DL Music HW1

R12942144 陳泓睿

Codes: Relative Codes

Methodology

- How I make it
  - Short Chunk CNN
  - Use the whole song to validation
- What I have tried
  - FCN
  - Simple CNN Network

Result

- Validation Accuracy: 0.569

```
Best Valid Accuracy: 0.56901041666666666, Get at Epoch 314
```

Findings

- The design of validation set can severely affect the training result
  - From only take one segment of song to taking full song, validation accuracy from 10%->56%
- Preprocessing is important
  - By discussion with other classmates, validation accuracy can be 80%+ if using preprocessing.
  - Unfortunately, I modify validation set so late that I cannot implement preprocessing like Spleeter in time.

Details: How I makes it

- Short Chunk CNN
  - Use shorter time slice, more focus on Instance-Level information
  - Transform audio to spectrogram, then use computer vision like model architecture to train model
- Training and Validation Set Design
  - For training dataset, randomly select one segment of song to train, and randomly select the other segment on the other epoch.
  - For Validation dataset, use multiple segments of a same song to form one validation batch, which can make validation more comprehensive

Details: What I have tried

- FCN and CNN
  - Longer time slice(compare to short chunk CNN), consider more on song-level information
  - Transform audio to spectrogram, then use computer vision like model architecture to train model

Result Analysis and Discussion

- Why modify validation set works?
  - Only take particular segment of the song for validation will makes the model choose the model base on only that particular segment.
- Why we don’t take full song to train?
  - Song-level information includes more information, but we only wants to consider vocal information foe singer classification.