Final Results

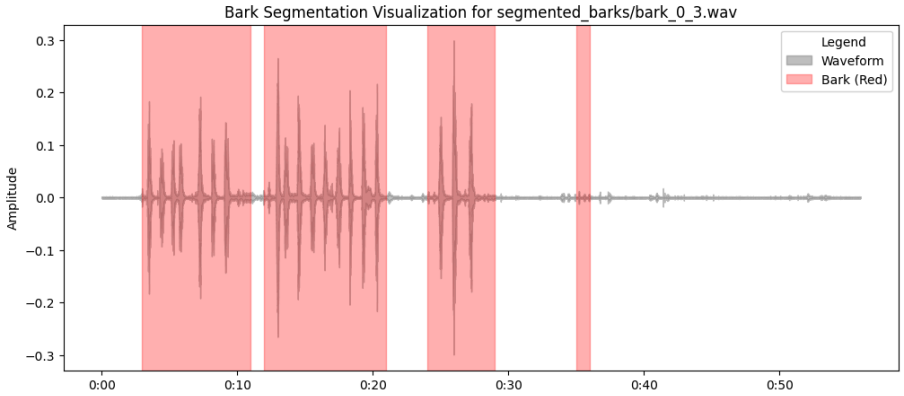
Project Title: Cross-Species Translation
Team Members: Ian Gifford, Deborah Sinishaw, Sara Yang
GitHub Repository: MAIS202_Final_Project
Assigned TPM: Emma

1. Final Training Results

Comparison with preliminary results

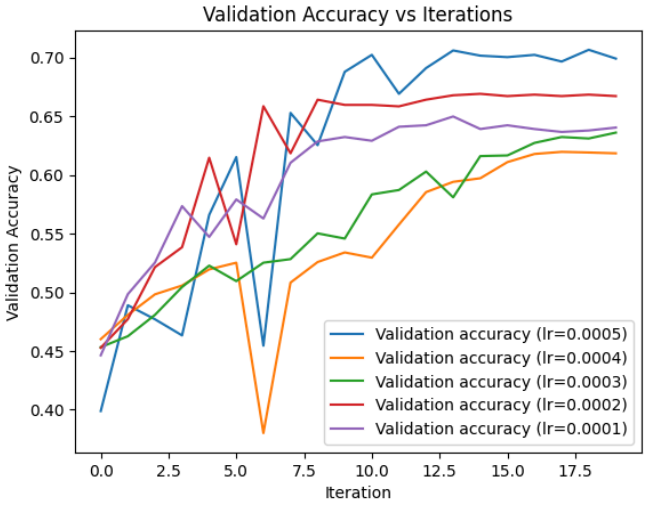
As mentioned in the previous deliverable, we further processed the data from the original dataset found in the github repo. This was done in the following three ways:

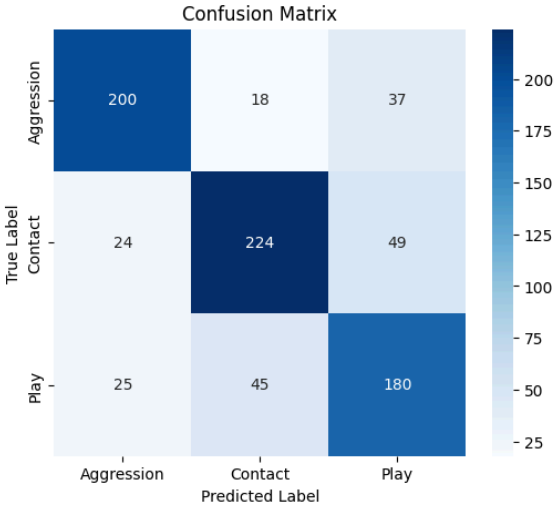
1. Convert the audio files into spectrograms and remove the frequencies that do not contain any significant information. This will increase the efficiency of our model.
2. Segment the barks using a k-means clustering. Note: consecutive barks with less than a ~1 second gap were kept in the same audio file. This is based on the fact that there might be information in strings of barks that cannot be detected if each "spike" in the waveform is segmented.

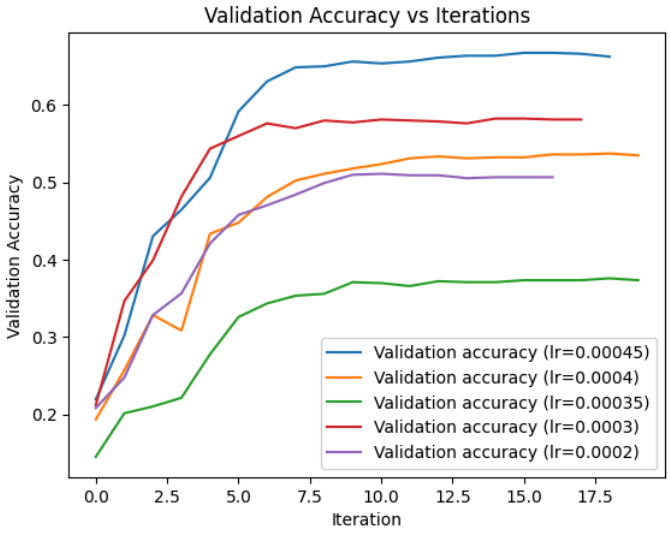


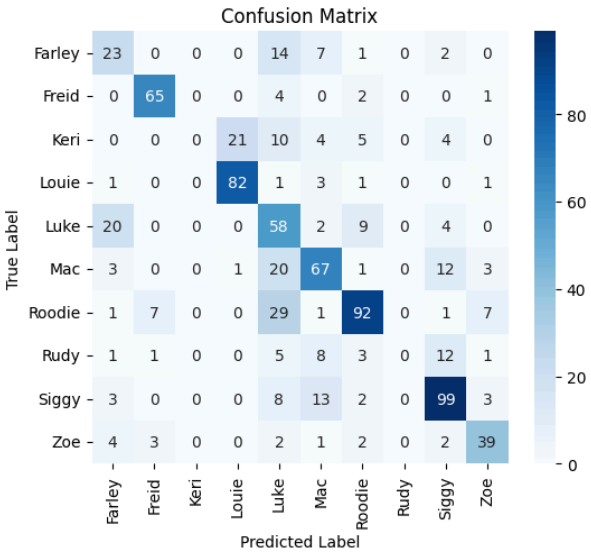
3. Shorten and pad the length of the segmented barks to some standard length, as all barks still varied widely (1s to 20s). The chosen length was 4s since the majority of the dataset did not go past 4s.

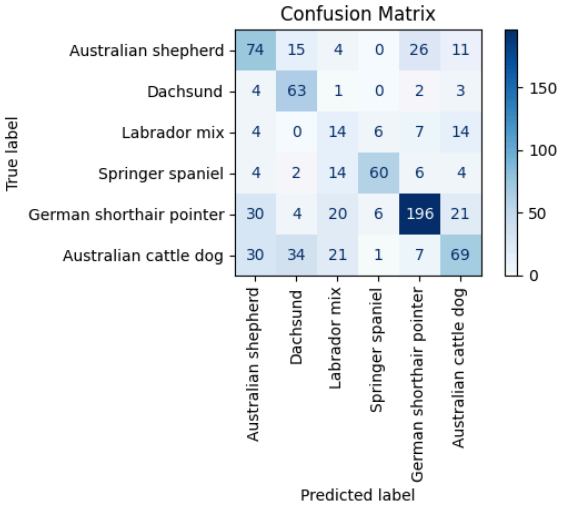
We also implemented a CNN model. This is directly seen in the google colab.

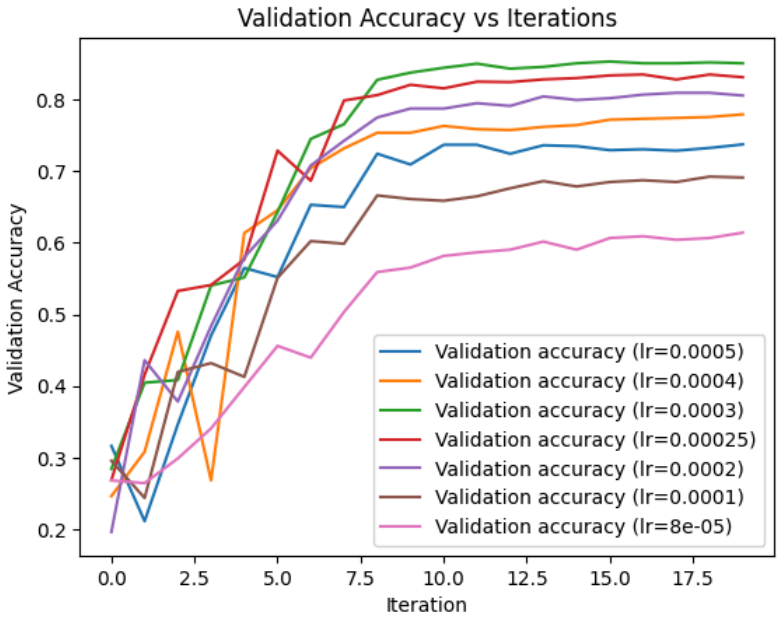
- Predicting the context of bark had the best accuracy of 0.7151 with 0.0005 learning rate.
  - (Will add a confusion matrix)
  - 

  - 

- Predicting dog name had the best accuracy of 0.6671 with 0.00045 learning rate
  - 

  - 

- Predicting dog breed had the best accuracy of 0.8317 with 0.0003 learning rate
  - 

  - 

2. Final Demonstration Preproposal

Overview

Animal communication is a complex system that humans have yet to fully decode. While some species exhibit structured vocalization patterns, interpreting meaning through structure remains an ongoing challenge. This project aims to advance our understanding by developing a

Cross-Species Translation Model, leveraging machine learning techniques to analyze classified animal vocalizations and generate human-readable interpretations of new vocalizations.

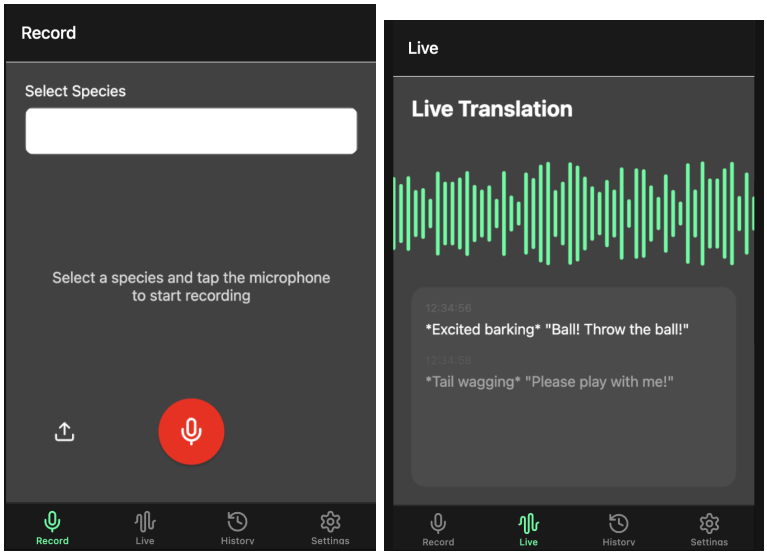
The model will provide users with three functionalities:

1. Identify the Dog's Name – If the dog is known, the model will attempt to recognize its unique vocal signature.

2. Determine the Breed – The system will classify the bark to determine the dog's breed.

3. Analyze the Context – The model will interpret the bark's meaning based on predefined categories (aggression, play, contact).

Design idea for front end:



To best showcase our model and make it easily accessible to users, we have chosen to develop a mobile application using Cursor. The nature of this project—working with real-time animal vocalizations—demands an interface that allows for on-the-go analysis, making mobile integration the most effective solution.

Here is the tech stack that we plan to implement this:

PyTorch → Model Training

Cursor → Designing and running the app

FastAPI (maybe) → Serving the model