Landon Buell

Kevin Short

Abstract v5

15 September 2020

Classifying audio signals ~~by source~~ with machine learning has become ~~a topic of much~~ [an important topic of] research in the past few years. Models often involve ~~the production~~ [input] of a ^[2-D] spectrogram or ^[1-D] feature vector ~~and passing either array~~ into a network of a single type such as a Convolutional Neural Network (CNN) or Multilayer Perceptron (MLP). In this study, we explore ^[automatic classification of musical instruments using] a new hybrid neural-network architecture that combines the CNN and MLP models ~~to produce a signal classifier with~~ [and gives] superior performance over models that rely solely ^[on] one or the other. This hybrid network uses two branches, one being a CNN to process an image-like 2D spectrogram, and the other being an MLP to process a 1D feature vector. Within the model, a hidden layer combines activations from the two branches by concatenating them into a single 1D dense layer, allowing for any predictions to be a product of both branches. We describe in detail the ~~production and usage~~ [creation] of the spectrogram and ~~predictors~~ [feature vectors (?)], as well as how they influence the chosen network architecture. We finish with a practical demonstration ~~in using~~ [that uses] this classifier model to match waveforms from a chaotic music synthesizer to real-world musical instruments. Training data was provided by Philharmonia Symphony Orchestra and University of Iowa's Electronic Music Studios.

Training data came from the U. of Iowa's Electronic Music Studios and the Philharmonia Symphony Orchestra.

Q: What is the Philharmonia? Did they record the audio for the Electronic Music Studio, or was it a separate data set? I don't want to say "provided by" since that imples a degree of collaboration with the Philharmonia.