Claws and Paws: Sound-Based Pet Identification

Progress Report

Aditya DSS & Aritra Deb

November 2023

GitHub Repository

Q1. Data Pre-processing : what data have you collected so far and what pre-processing have you done?

We are using ESC-50 dataset dataset to get the data. So we have created the data pipline. We have used augmentations like AddGaussianNoise, PitchShift, HighPassFilter, TimeStretch. We also used **MFCC** to extract the features from our audio files.

Q2. Training with the basic model, validation, and completion of the data pipeline?

Input data with 40 features. Three dense layers with 32, 64, and 128 neurons, using ReLU activation. Dropout layer to prevent overfitting. Additional dense layer with 32 neurons. Output layer with 2 neurons and softmax activation for classification. Here's a brief description for the LSTM RNN model:

1) Input data with 40 features. 2) Two LSTM layers with 64 units, using hyperbolic tangent (tanh) activation function. 3) Fully connected layers: one with 32 neurons (ReLU activation) and an output layer with 2 neurons (softmax activation for classification).

Fully Connected Network → Loss: 0.1001, Accuracy: 0.9732, Validation Loss: 0.3000, Validation Accuracy: 0.7500

Total training time: 2.84 seconds.

RNN → Loss: 0.1168, Accuracy: 0.9464, Validation Loss: 0.0885, Validation Accuracy: 1.0000

Total training time: 8.65 seconds.

Q3. Identification of the exact tasks you want to complete for the final submission.

1) We will be using CNN as they are capable of capturing energy modulation patterns across time and frequency when applied to spectrogram-like inputs.

2) We will implement YAMNet (Yet Another Multilabel Network) is a neural network model designed for audio analysis, particularly environmental sound classification.

What challenges you are facing and how you plan to address them. What will be your final deliverables?
As we are having overfitting we will try to reduce overfitting by applying cross-validation and try different models as mentioned above.

Q.C. Each group member needs to identify what they have done so far.

So far **Aditya** have built RNN, ANN and tried out different architectures and **Aritra** set the data pipeline, processing, applied MFCC and Augmentations.

Note: as we have as we mentioned earlier we both were heavily involved in the whole process and contributed equally.