

Classifying AudioMNIST using PyTorch

Jay Rodge

Illinois Institute of Technology



A 7x10 grid of handwritten digits from 0 to 9. Each row contains the digits 0 through 9 in a specific handwriting style. The styles vary significantly between rows, showing different levels of slant, thickness, and spacing. The digits are white on a black background.

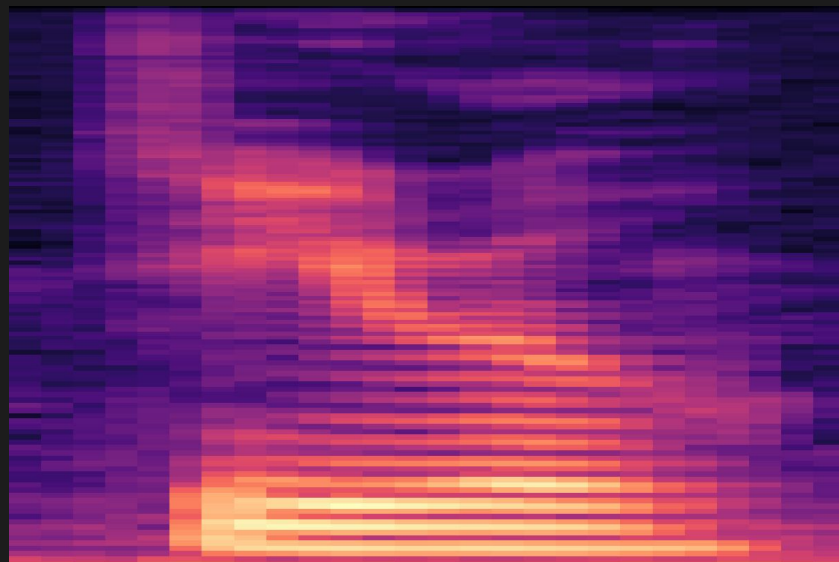
Audio-MNIST

Audio-MNIST

- 0-9 digits spoken by various people from different country
- 3000 audio samples for each class
- 30000 audio files in total

Spectrograms

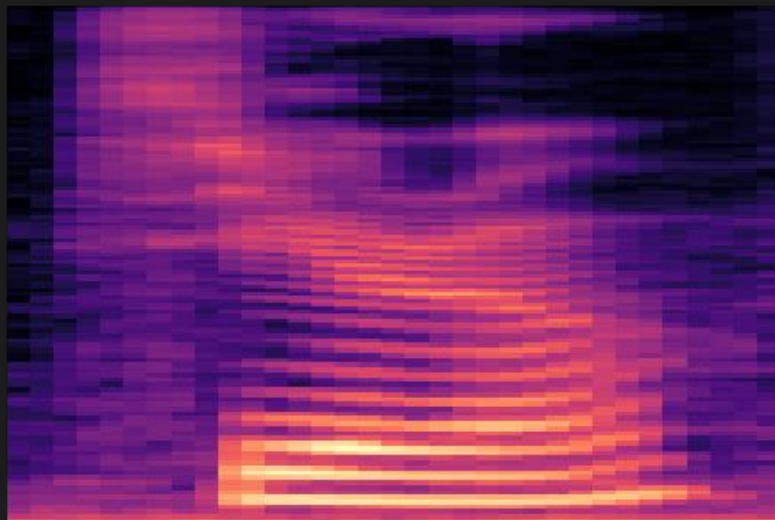
Spectrogram



Spectrogram

Librosa

```
y, sr = librosa.load('0_1.wav')  
data = librosa.feature.melspectrogram(y=y, sr=sr)  
librosa.display.specshow(librosa.power_to_db(data, ref=np.max), y_axis='mel', fmax=8000, x_axis='time')
```



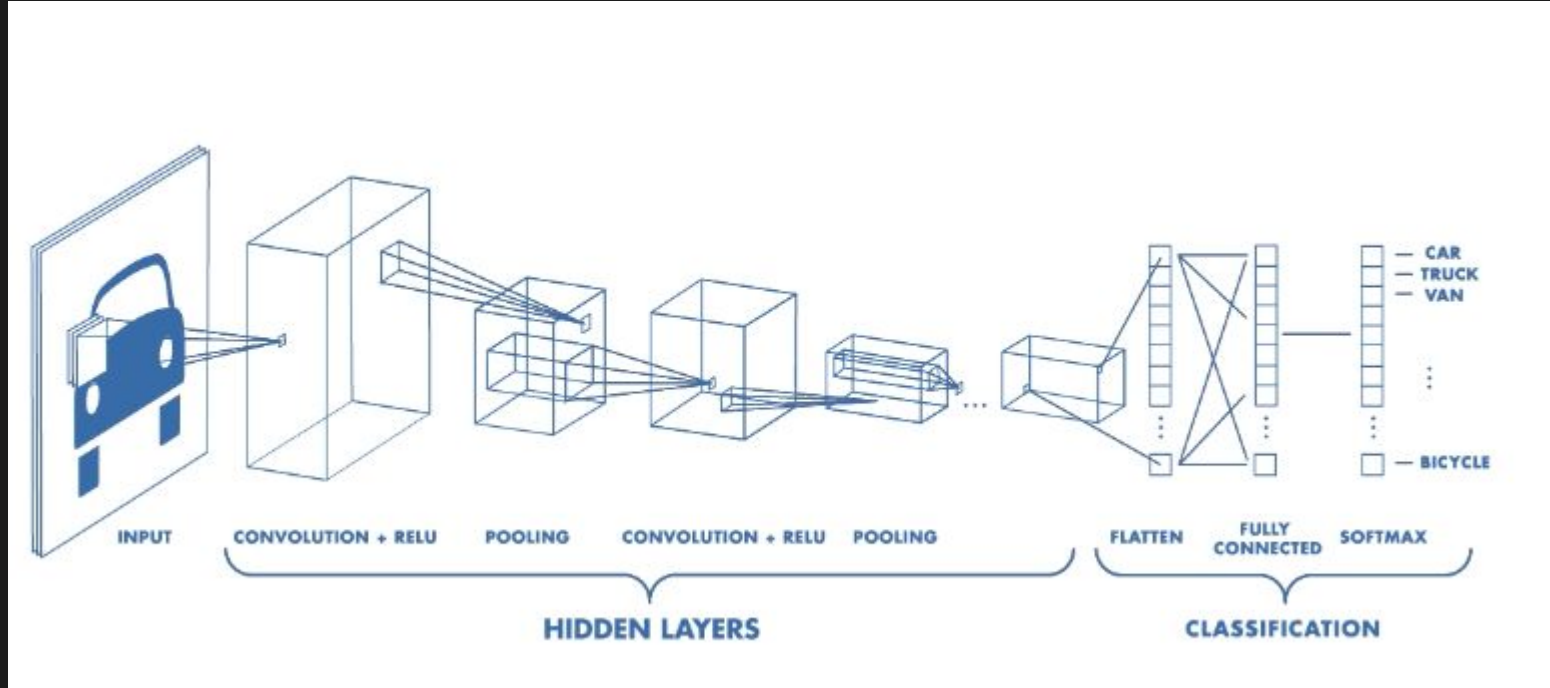
LET ME SPRINKLE SOME

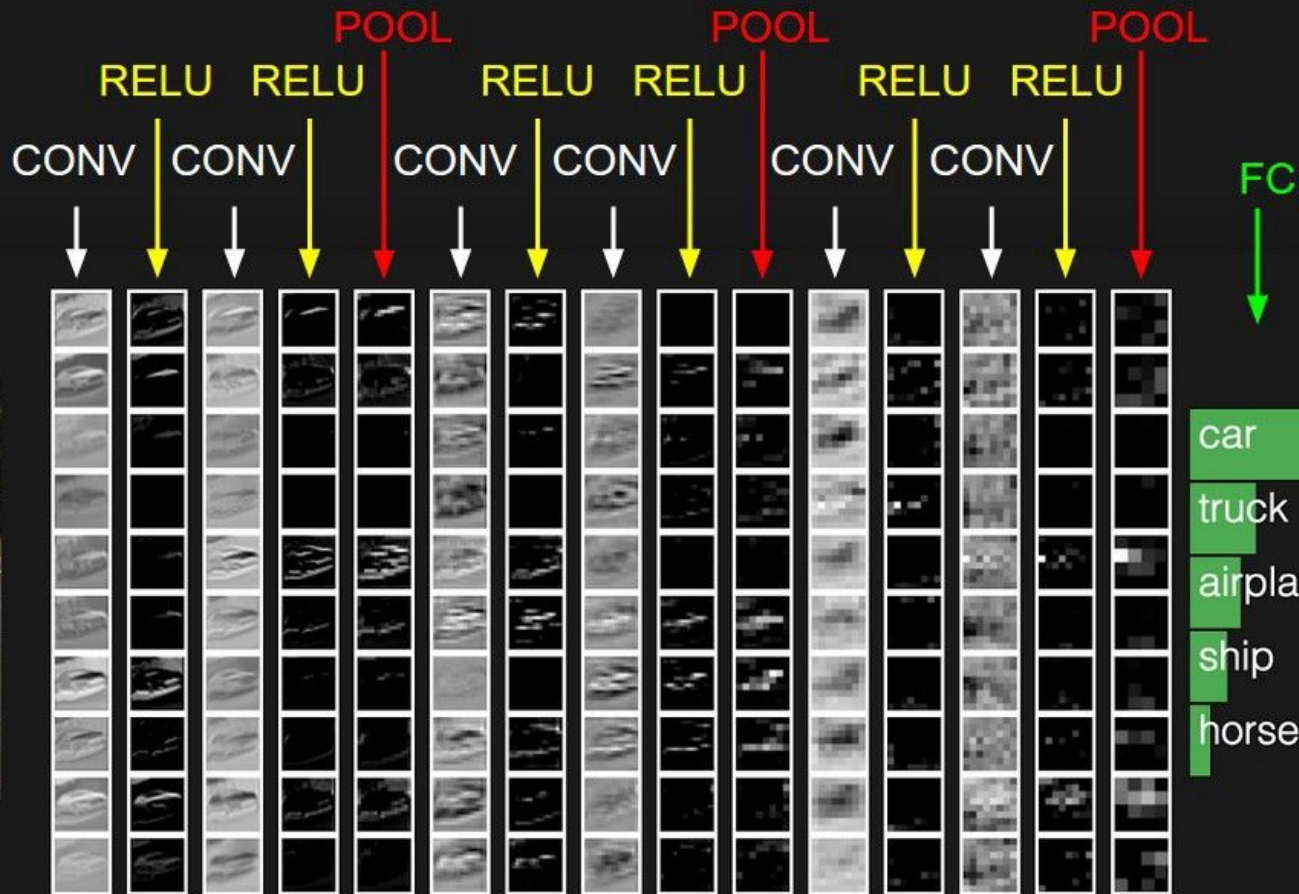


DEEP LEARNING MAGIC

memegenerator.net

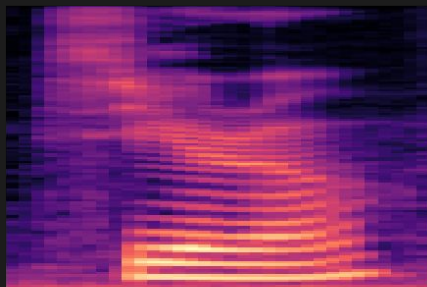
Convolutional Neural Network(CNNs)







Data



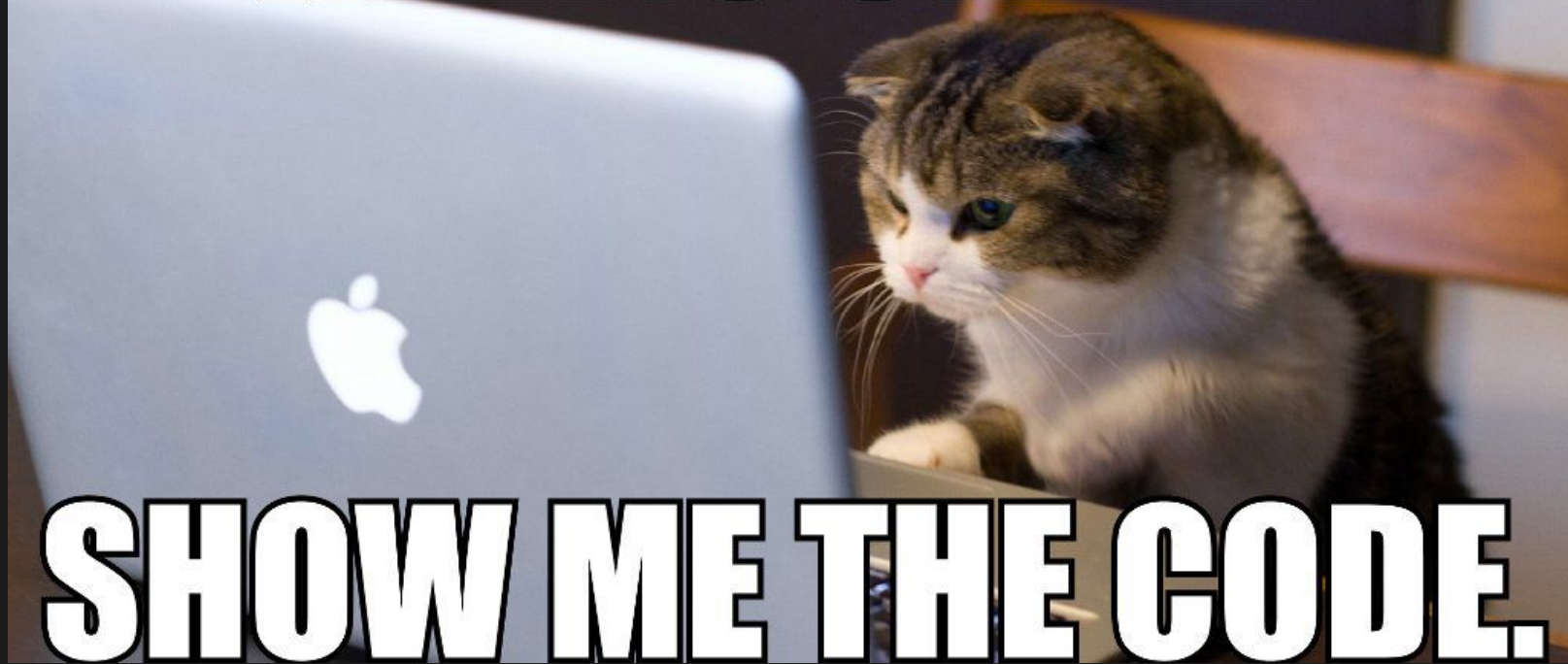
Model

CNN

Framwework

PYTORCH

TALK IS CHEAP.



SHOW ME THE CODE.

<https://github.com/jayrodge/AudioMNIST-using-PyTorch/blob/master/AudioClassifierMNIST.ipynb>



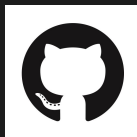
Thank You!

about.me/jayrodge

jayrodge@live.com



[/jayrodge](#)



[/jayrodge](#)



[/@jayrodge](#)