Efficient Music Genre Classification with Deep Convolutional Neural Networks

Suo, Wency (School: Boise High School)

In this project, classification of music genres with deep neural networks is explored. A Convolutional Neural Network (CNN) was trained to identify 10 different music genres. Using the GTZAN dataset, 1000 files were split into train data and test data. The accuracy, loss, and time were monitored for each epoch when training and testing. By comparing the results from an optimized CNN to the accuracy and speed of a human baseline model, the hypothesis that the deep neural network will outperform the baseline is tested. Data suggests that the CNN model is much more efficient when compared to a human baseline. The human baseline correctly classified 43.3% of the samples. The CNN achieved an accuracy of 98% in 37 sec for every 10 samples on the train set and 68.7% in 36 sec for every 10 samples on the test set. This study shows deep learning is an efficient tool to use when classifying music genres since the accuracy of the network is higher than that of the human baseline model. Using image recognition to search for similar patterns within each audio sample, a neural network can more quickly and accurately sort music samples into the correct genre. A recommendation pipeline implementing the CNN model is being constructed, exemplifying the applications of a neural network able to classify music genres. This project aims to prove that a machine can recognize and sort music into categories and reuse that information to give recommendations efficiently. There is still room for improvement with the model itself but experimental data shows the potential for CNN in music genre classification and recommendation.

Awards Won: