# The Development of a Convolutional Neural Network to Classify Drum Sounds

**Louis Larsen**
Dr. Brian Kernighan

# Music and Artificial Intelligence

- **Instrument groups are distinguishable**
  - **(Blaszke and Kostek, 2022)**
- **Limited classification within subset**

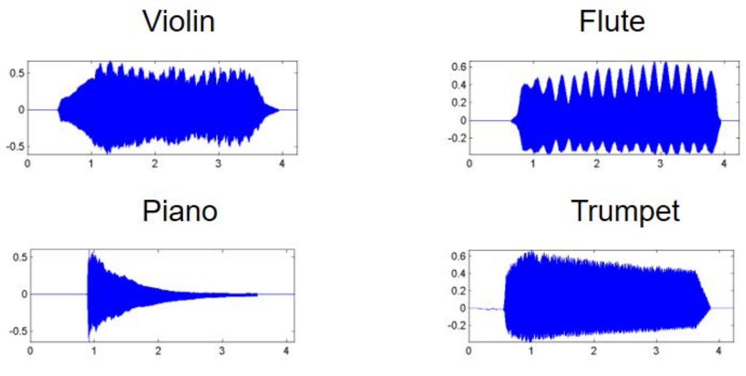



*Figure 1.* Waveform representations of C4 played on four instruments. Adapted from "Fundamentals of Music Processing" by M. Müller, 2021, *Springer*.

# Related Work

- **Violin and Viola**
  - **(Tan et al., 2022)**

- **Percussion and Drum**
  - **(Herrera et al., 2002)**
  - **(Chhabra et al., 2020)**

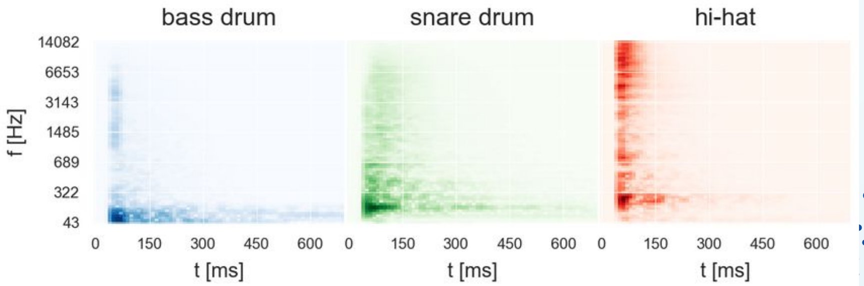



*Figure 2.* Spectrograms for drum sounds. Adapted from "Deep Learning Methods for Drum Transcription" by R. Vogl, 2018, Nov, *Computational Perception*.

“

To what extent can a convolutional neural network assist in the classification of drum samples?

# Hypothesis and Assumptions

**My CNN will maintain at least a 90% accuracy rating to classify percussion instruments on a drum set.**

**Unpitched instruments will be distinguishable by a computer utilizing MEL Spectrograms.**

# Creating the Database

**1440**

**Drum Samples (.WAV)**

**4**

**Drum Instruments**

# Creating the Spectrograms

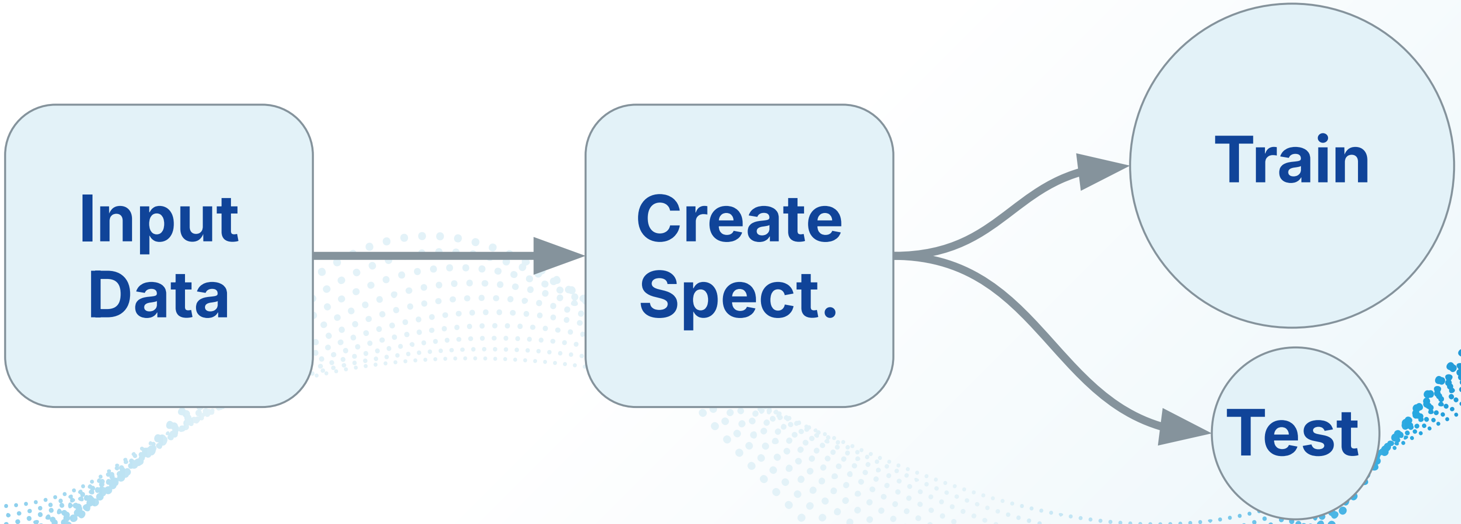



*Figure 3.* Flowchart for the structure of *createSpectrograms.py*, showing a 80/20 split between testing and training data.

# CNN Structure

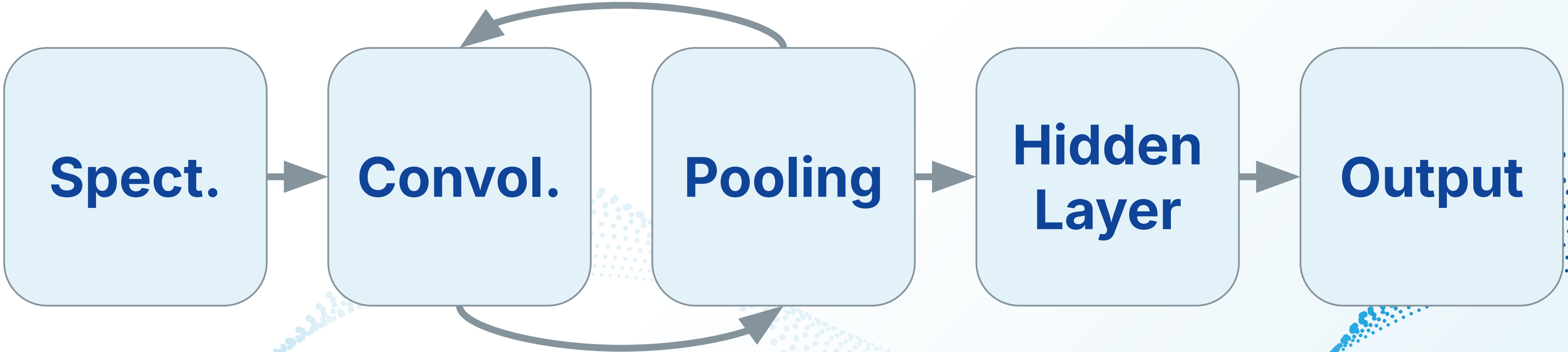



*Figure 4.* Flowchart for the structure of *CNNClassifier.py*, showing the structure of the convolutional neural network used for predicting drum instruments.

# Challenges

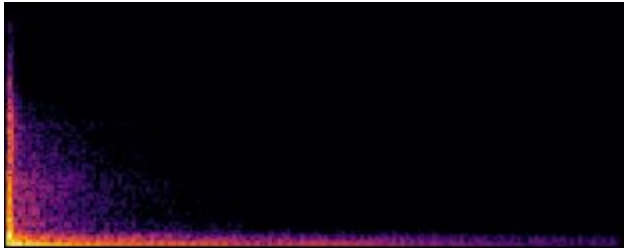
## Spectrogram

- Creating Database
- Stereo and Mono
- Saving to Directory

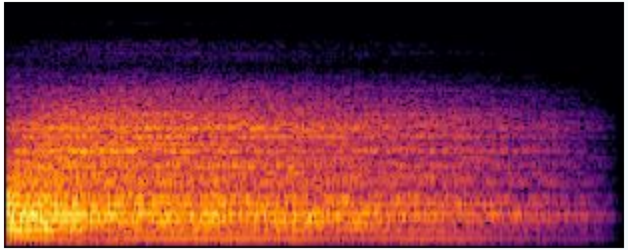
## CNN

- Mapping Labels
- Memory Intensive
- Learning Rate

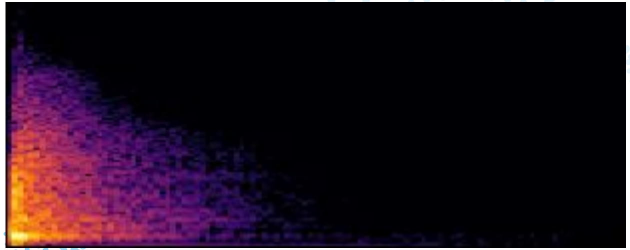
# Results - MEL Spectrogram



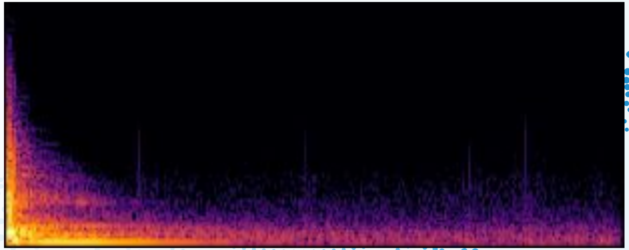
**Kick**



**Cymbal**



**Snare**



**Tom**

*Figure 5.* Shows example generated spectrograms for each of the four instrument types.

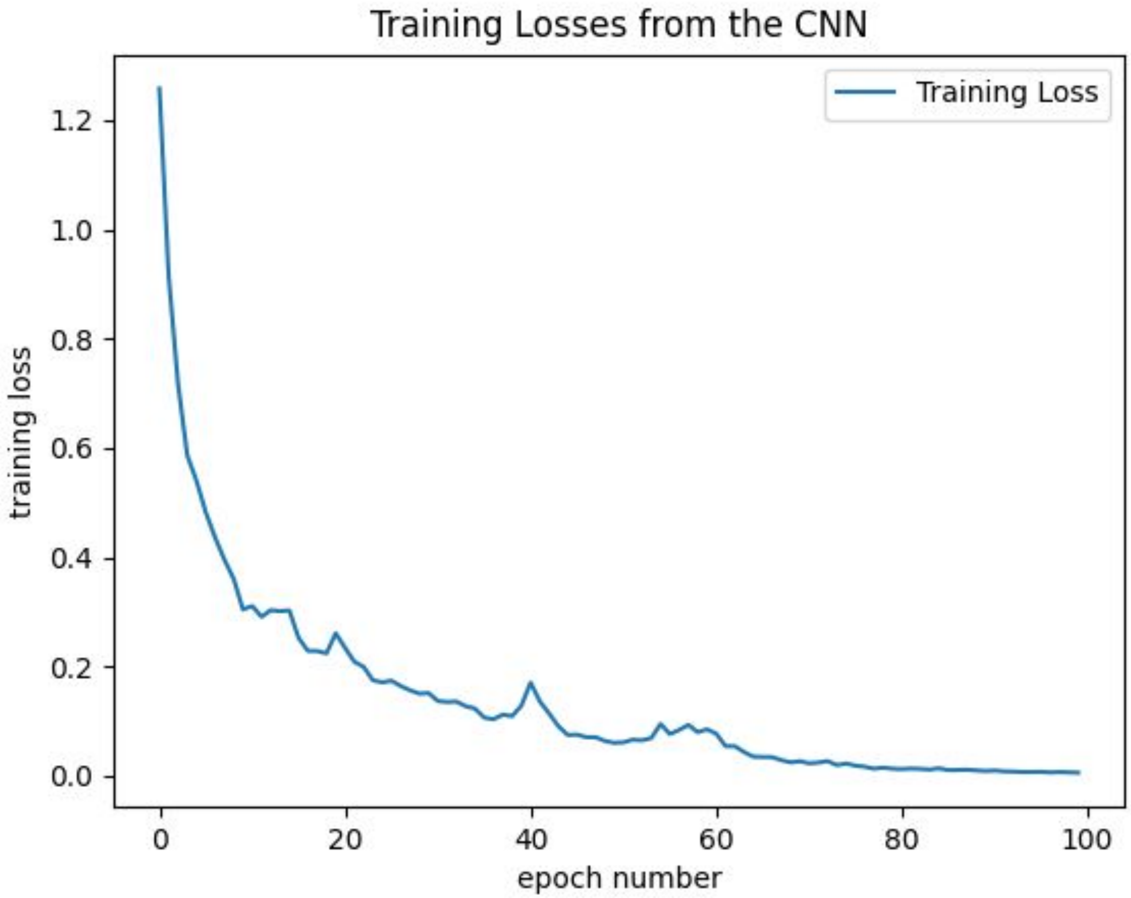



*Figure 6.* Shows Training Loss per epoch number, with a 1/x relationship.

# Results - CNN

**94%**

**Accuracy**

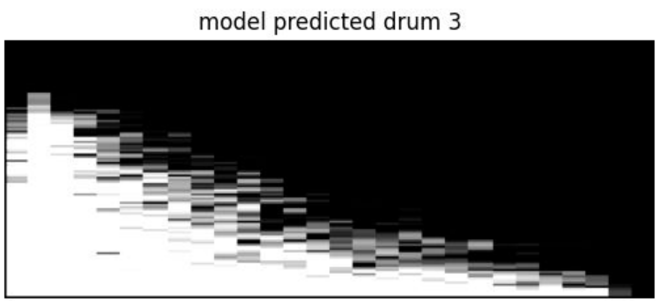
**.006**

**Training Loss**
***(Last Epoch)***

# Limitations

## Similar Frequency, Different Instrument

**Predicted Tom, Expected Kick**




## Corrupted Spectrogram

**Predicted Kick, Expected Cymbal**

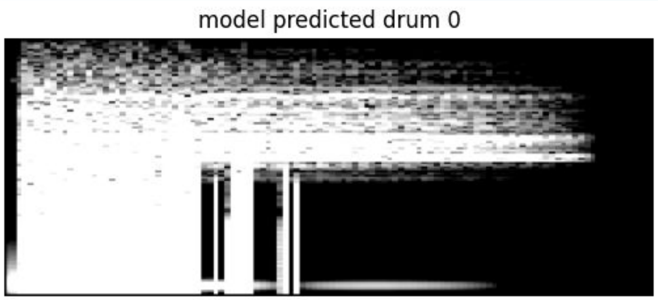



*Figure 8.* Shows edge cases of spectrogram failures, such as misclassification or corruption.

# Conclusions

- **Research classification within subset**
- **94% Accuracy**
- **Creative Applications**
    - **Sampling**
    - **Transcription**
    - **Generation**

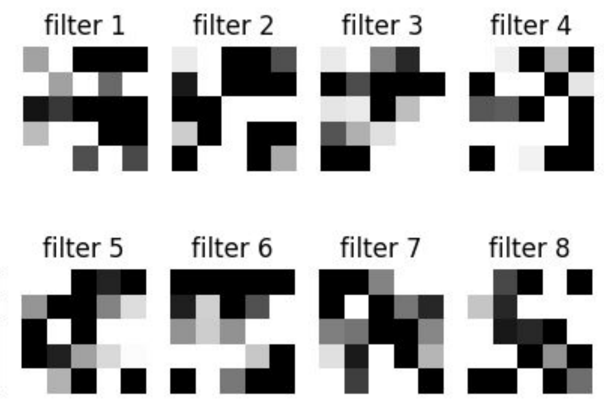



*Figure 9.* The weights assigned to each pixel of the first eight filters.

# Future Work

- Increase Accuracy
- Spectrogram Corruption
- Add More Classifications
- Work Toward Future Applications

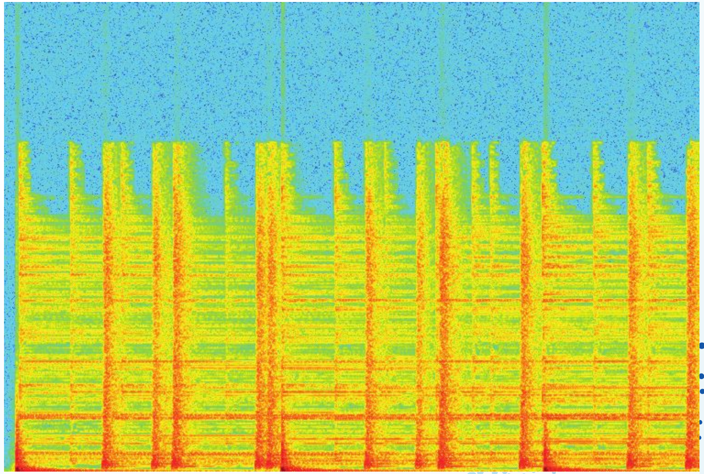


*Figure 10.* Spectrograms of percussive instruments. Adapted from "Phase-based Harmonic/Percussive Separation" by E. Cano, 2014, Oct, *Interspeech.*

# Acknowledgements


- Dr. Brian Kernighan
- Dr. Sierra Eckert
- COS IW classmates