Team #16: From Strings to Sequences — Classifying and Generating Music from Acoustic Guitar Notes

Camilo Martínez 7057573 | Dhimitrios Duka 7059153 | Honglu Ma 7055053

1. Task and Motivation

Automatic cord recognition (ACR) consists of recognizing the chords being played in a music piece. This information is quite valuable since it can later be used for music analysis, music transcription and other similar tasks. ACR was first introduces in 1999 by [8] where the author utilized lisp music to perform chord recognition at the signal level. Since then, many different approaches have been introduced []. However, these methods approached the task from an audio signal processing prespective which proved to be quite challenging and not very accurate.

However, to our knowledge, there is no research that tackles the problem of recunstruction the audio being played in a video.

Task statement and definitions

Motivation: Why do we need to explore this task?

Related work: How do existing papers solve this task or similar tasks (should include relevant citations)?

Challenge: What are the major challenges that have not been solved in this task?

[6] [4]

2. Goals

What challenges do you aim to address in this task?

What do you want to have completed by the mid-term? E.g., code for the task, data collection, results for baselines, etc.

3. Methods

What models/frameworks do you use to solve the challenges?

Why can the proposed method / analysis solve your problem?

What are the main differences between your method and existing methods (if applicable)?

What is the required computational budget for the training/analysis? (E.g. are you planning on using pretrained backbones?)

4. Datasets

What datasets are you going to use/collect and why?

4.1. Fretboard Detection

This one has both: [5]. Only fretboard: [7], [3].

4.2. Chord Recognition

For chord recognition: [9], [1], [2].

5. Evaluation

How is your method going to be evaluated?
What metrics are suitable?
Do you need to define your own metric for evaluation?

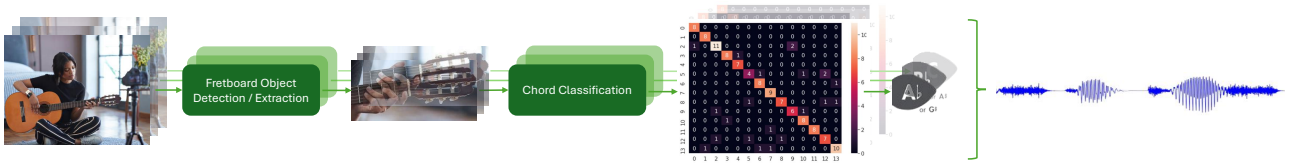

Figure 1. Example figure.

References

[1] GRC MASK custom dataset. Guitar chord bounding box dataset. https://universe.roboflow.com/grc-mask-custom-dataset/guitar-chord-bounding-box, jun 2024. visited on 2024-06-29. 1

[2] GRC MASK custom dataset. Guitar chord handshape dataset. https://universe.roboflow.com/grc-mask-custom-dataset/guitar-chord-handshape, apr 2024. visited on 2024-06-29. 1

[3] done. done dataset. https://universe.roboflow.com/done-ygt9y/done-npcll, apr 2024. visited on 2024-06-29. 1

[4] Yuexi Du, Ziyang Chen, Justin Salamon, Bryan Russell, and Andrew Owens. Conditional generation of audio from video via foley analogies. In *Conference on Computer Vision and Pattern Recognition 2023*, 2023. 1

[5] joaomarcoscrs. Guitar chords dataset. https://universe.roboflow.com/joaomarcoscrs/guitar-chords-daewp, jun 2024. visited on 2024-06-29. 1

[6] Yosi Kristian, Lukman Zaman, Michael Tenoyo, and Andreas Jodhinata. Advancing guitar chord recognition: A visual method based on deep convolutional neural networks and deep transfer learning. *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, 18(2):235–249, May 2024. 1

[7] SOEN357. guitar dataset. https://universe.roboflow.com/soen357/guitar-ppfil, may 2024. visited on 2024-06-29. 1

[8] Fujishima Takuya. Realtime chord recognition of musical sound: Asystem using common lisp music. In *Proceedings of the International Computer Music Conference 1999, Beijing*, 1999. 1

[9] My Work. Guitar chord dataset. https://universe.roboflow.com/my-work-3idwy/guitar-chord-tvon8, may 2024. visited on 2024-06-29. 1