional marks.

Submissions

After completing the task, create a private GitHub repository, push all code files into it, and then share the repository link with us: don@engeniasolutions.com.au, and chamika@engeniasolutions.com.au.

Your repository should include:

- All code files.
- Your model.
- Your dataset.
- Model evaluation results (precision, recall, accuracy score for each class).
- A simple document describing how to use your model to get results from user-input sound files.

Deadlines

Complete the task within four working days. If you cannot finish by the deadline, submit what you have done so far. Late submissions will not be accepted.

Evaluation and Awards

We will evaluate your model within two working days and send the results via email. The top three participants scoring above 80% will receive awards as follows:

- 1st place: 20,000 LKR
- 2nd place: 15,000 LKR
- 3rd place: 10,000 LKR

Two of these winners will receive the internship opportunity. Ensure to follow these guidelines closely for successful completion of the internship aptitude test. Best of luck everyone.

copyright @engeniasolutions