Identification of Instruments Through Sound Characteristics

Group Members: Avery Cameron, Raymond Knorr, Kegan Lavoy, Mason Lane

Project Objective

To train a model using the NSynth dataset, with the aim of distinguishing a note played on a violin with a note played on a piano. Additionally, various models could be trained using the NSynth dataset to distinguish between a wider variety of instruments. As well as testing our trained model on the NSynth dataset, we also plan on recording and labelling a smaller set of our own .wav files for various instruments to see if we can replicate results.

Description & Intended Achievements:

Sound in the form of .wav files are analyzed to identify different instruments. Using the NSynth dataset, we will train our models to identify between given instruments using various approaches, focusing on Random Forests which will allow for identification through the most common prediction from a set of predictions.

State of the Art Approaches:

NSynth:

Using the NSynth dataset, we can leverage train/test data of 300,000 four second long .wav files encompassing 1000 instruments. NSynth sounds may be categorized by sound type (acoustic, electronic, etc), instrument family, pitch, note, velocity, and more.

Algorithms:

Random Forests: an ensemble method consisting of a set of randomly generated decision trees, where the overall output class is the mode of the individual tree's predictions.

Support Vector Machines: SVM finds the hyperplane that best divides a dataset, support vectors (examples that would change the hyperplane) are used as critical elements of the data set to help determine the hyperplane. Higher dimension planes can be used for classification through kernelling.

AdaBoost: Uses an ensemble method (combines several base models to produce one optimal predictive model). AdaBoost sets weights of classifiers and data points, classifiers with higher accuracy receive higher weights, and misclassified data receives a higher weight for models to try and fit the data. Can be modified for use with more than two classes.

k-NN (k-Nearest Neighbours): algorithm takes an input and copies the label of the class that the majority of that inputs neighbours belong to.

Student Responsibilities:

Avery: Train and predict instruments using the AdaBoost model. Record and label real audio samples of piano for test with algorithms. Modify AdaBoost for more than two instruments.

Ray: Train and predict instruments using a Random Forest model. Record and label real audio samples of violin for test with algorithms.

Mason: Train and predict instruments using Support Vector Machines.

Kegan: Train and predict instruments using a k-NN model

Dataset:

https://magenta.tensorflow.org/datasets/nsynth

References:

Avci, K., Arican, M., & Polat, K. (2018). Machine learning based classification of violin and viola instrument sounds for the same notes. *2018 26th Signal Processing and Communications Applications Conference (SIU)*. doi: 10.1109/siu.2018.8404422

Patil, S. D., & Sanjekar, P. S. (2017). Musical instrument identification using SVM, MLP& AdaBoost with formal concept analysis. *2017 1st International Conference on Intelligent Systems and Information Management (ICISIM)*. doi: 10.1109/icisim.2017.8122157

Seipel, F (2018) *Music Instrument Identification using Convolutional Neural Networks* (Master's Thesis, Technische Universit ¨at Berlin, Berlin, Germany). Retrieved from https://www2.ak.tu-berlin.de/~akgroup/ak_pub/abschlussarbeiten/2018/Seipel_MasA.pdf