Audio Tagging Challenge
Group Proposal
Madison Turano, Jingjing Xu, Jingyi Xiao

Problem Selection

- Tagging/classifying audio recordings
- We thought we would learn more from using audio data because we've never worked with audio data.

Dataset

- This audio dataset is quite large, the train set is about 9.5k samples unevenly distributed across 41 categories, with the 94-300 samples in each categories, and duration ranges from 300 ms to 30 s. The test set is about 1.6k samples with manually-verified annotations, and similar category distribution. All audio samples have a single label.

Deep Network Framework

- We will use convolution network, as we are analyzing the similarities between audio samples.
- We decide to use Caffe, since we want to practice it, and gain more experience with it.

Reference Materials

- A Survey of Audio-Based Music Classification and Annotation
- Getting Started with Audio Data Analysis using Deep Learning (with case study)
- Neural Network Design, Chapter 25: Case Study 3: Pattern Recognition

Performance Metrics

- Accuracy both in terms of overall and each category.
- Confusion matrix.
- Time used.

Rough Schedule

- Apr. 12, Rough Network designed;

- Apr. 16, Network designed, and group report draft, start working on individual report, create GitHub repo;
- Apr. 20, Exploring different optimizer and functions, with final modification on the report, also finish individual report.