Acoustic Classification of Guitar Tunings with Deep Learning

Edward Hulme David Marshall Kirill Sidorov Andrew Jones

11th International Conference on Digital Libraries for Musicology

School of Computer Science & Informatics
Cardiff University

Introduction

Research question:

- Can neural networks be used for the acoustic classification of guitar tunings?

Definition:

- *Guitar tuning classification* (GTC) is the identification of a particular guitar tuning from a recording that contains a guitar performance.

Motivation

- Provide methods that facilitate the transcription of a vast corpus of non-notated guitar recordings

Altered Tuning Example

In the maskanda music of South Africa the tuning used by guitarists often varies from standard tuning; the high string is tuned to D_4 instead of E_4, and other tunings exist, “some pertaining to specific styles and others ‘invented’ by musicians to suit their individual characteristic styles” (Davies, 1994, p.122)

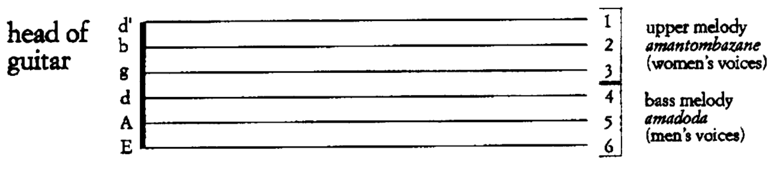
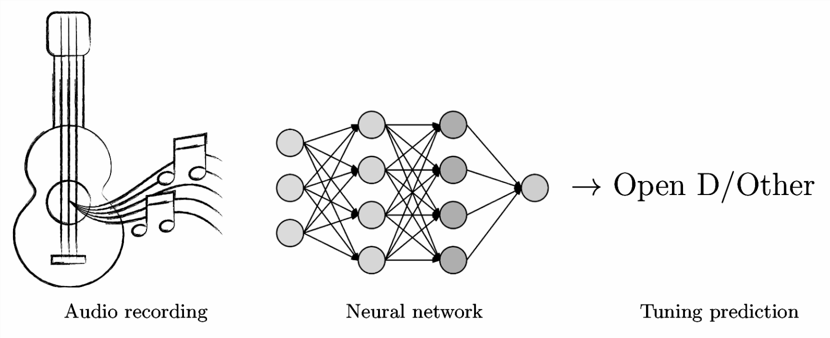


Figure 1: Common maskanda guitar tuning (Davies, 1994, p.121)

Task and Goals



Challenges

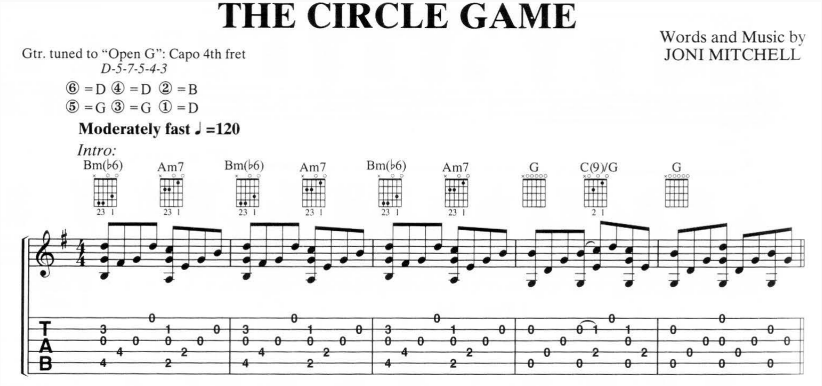


Figure 2: Guitar notation (Bernstein & Libertino, 1996)

Related Work

- Guitar tuning classification research is underdeveloped
- There is a wide body of research devoted to closely related tasks:
    - AMT (Benetos, Dixon, Duan, & Ewert, 2019)
    - Chord recognition (Barbancho, Klapuri, Tardon, & Barbancho, 2012)
    - String detection (Dittmar, Männchen, & Abeber, 2013)
    - Genre classification (Müller & Klapuri, 2014)

Guitar Characteristics

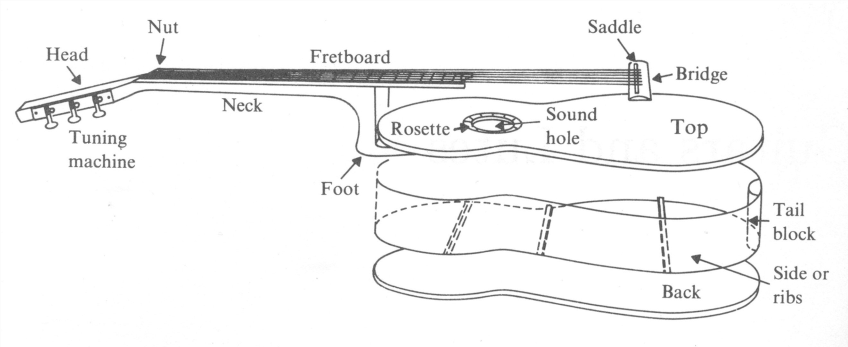


Figure 3: An exploded view of the guitar (Rossing, 1982)

Pitch Class, Pitch Range, Scale, and Key

Common guitar tunings (6th string to 1st string):

- Standard $E_2, A_2, D_3, G_3, B_3, E_4$
- Drop D $D_2, A_2, D_3, G_3, B_3, E_4$
- Open D $D_2, A_2, D_3, F\sharp_3, A_3, D_4$
- Open G $D_2, G_2, D_3, G_3, B_3, D_4$
- DADGAD $D_2, A_2, D_3, G_3, A_3, D_4$
- E♭ Standard $E\flat_2, A\flat_2, D\flat_3, G\flat_3, B\flat_3, E\flat_4$

Chord Type and Chord Voicing

Table 1: Open D chord in standard tuning and open D tuning.

Tuning	String Number (x = *no note*)					
	6th	5th	4th	3rd	2nd	1st
Standard	x	x	D_3	A_3	D_4	$F\sharp_4$
Open D	D_2	A_2	D_3	$F\sharp_3$	A_3	D_4

Harmonic Spectrum

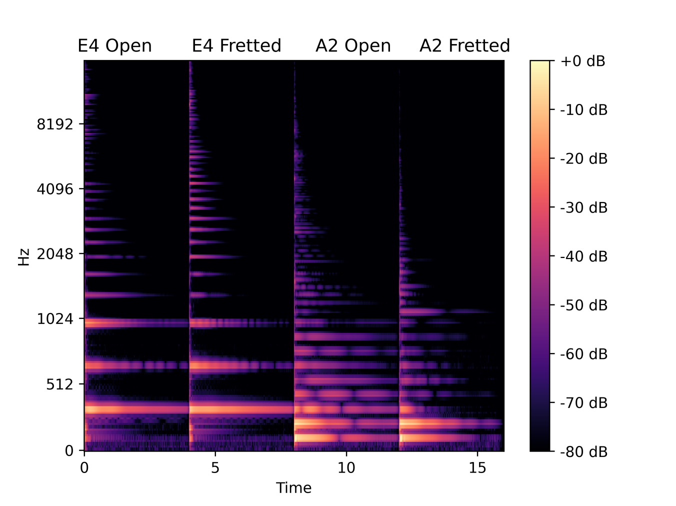


Figure 4: Log mel spectrogram representation of a guitar recording

Inharmonicity

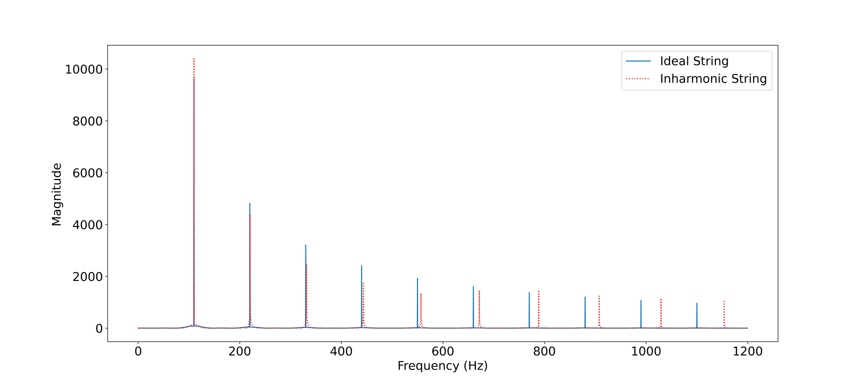


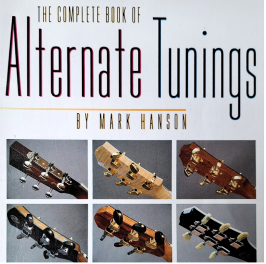
Figure 5: Spectrum plot of synthesised 'ideal' and inharmonic strings: $f_1 = 110$ Hz (i.e., A_2)

Method

To create a guitar tuning classification system we:

1. Compile an annotated audio dataset
2. Convert the audio samples into a suitable input representation
3. Train the model on the corpus of labelled samples
4. Evaluate the predictive performance of the trained model

Authentic Data



Source Separation

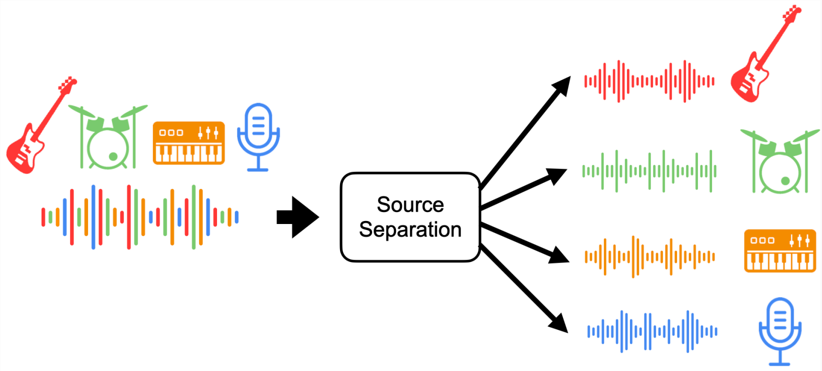


Figure 6: Music source separation (Manilow, Seetharaman, & Salamon, 2020)

Synthetic Data

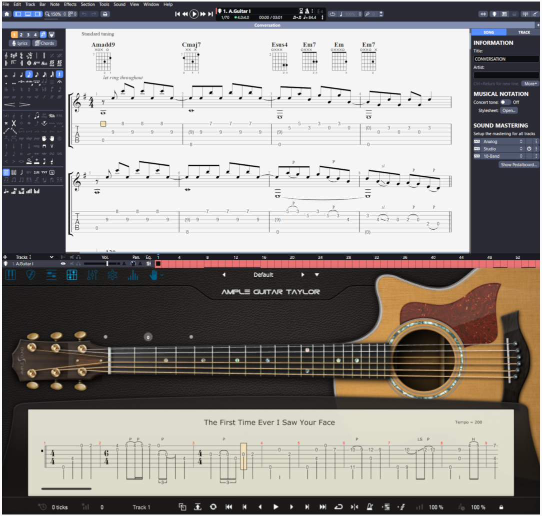


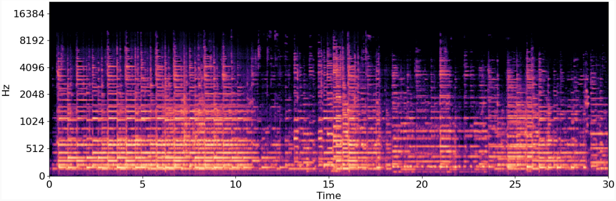
Figure 7: Software used to generate guitar audio

Labels

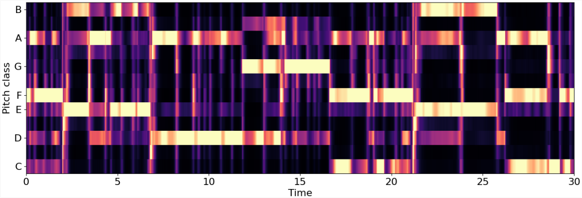
- Recordings assigned tuning/capo labels derived from:
  - *The Joni Mitchell Complete—Guitar Songbook Edition*
  - Official GuitarPro tablatures
- Tuning label examples:
  - Relative tuning: *x75435*
  - Absolute tuning: *EBEG#BE*
  - Capo position: *C2*

Input Representations

- Log mel spectrogram:



- Chromagram:



Model Selection

Two convolutional neural network (CNN) architectures were selected for the task:

- *CNN 1*:
  - Adapted from a bird audio detection system
  - Used in 2-class experiments
  - Trained on Joni Mitchell recordings
  - Log mel spectrogram input
- *CNN 2*:
  - Adapted from a keyword spotting system
  - Used in 5-class experiments
  - Trained on synthetic guitar audio
  - Chromagram input

Open D/Other: Sample Length Study

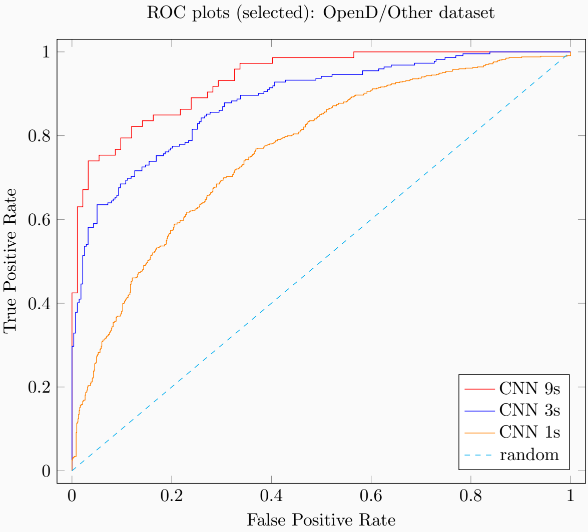
- CNN 1 trained on 29 Joni Mitchell recordings
- Spectrogram input representation
- Model predicts tuning class: *open D/other*
- Relative tuning predictions as audio contains capo/downtuning
- Model trained/tested with different clip lengths: 1s, 3s, 9s
- Evaluation on an unseen test set of 8 Mitchell songs

Results—Open D/Other: Sample Length Study

Table 2: AUC and ACC for Open D/Other Models

Model	AUC		ACC Clip		ACC Song	
	Mean	SD	Mean	SD	Mean	SD
9s	0.893	(0.03)	81.8%	(2.30)	**97.5%**	(5.00)
3s	0.823	(0.06)	74.2%	(5.37)	**87.5%**	(13.69)
1s	0.627	(0.12)	57.4%	(10.15)	**52.5%**	(14.58)

ROC Curve—Open D/Other: Sample Length Study



Open D/Open G Study

- CNN 1 trained on 19 Mitchell recordings
- Spectrogram input representation
- Model predicts tuning class: *open D/open G*
- Relative tuning predictions
- Evaluation on an unseen test set of 4 Mitchell recordings
- Evaluation on an independent test set of 8 recordings by different artists

Results—Open D/Open G Study

Table 3: AUC and ACC scores for Open D/Open G Models

Test Data	AUC		ACC Clip		ACC Song	
	Mean	SD	Mean	SD	Mean	SD
Mitchell	0.867	(0.05)	80.0%	(4.58)	**95.0%**	(10.0)
Multi-artist	0.577	(0.05)	57.3%	(4.30)	**65.0%**	(5.0)

Multiclass Study

- CNN 2 trained using 10 hours of synthetic guitar audio
- Chromagram input representation
- 5 tuning classes
- Absolute tuning predictions as capo/downtuning not present
- Evaluation on 2 hours of unseen synthetic guitar audio
- Evaluation on an independent test set of 46 authentic recordings by different artists

Results—Multiclass Study

Table 4: Average F-score for 5 Tuning Class Models

Test Set	F_1 Clips		F_1 Songs	
	Mean	SD	Mean	SD
Synthetic	0.649	(0.03)	**0.666**	(0.05)
Authentic	0.474	(0.06)	**0.516**	(0.10)

Table 5: Average accuracy for 5 Tuning Class Models

test set	ACC Clips		ACC Songs	
	Mean	SD	Mean	SD
Synthetic	65.4%	(0.03)	**67.8%**	(0.05)
Authentic	43.4%	(0.07)	**48.3%**	(0.11)

Results—Multiclass Model

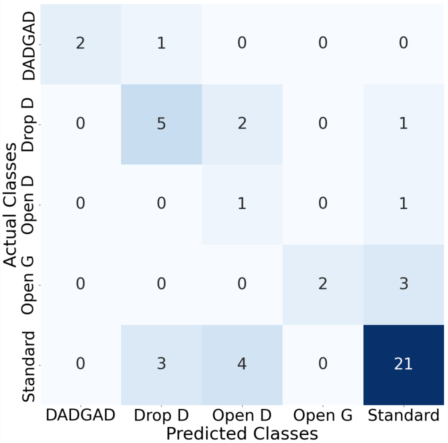


Figure 8: Predictions on an independent test set of real songs

Conclusion

In this work we:

- Provided evidence that deep learning can be used for the acoustic classification of guitar tunings
- Identified features of guitar audio that could be utilised by neural networks for GTC
- Created custom authentic and synthetic audio datasets for guitar tuning tasks
- Proposed dataset collection, generation and annotation methods

Future Work

- Investigate capo and open string detection
- Release a large synthetic audio dataset for GTC
- Create a robust *standard tuning* detection algorithm
- Develop a model that outputs a separate tuning prediction for each guitar string

Information

Thanks! Get more information about this research from:

`github.com/edhulme/guitar-tuning-classification`

References

Barbancho, A. M., Klapuri, A., Tardon, L. J., & Barbancho, I. (2012). Automatic transcription of guitar chords and fingering from audio. *IEEE Transactions on Audio, Speech, and Language Processing*, *20*(3), 915-921. doi: 10.1109/TASL.2011.2174227

Benetos, E., Dixon, S., Duan, Z., & Ewert, S. (2019). Automatic music transcription: An overview. *IEEE Signal Processing Magazine*, *36*(1), 20-30. doi: 10.1109/MSP.2018.2869928

Bernstein, J., & Libertino, D. (1996). *Joni mitchell complete guitar songbook edition*. Los Angeles, CA, USA: Alfred Publishing Co., Inc.

Davies, N. (1994). The guitar in zulu “maskanda” tradition. *The World of Music*, *36*(2), 118–137. Retrieved 2024-01-23, from http://www.jstor.org/stable/43561390

Dittmar, C., Männchen, A., & Abeber, J. (2013). Real-time guitar string detection for music education software. In *2013 14th international workshop on image analysis for multimedia interactive services (wiamis)* (p. 1-4). doi: 10.1109/WIAMIS.2013.6616120

Manilow, E., Seetharaman, P., & Salamon, J. (2020). *Source separation.* Retrieved from https://source-separation.github.io/tutorial/data/introduction.html

Müller, M., & Klapuri, T. A. (2014). *Automatic transcription of bass guitar tracks applied for music genre classification and sound synthesis* (Unpublished doctoral dissertation). Gustav-Kirchhoff-Straße 1, 98693 Ilmenau, Germany.

Rossing, T. D. (1982). *The science of sound*. Reading, Massachusetts: Addison-Wesley.