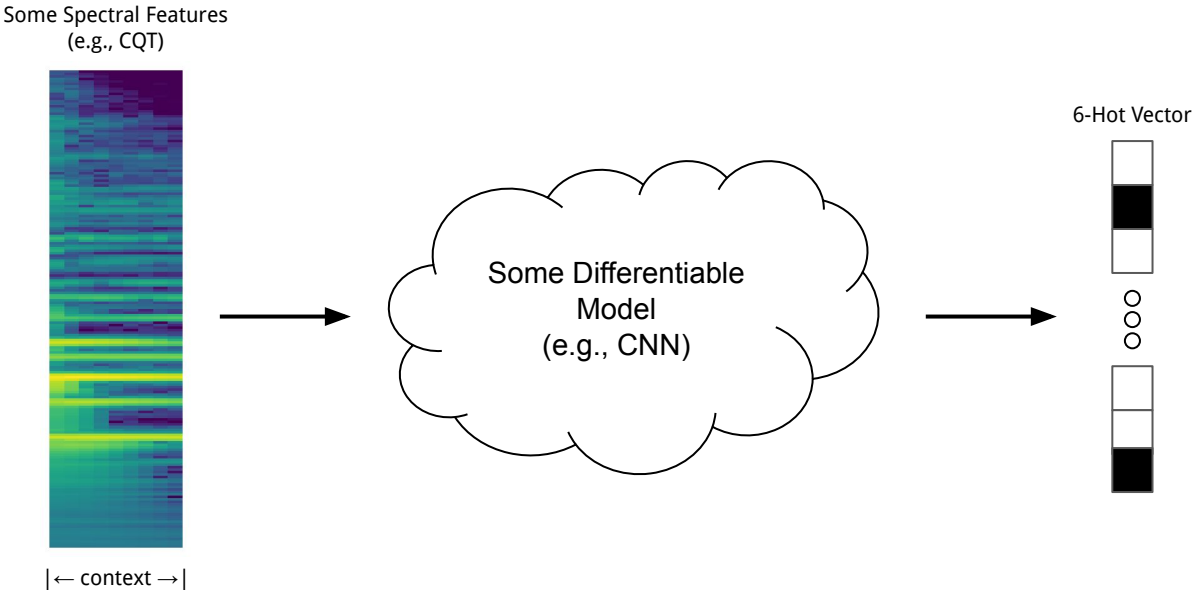
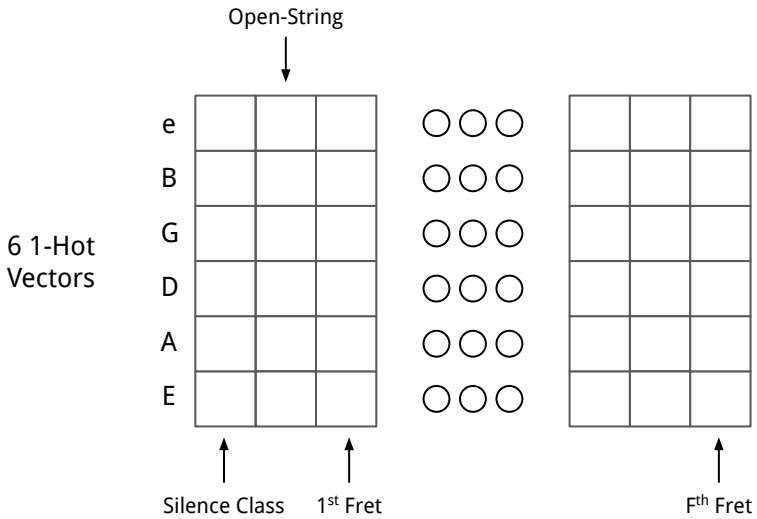
A Data-Driven Methodology for Considering Feasibility and Pairwise Likelihood in Deep Learning Based Guitar Tablature Transcription Systems

Frank Cwitkowitz, Jonathan Driedger, and Zhiyao Duan

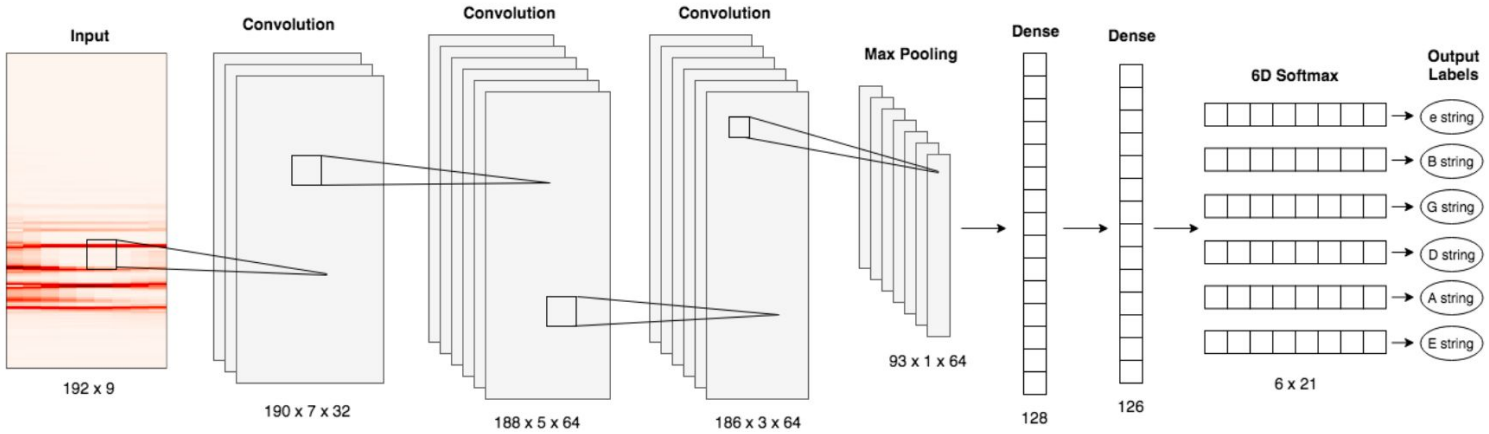
Guitar Tablature Transcription



Guitar Tablature Transcription



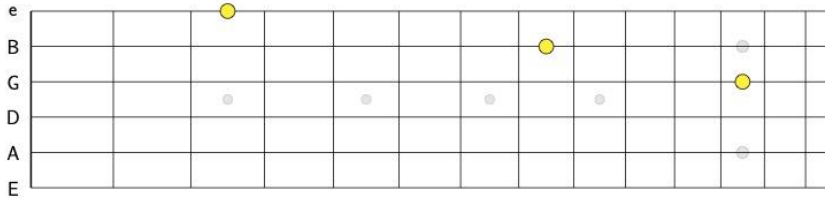
Example - TabCNN[1]



[1]Andrew Wiggins and Youngmoo Kim. “Guitar Tablature Estimation with a Convolutional Neural Network”. In: Proceedings of ISMIR. 2019.

6D Softmax Approach

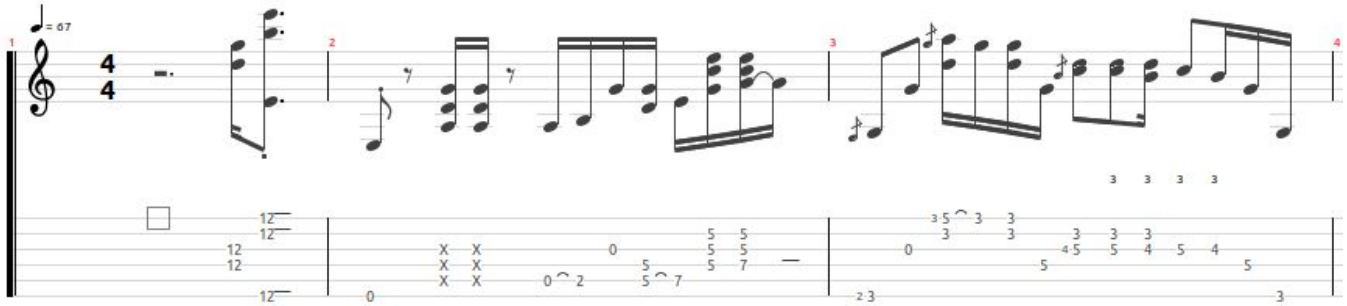
- Always generates valid predictions
- 6 independent classification tasks
- Ignores high correlation of fretting between strings
  - Physical limitations
  - Musical tendencies



The main obstacle for guitar tablature transcription is labeled data!

Symbolic Tablature

- Prescriptive playing notation for guitar
- Large digital collections available (e.g. DadaGP[2])



[2]Pedro Sarmento et al. “DadaGP: A Dataset of Tokenized GuitarPro Songs for Sequence Models”. In: Proceedings of ISMIR. 2021

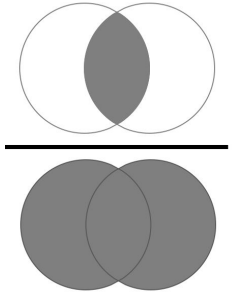
Proposed Methodology

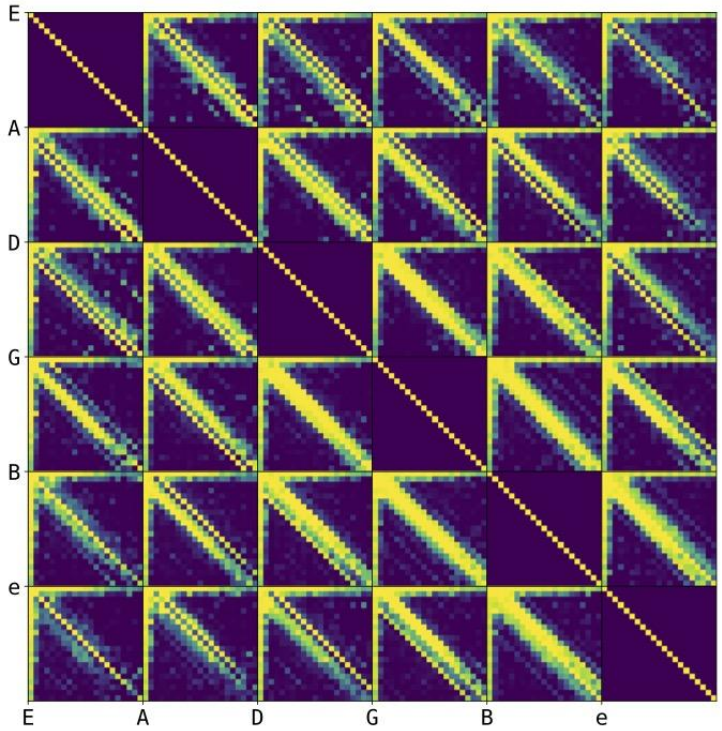
- Leverage tablature data
- Re-formulate output layer (sigmoid)
- Incorporate a novel inhibition loss
  - Discourage co-activation of unlikely note pairs
  - Informed by pairwise likelihoods estimated from data

Estimating Pairwise Likelihood

- Compute frame-level IOU

  of all string/fret (S/F) pairs





Inhibition Loss

$$w(c_i, c_j) = (1 - IoU(i, j))^b$$

Estimated likelihood for S/F *i* and S/F *j*.

Inhibition Loss

$$w(c_i, c_j) = (1 - IoU(i, j))^{b}$$

Boosting parameter to relax inhibition.

Frank Cwitkowitz | fcwitkow@ur.rochester.edu

Inhibition Loss

$$w(c_i, c_j) = (1 - IoU(i, j))^b$$

Inhibition weight for S/F *i* and S/F *j*.

Inhibition Loss

$$w(c_i, c_j) = (1 - IoU(i, j))^b$$

$$L_{inh} = \frac{1}{2N} \sum_{n=1}^{N} \sum_{i=1}^{C} \sum_{j=1}^{C} z_{c_i,n} z_{c_j,n} w(c_i, c_j)$$

Total number of frames.

Inhibition Loss

$$w(c_i, c_j) = (1 - IoU(i, j))^b$$

$$L_{inh} = \frac{1}{2N} \sum_{n=1}^{N} \sum_{i=1}^{C} \sum_{j=1}^{C} z_{c_i,n} z_{c_j,n} w(c_i, c_j)$$

Total number of distinct S/F combinations.

Inhibition Loss

$$w(c_i, c_j) = (1 - IoU(i, j))^b$$

$$L_{inh} = \frac{1}{2N} \sum_{n=1}^{N} \sum_{i=1}^{C} \sum_{j=1}^{C} z_{c_i,n} z_{c_j,n} w(c_i, c_j)$$

Activation produced by model for S/F *i* and S/F *j* during frame *n*.

Inhibition Loss

$$w(c_i, c_j) = (1 - IoU(i, j))^b$$

$$L_{inh} = \frac{1}{2N} \sum_{n=1}^{N} \sum_{i=1}^{C} \sum_{j=1}^{C} z_{c_i,n} z_{c_j,n} w(c_i, c_j)$$

Inhibition weight for S/F *i* and S/F *j*.

Frank Cwitkowitz | fcwitkow@ur.rochester.edu

Experiments

- Employ TabCNN as a baseline (extending the model)
  - 6-fold cross-validation on GuitarSet[3]
  - Evaluate with same transcription metrics
  - Introduce “distribution” metrics
  - Ablation study w.r.t. the inhibition loss

[3]Qingyang Xi et al. “GuitarSet: A Dataset for Guitar Transcription”. In: Proceedings of ISMIR. 2018.

Discussion

- Tablature performance does not increase significantly
  - Small size of GuitarSet
  - Some label noise in GuitarSet
- Distribution of predictions more closely matches DadaGP

Summary

- Pairwise likelihoods of S/Fs estimated using DadaGP
- Propose novel inhibition loss incorporating likelihoods
- TabCNN (baseline) augmented with the inhibition loss

All code is available at
https://github.com/cwitkowitz/guitar-transcription-with-inhibition