

Initial Plan: Deep learning to classify guitar tunings

By Arron Sehmbly

Supervisor: David Marshall

Module Number: CM3203

Module Title: One Semester Individual Project

Credits for Module:40

February 3rd, 2019

## **Project Description**

Songs have many instruments with different tunings to produce different sounds. The aim of this project is to classify different guitar tunings from a given dataset of audio files with their respective known tunings and recordings, which will be used in a deep learning model to determine classification. This model will be trained with training data that has the known tunings for the respective recordings in our training data set, it will then will be used on a test data set that has unknown labels for their tuning of the respective audio samples, in order to determine the accuracy of our model results.

The project will require the following things to be researched, first thing would be what is the best format for the input audio. For example, the dataset will be a list of audio files e.g. .wav. Every recording will be segmented based on a time step (the sample size of our audio files will be determined later in the project). In order for our deep learning model to classify the audio files into the respective tunings it will likely be more useful to use transformations(for example SIFT, constant Q transformations) on the data set to perform some spectrogram analysis using our deep learning model.

The second thing we will need to research is what type and configuration of network is best. This will include research on different neural network models, for example convolutional neural networks and research on what are the best model parameters in order to optimize the network for this problem domain. I will need to look at how many layers are needed to construct the network, the size of each layer, activation functions, we will also need to optimize hyperparameters like the learning rate of the model. The last main thing to research would be how to format class labels for our network. In this case there are only two possible ways, first is to have a n-class classifier trained for example there a class label for each tuning we chose to evaluate in our training data set. The second way to format class labels is to assign a tuning set for each string based on our training data set (i.e. a classifier for each string), for example String E would have standard tuning, dropped D tuning etc. Other research that may need to be included would be research on testing our model to evaluate the accuracy of our model in this domain such as cross validation. Another thing that may need to be considered is data augmentation of our data set to train the neural network more effectively before using our dataset with cross validation. Also, some filtering techniques in order to isolate the guitar signals.

## **Aims**

- A deep learning model such as a convolutional neural network that will take a set of audio files in a certain format.
- Create a neural network model for the classification of the guitar tuning for the corresponding audio file. Even if the audio file has unwanted audio, for example other instruments, voices.
- The model will be able to classify a balanced distributed data set of audio files to a certain degree of accuracy.

## Objectives

- Create a system that will take audio files and split it up based on a given time step to get short snippets of the respective sample to be processed.
- The neural network will classify an audio file by outputting a string representing the tuning for that file
- Apply transformations to audio to map the signal from amplitude time space to frequency-time space to get their respective spectrogram as input for the neural network.
- Investigate model parameters, hyperparameters and best classifier approach for our neural network.
- Investigate testing techniques to evaluate the effectiveness of the convolutional neural network.

## Work-plan:

### **Week 1 (27th Jan-2nd Feb):**

Acquire data set, software requirements, initial meeting, initial plan.

- For the first week I will acquire audio files that contain music with acoustic guitar, the primary data set that I will be using for this is the Joni Mitchell songbook. This has known tunings of each song and other details of describing the musical structures of her songs. I picked this dataset as it has ideally a low number of other instruments/voices and is mainly acoustic plus voice songs.
- I will also set up the api and other libraries needed to create the neural network, handle audio files and other processing of data such as Tensorflow, keras, librosa. I also will work on my initial plan and discuss projects aims and the plan with my supervisor during our meeting and over email. I will also do some research on CNN's to grasp a general idea of how I am going to complete this project.

### **Week 2-4 (3rd Feb-9th Feb):**

Submit initial plan, research on other convolutional neural networks and different classifier methods, tutorials on CNN's, sample large data set into a subset of it so we have a balanced distributed data set to be used as my training data set.

- During this week I will be focusing on researching other Convolutional neural networks that have been used in audio classification such as audioNet, wavNet and Anthem score (used to transcribe piano music). There is also a tutorial video of constructing a CNN in Tensorflow using keras environment for audio classification of different sounds. This will give me knowledge of the methods required to construct a CNN for basic classification of audio and therefore will allow me to modify and adapt the model to classification of guitar tunings. This will also give me knowledge on how to divide the frequencies into normalized bins for our input data to the neural network and transformation methods used on audio, for example SIFT, dynamic Q transforms.

- I will also not use the whole of Joni Mitchell songbook but a subset of the songbook containing n number of songs with similar length of time for each tuning that the system will be able to classify. Selection can be done via random sampling or other sampling methods to get a balanced distributed data set.
- Finalise and submit initial plan.
- Research classifier approaches used in preexisting CNN's for a similar domain to find out classifier methodologies best suited for this project. Also, research how they do their feature extraction method and processing to grasp an idea of how to apply it to this project.

#### **Week 4 (10th Feb-16th Feb):**

Process data in the right format for neural network i.e. transformations, sample size, provide class labels for our data set.

- During this week I will devise a method to segment audio recordings into snippets based on a time step interval for the sample size. I will explore different sample sizes to find the most optimal size that is easy for the neural network to handle.
- I will find the best transformation method to use to get the respective values going to be used to produce and analyze the spectrograms of the audio samples in order for the system to learn how to classify different guitar tunings. It will do this through some feature extraction method to get the feature vectors in the spectrograms representing the tuning of the guitar.
- Setting class labels for our training data set.
- Finish the CNN tutorial on audio classification.

#### **Week 5 (17th Feb- 23rd Feb):**

Use filtering techniques to remove unwanted audio in the audio files of the data set

- Research and test with different filtering techniques on audio before given as input to the CNN.

#### **Week 6 (24th Feb - 1st March):**

Modify Convolutional neural network using tutorials learnt to classify guitar tunings, software environment is tensorflow using keras library, adapt network from tutorial to our problem domain.

- Use CNN created from tutorial and other preconstructed models used in similar domains of classification to adapt to our problem domain by modification of the Neural network.

#### **Week 7-8 (2nd March-15th March):**

Testing different classifier approaches with our neural network and complex audio. I need to investigate different class models, this would include a classifier for each tuning, or a classifier for each string. I will initially focus on the first method as it is easier to implement.

- Once a classifier model is picked and trained with simple training data can test our system on complex audio, this will be audio with more instruments or voices in them.

#### **Week 9-12 (16th March-12th April):**

Tuning and Training and testing-so rest of time is used to tune the neural network parameters and train it on more data to make effective and testing with cross validation.

- Tune Neural network and optimize its model parameters and hyperparameters.
- Can use data augmentation on pre-existing training data set to get a new dataset to train with.
- Use system on a data set that has a balanced distribution with combination of simple and complex audio, record results, evaluate the model performance using cross-validation on the data set.

**Week 13-15 (13th April-3rd May):**

Finalize final year report.

Throughout this project, during every week of implementation I will write drafts of the corresponding sections in my report for those methods implemented.

**Meetings:**

I have arranged regular meetings with my project supervisor each week on Tuesday at 11am. The two scheduled review meetings have been arranged on 3<sup>rd</sup> march and 3<sup>rd</sup> may, although these dates may be rearranged with my supervisor later.