

Synopsis

1 - Title of project: Audio Event Research

2 - Name and study ID of all group members:

- Stefan Frederiksen - s144469
- Jacob Johansen - s103808

3 - Background and motivation:

In this project we examine the possibility to use a neural network to classify environments in sound files. We focus on classifying sounds that are mostly present as background noise, in other sound files, such as car horns, sirens, street music, etc.

This is interesting for two reasons. In some situations it might be desirable to know the environment where a sound was recorded, such as for instance in an investigation of some sort. However, classifying background noise might also be used for better searching in sound files. For instance we might be able to search for YouTube videos containing car horn sounds, even though the name of the video might not be related to car horns at all.

4 - Milestones:

- Implement a (simple) CNN network that gets a descent accuracy on the native Urban8k data.
- Modify CNN such that we get a decent, or as high as possible, accuracy on Urban8k data altered with noise (white-noise & speech).
- Modify CNN such that we can classify multiple labels, data is Urban8k added together.
- Alter multi label data with noise and evaluate performance

References

- [1] Dan Ellis, Google Research Scientist. *Announcing AudioSet: A Dataset for Audio Event Research*, <https://research.googleblog.com/2017/03/announcing-audioset-dataset-for-audio.html>. March 30th, 2017.
- [2] Justin Salamon, Christopher Jacoby, and Juan Pablo Bello. *Urban Sound Datasets (Namely URBANSOUND8K)*, <https://serv.cusp.nyu.edu/projects/urbansounddataset/urbansound8k.html>, 22st ACM International Conference on Multimedia (ACM-MM'14), November 2014.
- [3] Karol J. Piczak. *Environmental Sound Classification with Convolutional Neural Networks*, <http://karol.piczak.com/papers/Piczak2015-ESC-ConvNet.pdf>, 2015 IEEE International workshop on machine learning for signal processing, September 1720, 2015, Boston, USA.