Sound Classification using 1D- Convolutional Neural Networks

Arshiya Khan, ishaani Priyadarshini, Dr. Chase Cotton
Electrical and Computer Engineering University of Delaware, Newark Delaware

Problem Statement

Analysis of sound waves has become convenient since the arrival of machine learning. But, in past sound analysis has been performed using statistical techniques. The project demonstrates the power of neural networks in analyzing time series data without requiring any introductory knowledge of the signal statistics.

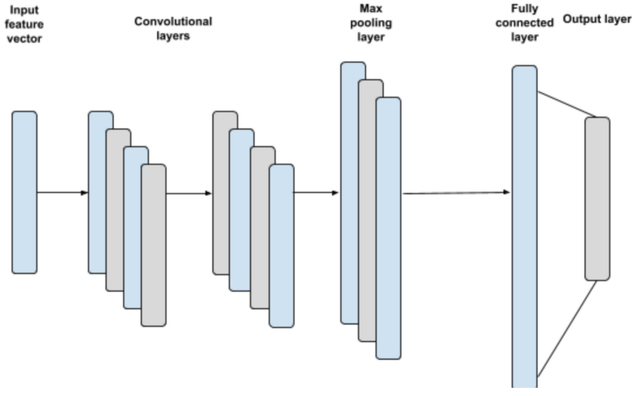
Theory

Classification in machine learning assigns labels to unseen data by learning characteristics from the training set. In our case the items are audio signals. The characteristics are the features we extract from them.

The classes for the problem definition are single tone and tritone

The complexity lies in finding an appropriate relationship between features and classes.

The given figure is a depiction of 1D Convolution Neural Networks. Input vector is the length of one dataset.



Proposed Solution

In the past, several methods have been suggested for sound classification. Normal convolutional neural networks or CNNs which are widely used in image classification all over the world, are two dimensional. They have been very effective in providing accurate results. On the other hand, one-dimensional CNNs are not very common. One very common example is natural language processing problems. But in recent times, 1D CNNs are proving useful in analyzing a multi-dimensional time series data to find patterns or detect anomalies. In our case, it is amplitude of sound waves are a set of 2000 values in quarter of a second.

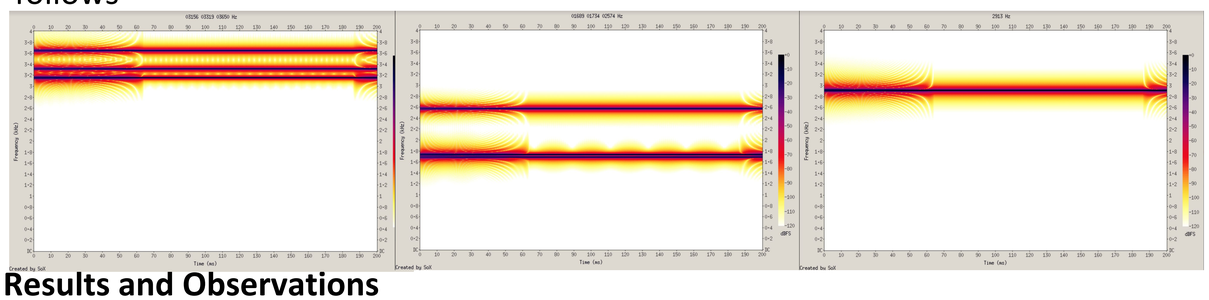
Approach

We are using 1D CNN on raw magnitudes of a wav file. A sample in the training data is a set of magnitude recorded and the instance when it was recorded along wiht labels. We have taken ten samples of each tone at different frequencies.. We have used Keras API to make the layers and stacked them using Sequential module. All the experiments were performed using Python3.

Novelty : Although sufficient research has been performed on Sound Classification, **none of the past researches differentiate between single tone and tritone data. 1D- CNNs have not yet been used for sound classification. Experimental analysis conducted for this study has been performed by generating sound data and not just by relying on available datasets.**

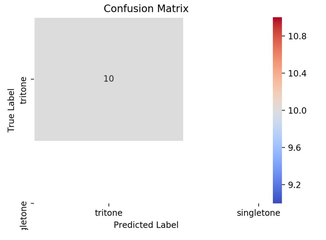
Experimental Analysis

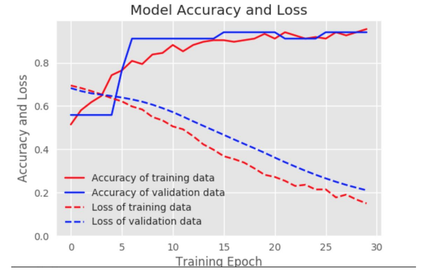
To perform the experimental analysis we have generated our own sound data using Sound Exchange (SOX). The data generated is in form of .wav files with specific ampliude. The spectogram of the sound aka wav files is depicted as follows



Results and Observations

Results have been validated using Confusion Matrix and Classification Report. The **Accuracy percentage is 94%, which validates the effectiveness of our research**





Conclusion and Future Work

This experiment proves that 1DCNNs have been very good at finding patterns in data. The future of this experiment is to make models which can be used to detect anomalies in on recorded data. Moreover, the dataset has been generated by us using SOX due to non availability of the required dataset**.** The dataset is small, therefore in future, a larger dataset may ensure better results.

Thanks to Department of Electrical and Computer Engineering for the support