Visualization of Deep Networks for Musical Instrument Recognition

Charis Cochran, Youngmoo Kim {crc356 , ykim} @drexel.edu
Electrical and Computer Engineering Drexel University, USA

DREXEL UNIVERSITY ExCITe Center Expressive and Creative Interaction Technologies

METlab

Abstract

We present a visualization tool for Convolutional Neural Networks focused on the task of instrument recognition. This tool allows you to visualize the network response layer by layer to a specific input sample as an array of animated activation plots corresponding to nodes, or filters, in the network, as seen in Figure 2.

CNN Visualization Tool

- Visualization of CNN networks can aid in better understanding of prediction results and learned features in the network.
- When looking at networks for music specific tasks, such as musical instrument recognition, more relevant information on how the network responds to the music over time may be obtained by providing animated visualizations that can be compared with input audio.
- We implemented this visualization in Google Colab (**Code/Demo**: http://bit.ly/CNNVisualization) using python matplotlib and a pretrained Keras model.

Initial Results and Future Work

- From the initial results we can see some possible areas of interest in the network with clear differences in activation based on instrument class or prediction correctness.
- In the future, we are looking to extend the tool to other CNN visualization and deconvolution techniques to further improve our understanding of learned features in MIR specific deep networks as no such tool exists within current libraries for deep learning.

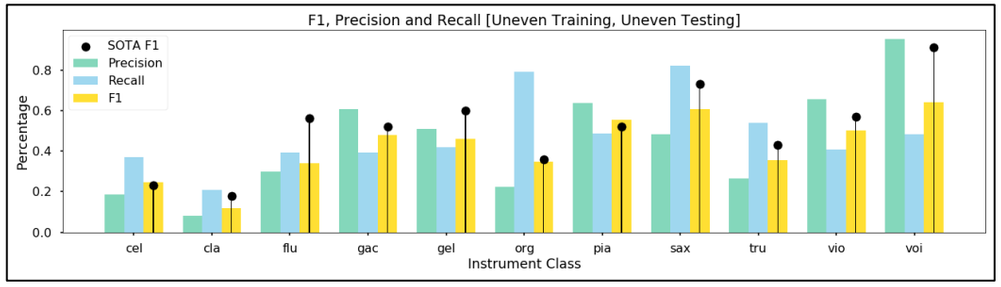

Figure 1. Model Performance As Compared to State of the Art

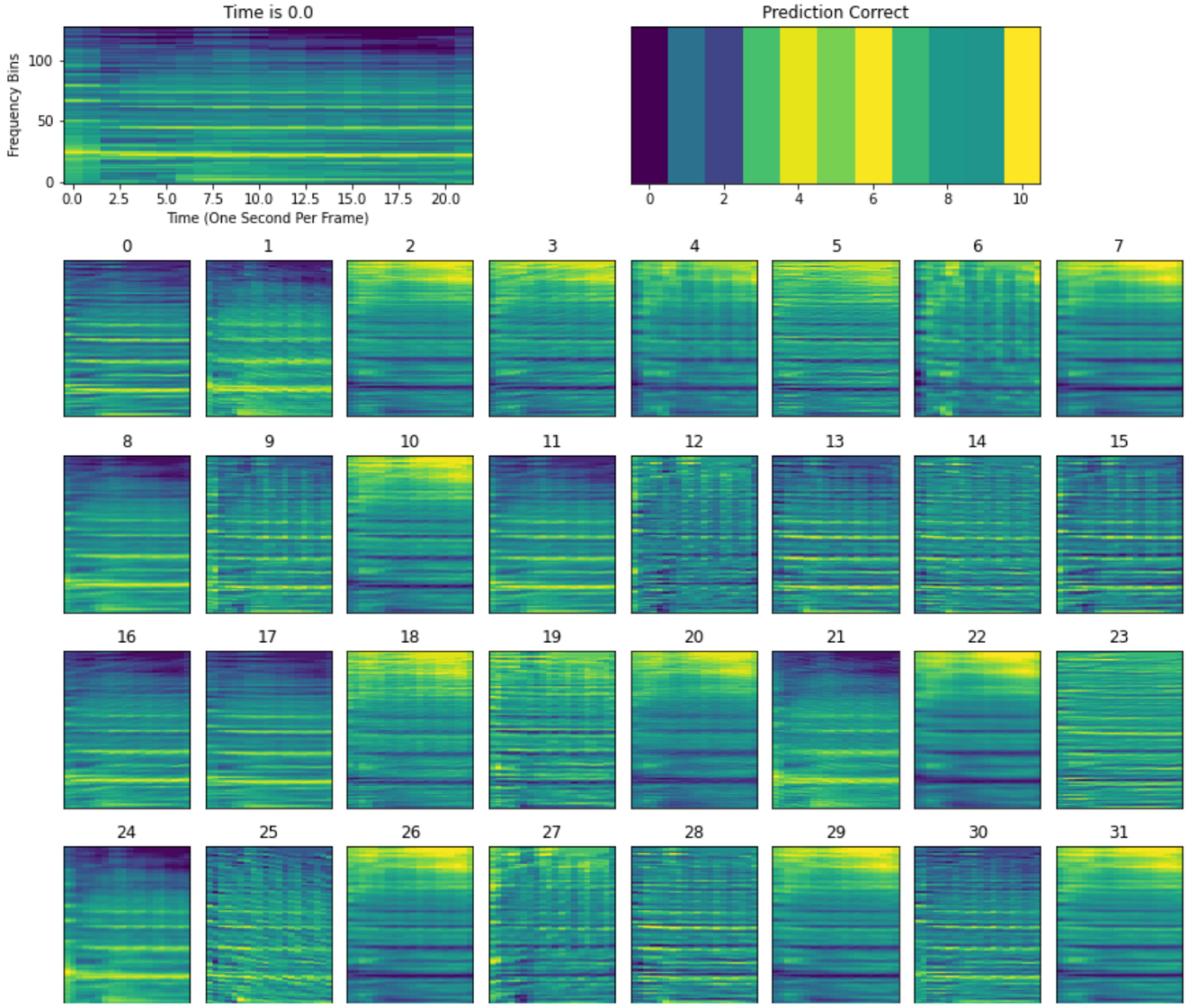

Figure 2. Single Frame of CNN Visualization Tool - Each animated visualization shows the original input Mel-spectrogram (top left), the final classification layer (top right), and an array of numbered nodes, or filter, activations within the CNN.