Project -1 (Notes: 09/26/18)
A Dataset & Taxonomy for Urban Sound Research (Paper Overview)

Aim: Not to maximize accuracy but to study the dataset itself

- Paper defines Taxonomy of the Urban Sound data-set
- Experimentation (in this paper) done on dataset limited to 10 classes.
- Formulated dataset collected from FreeSound (online)
- The main goals while deciding dataset – 1) Contains sound in the urban field environment. 2) Real – Field recording. 3) Large and varied in terms f sounds and recordings
- Dataset used for experimentation – UrbanSound8K (8.75 hours of recording)
- The paper experiments the fact ([5] as cited in paper) that 4 seconds sufficient to identify environmental sounds with 82% accuracy
- Feature extraction: Mel-Frequency Cepstral Coefficient (MFCC) – characteristics genrally used for environmental sound analysis. Essentia Audio analysis library
- Standard classification algorithm used (Decision Tree, KNN, random forest, SVM, ZeroR) on 10-fold cross validation of dataset.

Proposal outline:

- Paper – Using standard classical algorithm to study datasets characteristics

Our Opportunities – To implement deep learning algorithm to improve accuracy of results

- We can implement one or more deep learning algorithm to show accuracy improvement.
- We can examine Google AudioSet and Urban Dataset to characteristics as done in the paper. Examining different algorithms to see what suits what and why?
- Classifier confuses between some relatable pairs of sound. We can identify if same is true for DL algorithms
- Is the pattern seen in the salience feature true in foreground and background if we resort to DL algorithm
- How does more improved classification effects salience features