Guitar Tuning Detection

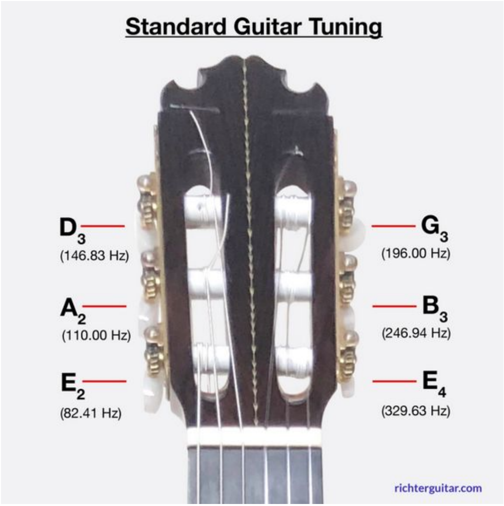
McCormack Chew
Declan Parker

Guitar Tuning Background

- Using alternate tunings involves changing which notes the open strings on a guitar are tuned to

- It's essential for a guitarist to know before learning a song

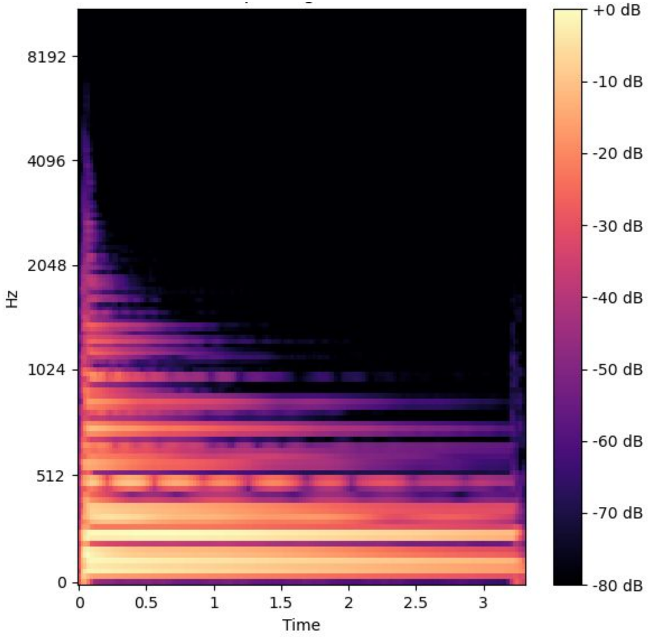
String	6	5	4	3	2	1
E Standard	E	A	D	G	B	E
Drop D	D	A	D	G	B	E
Drop C	C	G	C	F	A	D
D Standard	D	G	C	F	A	D
Open D	D	A	D	F#	A	D
Open G	D	G	D	G	B	D
F Maj 9	F	A	C	G	C	E



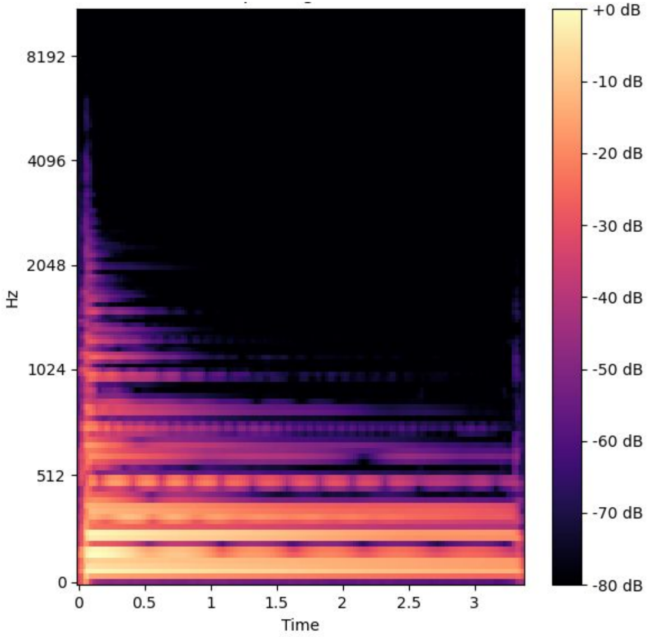
	E Standard	Drop C
Notes		
E Chord		

Nuances in Spectral Content Between Different Tunings

E2 Played in E Standard



E2 Played in Drop C



Goals

- Determine if a CNN and shallow classifier model is capable of distinguishing between different tunings using audio alone
  - We looked at the 7 different tunings
- Explore what might make tunings distinguishable
- Find ways to maximize the accuracy of the classifier
- Improve upon previous work, which found similar approach was promising with a limited dataset

String	6	5	4	3	2	1
E Standard	E	A	D	G	B	E
Drop D	D	A	D	G	B	E
Drop C	C	G	C	F	A	D
D Standard	D	G	C	F	A	D
Open D	D	A	D	F#	A	D
Open G	D	G	D	G	B	D
F Maj 9	F	A	C	G	C	E

Dataset Collection

- No suitable dataset for our project existed, so we had to create our own
- All audio used contains only guitar as not to confuse the model in training
- Audio across all tunings contain clean tones, distortion, and other effects
- The dataset includes:
    - 96 ten-second clips of solo guitar, sourced from YouTube videos
    - 164 ten-second clips of stem-separated guitar, sourced from songs of many genres
- Clips were manually selected in an effort to make sure they were characteristic of their tuning

Model

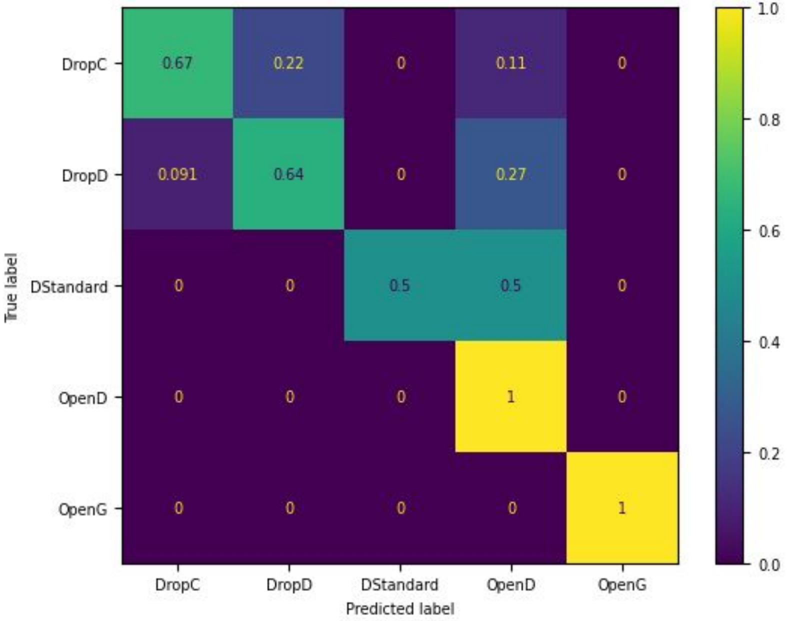
- Generate feature embeddings with pre-trained OpenL3 model (256 Band Mel Spectrogram, embedding dimensionality of 6144, .5 second hop, 1 second window)

- Classify each frame of embeddings with Support Vector Machine (polynomial kernel)

- We can train multiple classifiers based on lowest detected pitch

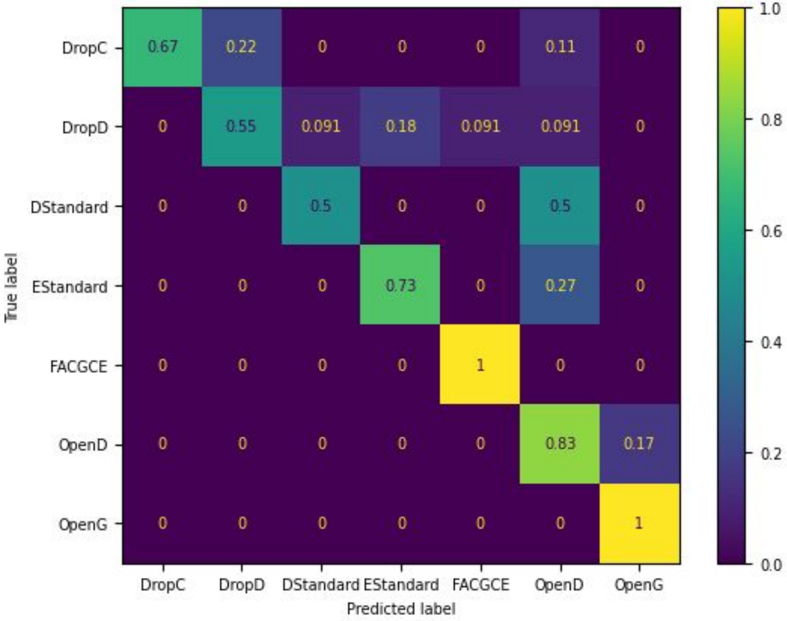
Results

Acc = 77.5%



Confusion Matrix of Classifier With Pitches Below E2

Acc = 71.7%



Confusion Matrix of Classifier With All Pitches

Conclusions

- Achieved moderate success, but may be hitting limits of OpenL3
- Our model was trained on short frames of audio, which may introduce error, especially with single notes
- Differentiating between closely related tunings (multiple strings tuning to same pitch) might require music theory and/or genre information
  - Open Tunings: predominantly major chords and Folk/Indie genres
  - Drop Tunings: predominantly power chords and Rock/Metal genres

References

[1] D. Roy, "Deep Learning Guitar Tunings," in *One Semester Individual Project.*, Cardiff, Wales. May, 2019.

[2] https://www.ultimate-guitar.com/ (accessed Dec. 1, 2023).

[3] https://gtdb.org/ (accessed Dec. 1, 2023).

[4] Kim, J. W., Salamon, J., Li, P., and Bello, J. P., "CREPE: A Convolutional Representation for Pitch Estimation", <i>arXiv e-prints</i>, 2018. doi:10.48550/arXiv.1802.06182.

[5] Aurora Cramer, Ho-Hsiang Wu, Justin Salamon, and Juan Pablo Bello. "Look, Listen and Learn More: Design Choices for Deep Audio Embeddings", IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), pages 3852–3856, Brighton, UK, May 2019.

[6] F. Gemmeke, D. P. W. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. C. Moore, M. Plakal, and M. Ritter, "Audioset: An ontology and human-labeled dataset for audio events," in IEEE ICASSP, 2017, pp. 776–780.