

**Team Name**

MLPD

**Project Title**

Amadeus Ex Machina - Deep Learning Chord Transcription

**Project Summary**

Chord and chord progression classification is a task where traditional DSP-based approaches are typically lackluster in performance, presenting an opportunity for utilizing deep learning neural networks. Within the deep learning sphere, the task has not received as much attention relative to more big ticket items such as image classification and speech recognition, and presents room for novel contributions. We therefore endeavor to conduct a multi-faceted exploration into the application of deep learning techniques in the context of chord classification. Our hope is to potentially uncover previously unpublished insight in this application.

**Approach**

Our project will focus on the following areas of exploration:

1. **Learning Bias:** Inspired by previous research by Geirhos et al. on texture bias in CNNs [5], we will examine if such bias for timbre exists in the audio domain. Additionally, we will explore the effect of training on various transformed versions of the dataset on model performance and generalization.
2. **Transfer Learning:** We will explore if training the model on standalone chords (with transformed variants as appropriate) allows for transfer learning in classifying chord progressions.
3. **Model Architecture Exploration:** As a stretch goal, we will also experiment with different architectures from both within and outside the audio domain. We can extend this further by exploring integration of audio domain techniques [6] (Chroma, FFTs, DSPs, etc.) in our model and their effects on both accuracy and computational performance (which has implications in real-time applications).

**Resources & Related Work**

Work in this exact field of study is relatively sparse outside of proprietary algorithms developed by music software companies [4]. As such, we turn to adjacent areas of study related to high-input deep learning (especially convolutional)[1][3], as well as the field of audio processing [5] and audio deep learning in general [2].

1. "A Fully Convolutional Deep Auditory Model for Musical Chord Recognition", Korzeniowski and Widmer
2. "Neural Networks for Musical Chords Recognition", Osmalskyi et al.
3. "Chord Detection using Deep Learning", Zhou and Lerch
4. "Chordify: Chord Transcription for the Masses", Haas et al.

5. “ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness”, Geirhos et al.
6. “Deep Learning for Audio Signal Processing”, Purwins et al.

### **Datasets\***

We plan to generate our datasets via automation. Musical transcription data will be parsed from transcription software files available from various subscription-based websites\*\*. The accompanying MIDI tracks from these files will provide the basis for modifiable synthetic versions of the audio dataset, with song title information used to fetch original recorded versions from various platforms.

The standalone chord dataset will be generated using a combination of MIDI and VST instrument libraries. Transformations on audio data will be performed en masse using various scriptable audio plug-ins.

### **Team Members\***

Jun Zhu  
Dana Tareq Alnabulsi  
Maxwell Steven Herron  
Matthew David Bisson  
Edward Hunter Mortimer-Tsushima

### **\*NOTE**

Approval for 5 members and generating our own dataset was granted in this ed post:  
<https://edstem.org/us/courses/60909/discussion/5194465>

\*\*Tool for this was previously developed by one of our members for an unrelated application.