Landon Buell

Kevin Short

PHYS 799

29 August 2020

Bachelor of Science In Physics Senior Thesis Outline

1. Introduction
   1. The problem that I am trying to solve
   2. Background of Signal Processing / Neural Networks
   3. Current State of research/work in the field
   4. Gaps in knowledge of the field
   5. Hypothesis, concluding Remarks
2. Methods
   1. Design a Program to Map soundwaves to musical instruments
   2. Function:
   3. Use Neural networks as Function *F*
   4. Develop features for input
   5. Map to *k* possible output classes 0,1,2,…,k-1
3. Using Neural Network
   1. Some problems are exceedingly difficult to solve with conventional programming techniques
      1. A neural network is just composite a mathematical function
      2. Composed of *layers* which are also function
      3. Each Layer feeds into the next
      4. Layers are made up of activations, a weighting matrix, and a bias vector
   2. Various architecture types, MLP, CNN, ensemble methods
      1. Discuss particular used outputs
      2. Use MLP for one set of inputs (features)
      3. Use CNN for more inputs (spectrogram image)
   3. Output of a neural network is a prediction (activations in final layer)
   4. Trained through *Optimization*
      1. Optimization in 33,000 dimensions!
      2. We optimize a cost/loss/objective function to have the *smallest* possible value.
      3. This is how a Neural network learns
      4. Back propagation, SGD
      5. ADAM optimizer – Hyper-parameters, initial conditions, regularizes, etc.
   5. We Optimize a neural network, which adjusts a set of parameters
   6. For a classifier network, we are producing hyper-planes of best fit
4. Designing the Features
   1. Importance of Feature set
      1. THIS IS THE PHYSICS!
      2. We use the physical properties and mechanics of soundwaves to develop features
      3. The best student is a terrible student if they don’t have the right material
      4. I.I.D. Assumption?
   2. Spectrogram Input
      1. Spectrogram tracks energy as a function of time in space.
      2. Divide waveform into “Time Frames” – 4096 samples, 75% overlap
      3. Take DFT along last axis – transpose
      4. Frequency on Y-axis, time on X-axis (Images and examples)
      5. This is a gray-scale image to a human, but a matrix to a computer
   3. Time-Domain-Features
      1. Time-Domain Envelope
      2. Time-Center of Mass
      3. Zero-Crossing Rate
      4. Auto-Correlation Coefficients
   4. Frequency-Domain Features
      1. MFCC’s
      2. Frequency-Center of Mass
   5. Uses of Features:
      1. Spectrogram is given to CNN branch – Convolution Layer Groups
      2. Features are given to MLP branch – Linear Dense Layers
      3. Result of each arm is concatenated and used to produce a single prediction
5. Interpreting the output
   1. For a “K-Classes” Classifier – there are K Neurons in the output layer
      1. Each one corresponds to a class
      2. Inputs are one-hot-encoded
      3. Neuron with the highest activation is the final prediction
   2. Optimization hopefully causes the features & spectrogram to allow the correct neuron to have the highest activation
      1. Each sample (audio file) gets its own prediction
   3. Classification Metrics
      1. Importance of a Confusion Matrix – What samples are *confused*
      2. Precision Score – Definition & Use
      3. Recall Score – Definition & Use
      4. F1 Score -Definition & Use
6. Experimental Results
   1. Example output and Predictions over a few unique classes fromt rain/test data set
   2. Performance of Model Overall
      1. Final Loss, Precision, Recall
   3. Examples from Chaotic Synthesizer Outputs
   4. Extra Data (Bird Song?)
7. Conclusions / Citations