Guitar Tablature Estimation

Nick vd Linden s2971976

What is guitar tablature?

Why Automatic Guitar Transcription?

- Problems with human-made tabs

Why this is challenging

Previous work

- Gregory Burlet and Ichiro Fujinaga. Robotaba guitar tablature transcription framework. In Proceedings of the 14th International Conference on Music Information Retrieval (ISMIR), Curitiba, Brazil, 2013

- Gregory Burlet and Abram Hindle. Isolated guitar transcription using a deep belief network. PeerJ Computer Science, 3:e109, 201

- Kazuki Yazawa, Katsutoshi Itoyama, and Hiroshi Okuno. Automatic transcription of guitar tablature from audio signals in accordance with player's proficiency. In 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 3122–3126. IEEE, 201

- Kazuki Yazawa, Daichi Sakaue, Kohei Nagira, Katsutoshi Itoyama, and Hiroshi G Okuno. Audio-based guitar tablature transcription using multipitch analysis and playability constraints. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 196–200. IEEE, 201
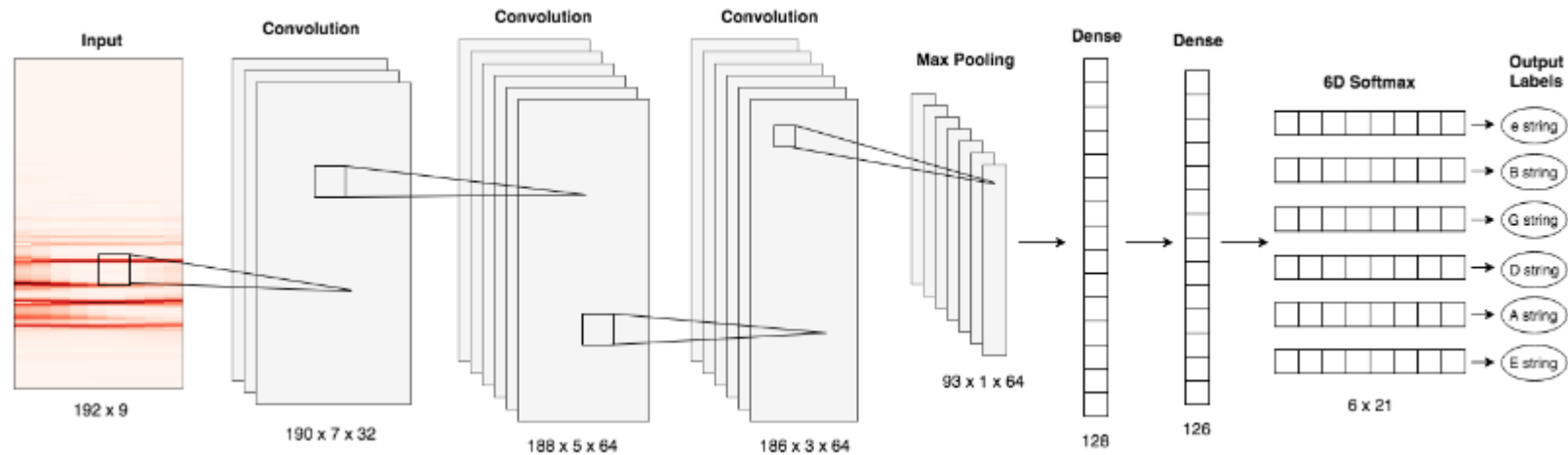
Table of Contents

- Methodology

- Evaluation

- Results

- Discussion

- Conclusion

Dataset

- Recording

- Audio pre-processing

- Label pre-processing

The network

Evaluation

- Multi pitch Metrics

$$p_{\text{pitch}} = \frac{e^T(Y_{\text{gt}} \odot Y_{\text{pred}})e}{e^T Y_{\text{pred}} e} \qquad r_{\text{pitch}} = \frac{e^T(Y_{\text{gt}} \odot Y_{\text{pred}})e}{e^T Y_{\text{gt}} e} \qquad f_{\text{pitch}} = \frac{2p_{\text{pitch}} r_{\text{pitch}}}{p_{\text{pitch}} + r_{\text{pitch}}}$$

- Tablature Metrics

$$p_{\text{tab}} = \frac{e^T(Z_{\text{gt}} \odot Z_{\text{pred}})e}{e^T Z_{\text{pred}} e} \qquad r_{\text{tab}} = \frac{e^T(Z_{\text{gt}} \odot Z_{\text{pred}})e}{e^T Z_{\text{gt}} e} \qquad f_{\text{tab}} = \frac{2p_{\text{tab}} r_{\text{tab}}}{p_{\text{tab}} + r_{\text{tab}}}$$

- TDR

$$TDR = \frac{e^T(Z_{\text{gt}} \odot Z_{\text{pred}})e}{e^T(Y_{\text{gt}} \odot Y_{\text{pred}})e}$$

Results

System	p_{pitch}	r_{pitch}	f_{pitch}
TabCNN	**0.900** $\pm$ 0.016	**0.764** $\pm$ 0.043	**0.826** $\pm$ 0.025
Deep Salience	0.778 $\pm$ 0.092	0.562 $\pm$ 0.099	0.646 $\pm$ 0.078

Table 1. Multipitch estimation metrics for our system, TabCNN, compared against a baseline, the Deep Salience f0-estimation algorithm introduced in [4], from experiments carried out in [24]. For all metrics, we report the mean and standard deviation over the entire dataset.

Results

System	p_tab	r_tab	f_tab	TDR
TabCNN	**0.809 ±** 0.029	**0.696 ±** 0.061	**0.748 ±** 0.047	**0.899 ±** 0.033

Table 2. Tablature estimation results for the proposed system, TabCNN, using the metrics we introduce to measure performance in fingering prediction. For all metrics, we report the mean and standard deviation over the entire dataset.

Discussion

- Common errors
    - False Alarms
    - Missed Detection
    - Miss Frettings
- Temporal Smoothing

Conclusion

- Future work
  - end-to-end transcription program
  - temporal smoothing