Guitar Tab Generation From Audio

Ears to Frets: Crafting Guitar Tabs from Audio Inputs

Ashhwath Chandramohan
Nandhini Devaraj
Austin McCall

Unveiling the Future of Music

Welcome to a Harmonic Revolution!

Imagine a world where the beautiful complexities of guitar music are instantly transcribed from mere soundwaves into accurate, playable guitar tabs.

Today, we stand on the brink of making this a reality, merging the realms of audio processing and musical expression through advanced technology.

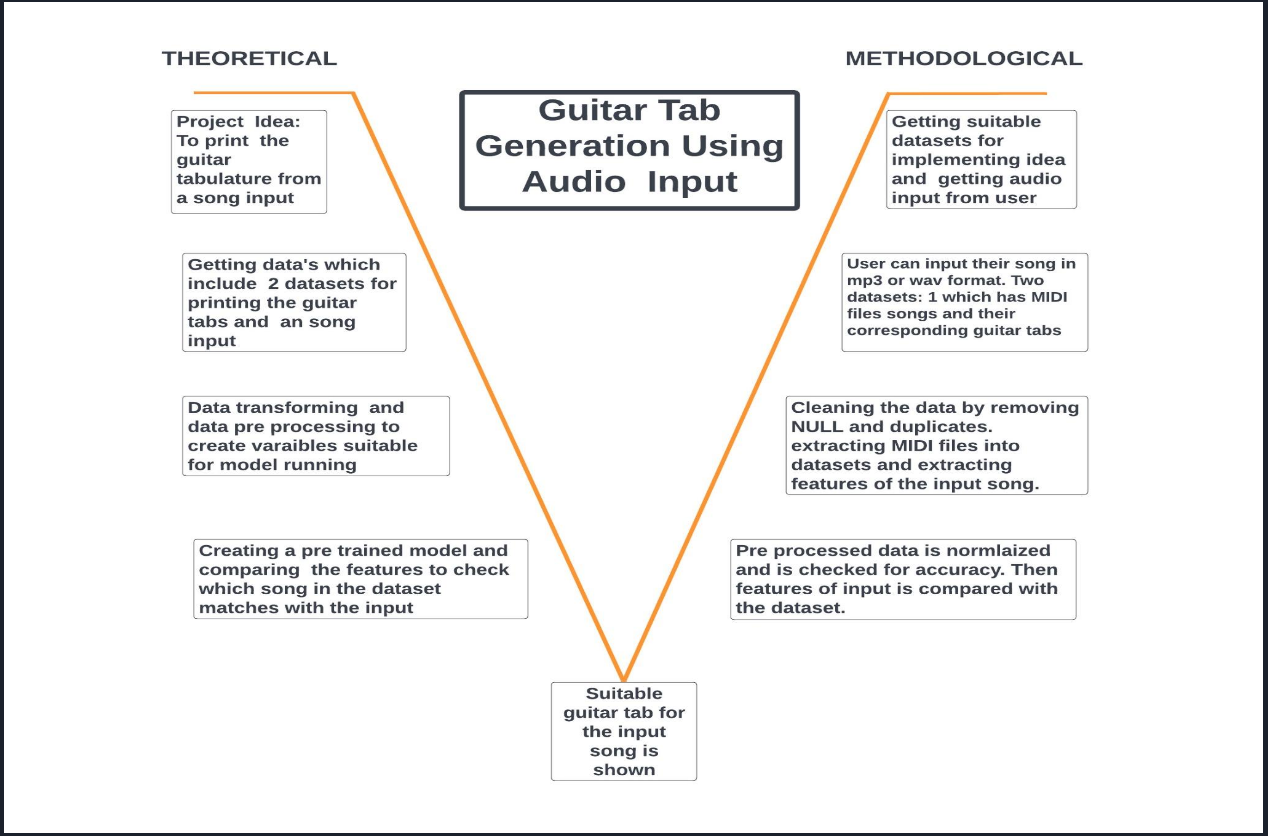
Project Overview

Automatic Guitar Tab Generation from Audio

Goal: Develop a system that accurately transforms audio recordings of a song into guitar tablature.

Purpose: Enable musicians of all skill levels to learn and play songs more easily by providing an automated tool to transcribe music.

Process Flow



Datasets

Input song features

```
input_features

[76]:  [62.96210268948655,
        0.2656684294926658,
        51.35085574572127,
        651.9503259750018,
        [0.0,
         0.16666675,
         0.19444454166666666,
         0.17222230833333335,
         0.650000325,
         0.750000375,
         0.816667075,
         0.9611115916666667,
         0.9277782416666667,
         0.9388893583333333,
         0.9388893583333333,
         2.5111123666666666,
         2.5111123666666666,
```

MIDI Dataset

[91]:

	Name	features
0	GlasgowKiss	[{'pitch': 4, 'start': 0.0, 'end': 0.14, 'velo...
1	Swear	[{'pitch': 1, 'start': 2.82, 'end': 3.18, 'vel...
2	HowDoYouKeepTheMusicPlaying	[{'pitch': 3, 'start': 0.43, 'end': 0.82, 'vel...
3	MessageOfLove	[{'pitch': 7, 'start': 1.71, 'end': 5.12, 'vel...
4	Spiderman	[{'pitch': 0, 'start': 2.58, 'end': 2.74, 'vel...
...	...	...
788	HearOnMyOwn	[{'pitch': 4, 'start': 7.25, 'end': 7.28, 'vel...
789	SinceIDontHaveYou	[{'pitch': 5, 'start': 3.62, 'end': 3.82, 'vel...
790	AllIEverNeedIsYou	[{'pitch': 0, 'start': 0.0, 'end': 0.36, 'velo...
791	Invitation	[{'pitch': 11, 'start': 5.45, 'end': 6.87, 've...
792	ItsOver	[{'pitch': 1, 'start': 0.0, 'end': 0.04, 'velo...

793 rows × 2 columns

Guitar Tablature Dataset

[71]:

	Name	Notes	Len_Sequence	Unique_notes	len_Uni_Notes
0	BreezeAndI	['E-2', 'B-2', 'B-2', 'E-3', 'B-2', 'B-2', 'E-...	427	{'C3', 'F1', 'A2', 'B-0', 'E-1', 'E3', 'B1', '...	27
1	IfIHadYou	['10.2.5', '10.2.5', '2.5.8', '2.5.8', '0.3.7'...	204	{'7.10.2', '6.9.0.2', '9.10.0.2.4', '6.9.0', '...	37
2	IllBeSeeingYou	['E-2', 'E-2', 'B-2', 'B-2', 'G2', 'D3', 'F2',...	455	{'C3', 'F1', 'B-0', 'A2', 'E-1', 'D3', 'B1', '...	29
3	JustAGame	['B3', '6.11', 'B1', 'E4', 'B1', 'F#4', 'B1', ...	1572	{'C3', '9.2', 'A2', 'B4', 'E3', 'A3', 'D5', 'B...	39
4	Unforgettable	['G2', 'G1', 'G2', 'G1', 'G2', 'G2', 'G2', 'C#...	421	{'A0', 'C3', 'F1', 'A2', 'E-1', 'E3', 'G1', 'D...	28
...	...	...	...	...	...
813	ThatsAllGreatAmericanSongbook	['B-1', 'B-1', 'B-1', 'B-1', 'B-1', 'B-1', 'B-...	280	{'G1', 'F#2', 'G2', 'E-2', 'F1', 'C#2', 'A2', ...	13
814	Respect	['C2', 'C2', 'A1', 'G1', 'C2', 'C2', 'G1', 'F1...	437	{'G1', 'B1', 'F#2', 'E2', 'D2', 'G2', 'E-2', '...	17
815	ATimeForUs	['D3', '2.5.9', '2.5.9', 'E-3', '3.7.10', '3.7...	193	{'D3', '7.10.2', '8.0.3', 'G#2', '2.5.9', 'G2'...	15
816	TwentyFiveToMidnight	['F#4', 'A4', 'C#5', 'A4', 'A4', 'A4', 'F#4', ...	556	{'E4', 'F#5', 'A4', 'F5', 'G#5', 'F#4', 'E-5',...	14
817	UniversalMind	['A5', '0.5', 'F5', 'B-4', 'A4', 'B-4', 'F5', ...	2981	{'E6', '6.7.11.2', 'C#6', '0.4.7', 'E3', 'B5',...	89

818 rows × 5 columns

Pre-processed Data

	Name	features	Notes	Len_Sequence	Unique_notes	len_Uni_Notes	extracted_features	average_pitch	average_duration
0	GlasgowKiss	[{'pitch': 4, 'start': 0.0, 'end': 0.14, 'velo...	['E3', '11.4', 'B3', 'E4', 'F#4', 'E4', 'B3', ...	4993	{'4.6.11', 'E6', '6.8', '3.4', '1.6', '9.2', '...	66	{'average_pitch': 6.197090310183914, 'average_...	6.197090	0.302015
1	Swear	[{'pitch': 1, 'start': 2.82, 'end': 3.18, 'vel...	['6.9.1', '6.9.1', '6.9.1', '9.1.4', '6.9.1', ...	785	{'1', '4.8.11', '4.5.9.11', '0.2.5.8', '8.9.1....	72	{'average_pitch': 5.598831427150671, 'average_...	5.598831	0.347427
2	HowDoYouKeepTheMusicPlaying	[{'pitch': 3, 'start': 0.43, 'end': 0.82, 'vel...	['E-5', 'D5', 'C5', 'D5', 'E-5', 'C5', 'B-4', ...	387	{'E6', 'A6', '9.2', '7.10', 'A3', 'D5', 'F6', ...	32	{'average_pitch': 5.315254237288135, 'average_...	5.315254	0.653247
3	MessageOfLove	[{'pitch': 7, 'start': 1.71, 'end': 5.12, 'vel...	['B-2', 'B-2', 'B-2', 'C2', 'C2', 'E2', 'C2', ...	2427	{'0.4.5', '4.7.11', 'C3', 'A2', '0.4.7', 'B4',...	43	{'average_pitch': 4.675894799288115, 'average_...	4.675895	0.441422
4	Spiderman	[{'pitch': 0, 'start': 2.58, 'end': 2.74, 'vel...	['C2', 'G#2', 'C3', 'C2', 'A2', 'G#2', 'G2', '...	1687	{'9.0.1.2.5', '8.10.3', '9', 'B-0', '6', '1.4....	202	{'average_pitch': 4.961133069828722, 'average_...	4.961133	0.215471
...	...	...	...	...	...	...	...	...	...

Challenges & Solutions

Tackling Technical Hurdles

- **File format complexity:** The goal was to convert MP3 files to MIDI, but dealing with MIDI files proved challenging due to their large size and complexity.
- **Library compatibility:** Ensuring library compatibility was crucial, requiring installation outside of Jupyter Notebook.
- **Data pre processing:** Data preprocessing involved extensive cleaning of the dataset, which contained numerous strings that needed removal.
- **Model Execution:** Implementing the transformer model posed another hurdle, as it didn't accept arrays or dictionaries as input.

Technology Overview

Leveraging MIDI and Transformer Models

MIDI: Essential for Music Analysis

The project leverages MIDI (Musical Instrument Digital Interface) technology to extract and process musical data from audio, transforming it into a format that's easier to handle computationally. By using MIDI, the system accurately determines pitch, timing, and note velocity, which are crucial for converting audio accurately into guitar tabs.

Transformer Model: Advanced Machine Learning

The Transformer model utilizes multiple layers of attention mechanisms to adeptly handle sequences and interpret complex musical nuances, including overlapping notes and harmonics. It is trained on extensive datasets of MIDI-encoded music, enabling it to accurately predict guitar tabs from diverse musical patterns in new audio inputs.

MODEL OVERVIEW

Core: Leveraged a Transformer-based neural network optimized for decoding complex sequences of musical notes and associated numerical features.

Input Processing: Utilizes a Tokenizer to convert guitar notes into numerical sequences, enhancing data interpretation.

Feature Integration: Merges tokenized sequences with scaled numerical features like duration and velocity for comprehensive training input.

Transformer Encoder Blocks: Features multi-head attention for focus across sequences, coupled with feed-forward networks for refined processing and layers for normalization and dropout to ensure stability and prevent overfitting.

Classification Approach: Utilizes a softmax activation layer for classifying sequences into musical pitches.

RESULTS

```
Epoch 1/10
16/16 [==============================] - 13s 181ms/step - loss: 0.7706 - accuracy: 0.6824 - val_loss: 0.7274 - val_accuracy: 0.7008
Epoch 2/10
16/16 [==============================] - 2s 97ms/step - loss: 0.6067 - accuracy: 0.7515 - val_loss: 0.7051 - val_accuracy: 0.6929
Epoch 3/10
16/16 [==============================] - 2s 117ms/step - loss: 0.5257 - accuracy: 0.7692 - val_loss: 0.7127 - val_accuracy: 0.7008
Epoch 4/10
16/16 [==============================] - 1s 86ms/step - loss: 0.5033 - accuracy: 0.7850 - val_loss: 0.7398 - val_accuracy: 0.7087
Epoch 5/10
16/16 [==============================] - 2s 133ms/step - loss: 0.4866 - accuracy: 0.7732 - val_loss: 0.7970 - val_accuracy: 0.7008
Epoch 6/10
16/16 [==============================] - 2s 109ms/step - loss: 0.4626 - accuracy: 0.7988 - val_loss: 0.7796 - val_accuracy: 0.6929
Epoch 7/10
16/16 [==============================] - 1s 52ms/step - loss: 0.4391 - accuracy: 0.8126 - val_loss: 0.7702 - val_accuracy: 0.6929
Epoch 8/10
16/16 [==============================] - 2s 130ms/step - loss: 0.3821 - accuracy: 0.8422 - val_loss: 0.7790 - val_accuracy: 0.6693
Epoch 9/10
16/16 [==============================] - 1s 90ms/step - loss: 0.3672 - accuracy: 0.8501 - val_loss: 0.8682 - val_accuracy: 0.6693
Epoch 10/10
16/16 [==============================] - 1s 96ms/step - loss: 0.3548 - accuracy: 0.8501 - val_loss: 0.9155 - val_accuracy: 0.6693
5/5 [==============================] - 0s 6ms/step
              precision    recall  f1-score   support

     Class 0       0.32      0.16      0.21        38
     Class 1       0.75      0.89      0.81       118
     Class 2       0.00      0.00      0.00         3

    accuracy                           0.70       159
   macro avg       0.36      0.35      0.34       159
weighted avg       0.63      0.70      0.65       159
```

The classification report and training logs shown in the image indicate that the model trained for 10 epochs, with improvements in accuracy and loss across both training and validation sets.

Notably, the final epoch shows a training accuracy of 85.01% and a validation accuracy of 66.93%, but the model displays varied performance across different classes, with Class 0 and Class 2 showing particularly low precision and recall.

Future Strums: Enhancements and Vision

Envisioned Enhancements

- **Real-Time Tab Generation:** Develop capabilities for instant transcription as music is played, enhancing live learning and composing experiences.
- **Expanded Instrument Support:** Extend the technology to include a variety of stringed instruments, broadening the application scope.
- **Integration with Digital Audio Workstations (DAWs):** Seamlessly connect our technology with popular music production software to streamline the music creation process.

Advanced Research Areas

- **Algorithmic Improvements:** Invest in research to refine machine learning algorithms, improving accuracy and speed in complex musical contexts.
- **Music Composition Tools:** Integrate intelligent composition features that suggest musical ideas based on user style and past inputs.

References

https://arxiv.org/pdf/2301.05295.pdf

https://project-archive.inf.ed.ac.uk/msc/20172175/msc_proj.pdf

https://towardsdatascience.com/audio-to-guitar-tab-with-deep-learning-d76e12717f81

https://github.com/Josuelmet/Metal-Music-Interpolator/tree/main

https://www.kaggle.com/code/benfitzgerald3132/sample-binary-classification-of-audio-input

https://github.com/krishnaik06/Huggingfacetransformer/blob/main/Custom_Sentiment_Analysis.ipynb

https://www.kaggle.com/datasets/saikayala/jazz-ml-ready-midi?select=Jazz-midi.csv(Dataset)

https://www.kaggle.com/code/toddpp/labeling-and-visualization-of-jazz-music-in-midi

https://www.kaggle.com/code/kerneler/pitch-detection-with-spice

Thank You