NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
College of Engineering

Animal Hunt

AI-Based Animal Sound Recognition Application

Student: Lin Yi Heng

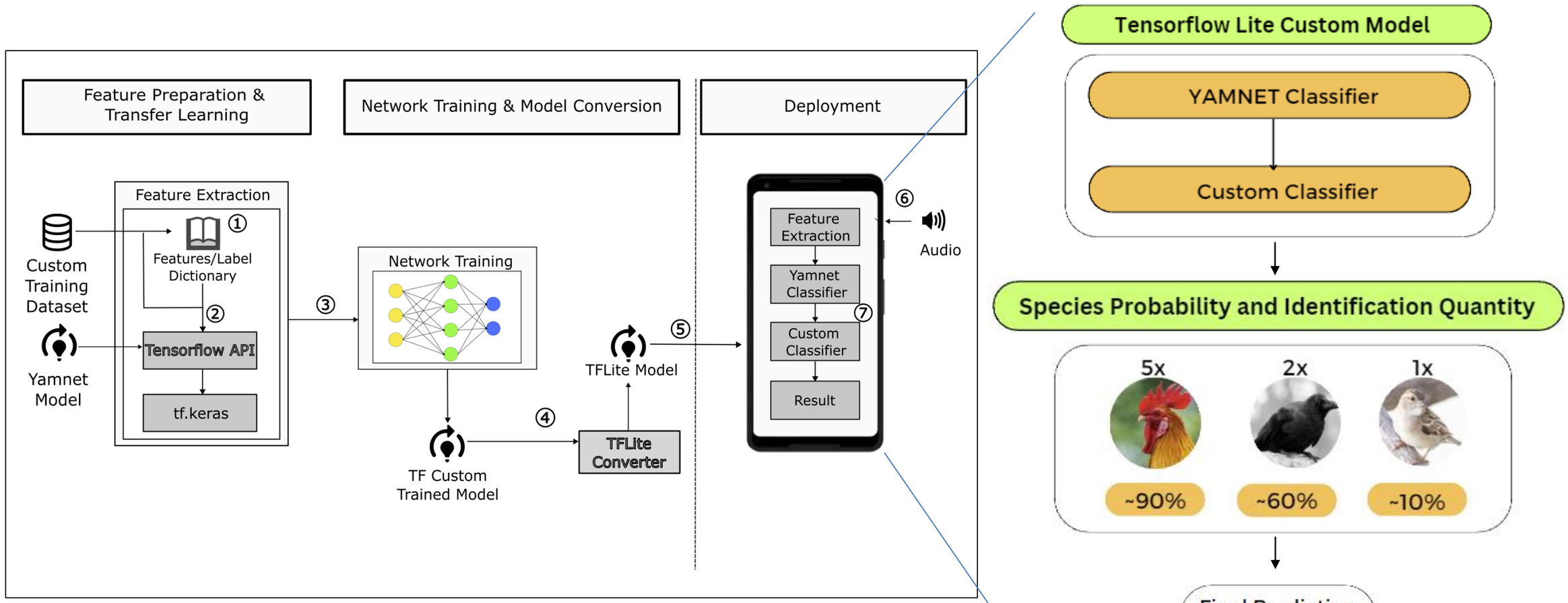
Supervisor: Dr Owen Noel Newton Fernando

Introduction

Measuring unique acoustic features is crucial in analyzing animal sounds, particularly in bioacoustics research where classifications are manually extracted from spectrogram plots. However, this identification process can be time-consuming, taking up to ten times longer than the recording time due to noise pollution. To speed up analysis, most acoustic signal analysis systems utilize complex techniques such as short-time Fourier transform and wavelets. However, these methods are computationally intensive, require a strong network connection, and are challenging to implement on low-cost microcontroller-based systems. These drawbacks show that a new system is needed to accelerate and ease the process of continuous real-time biodiversity monitoring.

Project Objective

As such, the goal of this project is to present a new system for animal sound classification by implementing a pre-trained machine model, Yet Another Mobile Network (YAMNet) to detect noises and animal sounds using transfer learning followed by a custom trained machine model that uses Keras a deep learning-based neural networks that employs TensorFlow, an open-source library for numerical computation and large-scale machine learning to create a custom machine model that identifies which animal is present in real-time. To make the models more efficient, compact and overcome the constraints stated previously, both will be converted into Tensor Flow Lite (TFL) using TensorFlow Lite Converter (TFLC). Lastly, the accuracy of prediction and identification will be evaluated using a mobile device's trained model against test datasets.



Project Method

The figures above outlines how AnimalHunt created, loaded and utilized to make inference. It begins the Feature Preparation phase where Datas are collection (Step 1) The extracted data is then used to train a custom TF model that includes the YAMNet Model as a noise detector and filter, as our custom data pertains only to animals (Step 2).The model trained and converted into a TFLite model that can be embedded into a mobile phone (Steps 3, 4, & 5). During deployment, audio is loaded in real-time (Step 6) and classified using the YAMNet Model as the first inference, followed by our custom classifier for a second inference to identify the animal (Step 7).