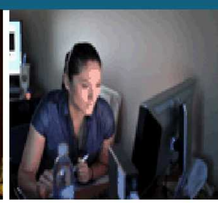
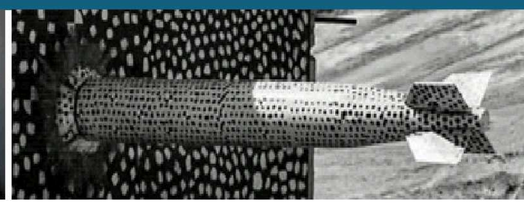
SAND2018-8214PE

Unlocking Real Time Infrasound Event Classification Abilities Using Machine Learning





PRESENTED BY

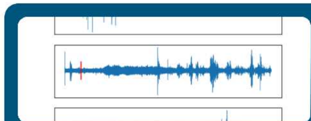
Sarah Albert (06752) and Lisa Linville (06363)

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

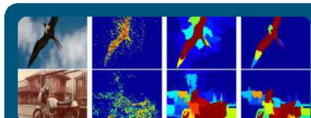
SAND 2018-XXXX

Outline


Introduction


Problem


Event Catalog

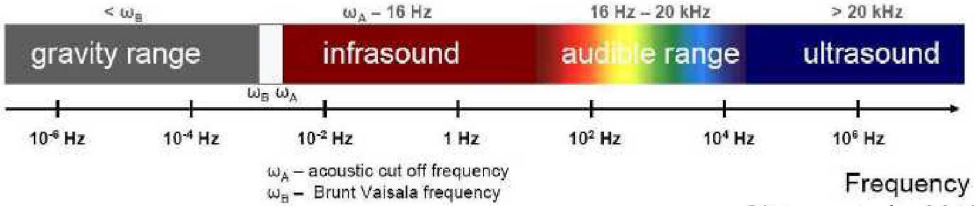

Approaches

Preliminary Results

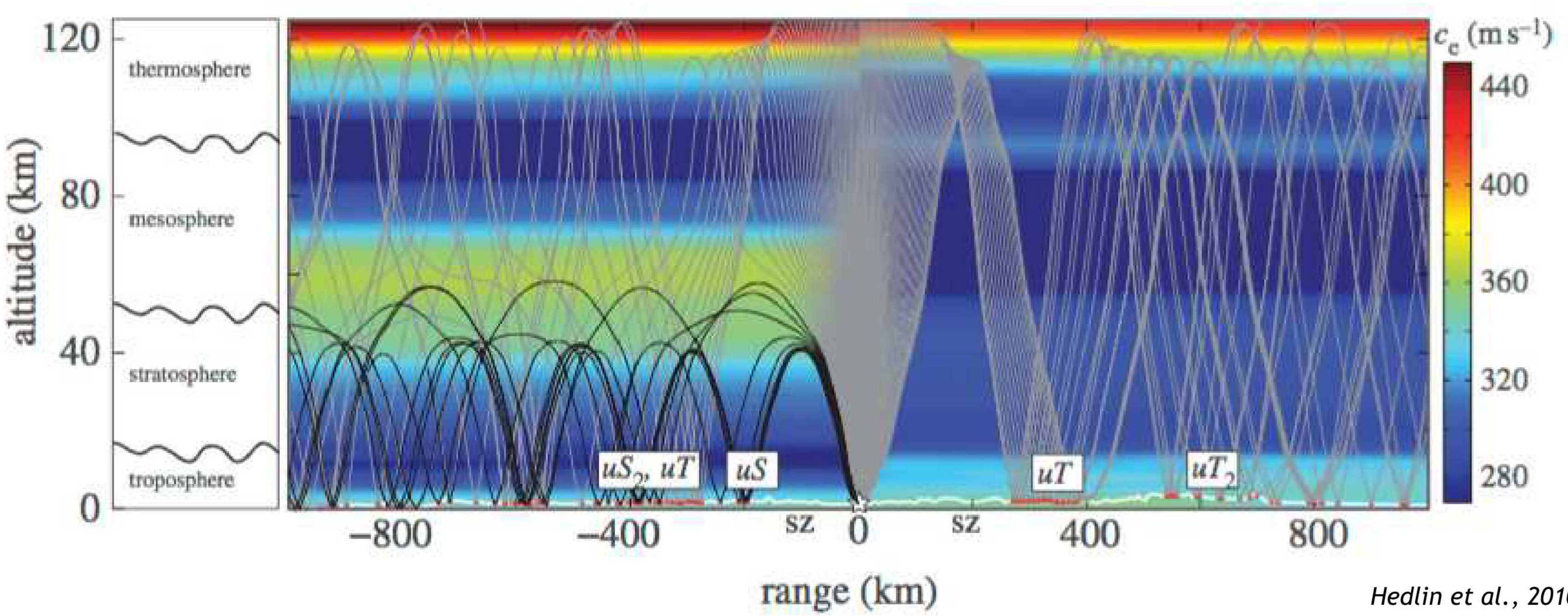

Conclusions and Future Work

What is Infrasound?

- Low-frequency sound
- Can travel thousands of kilometers



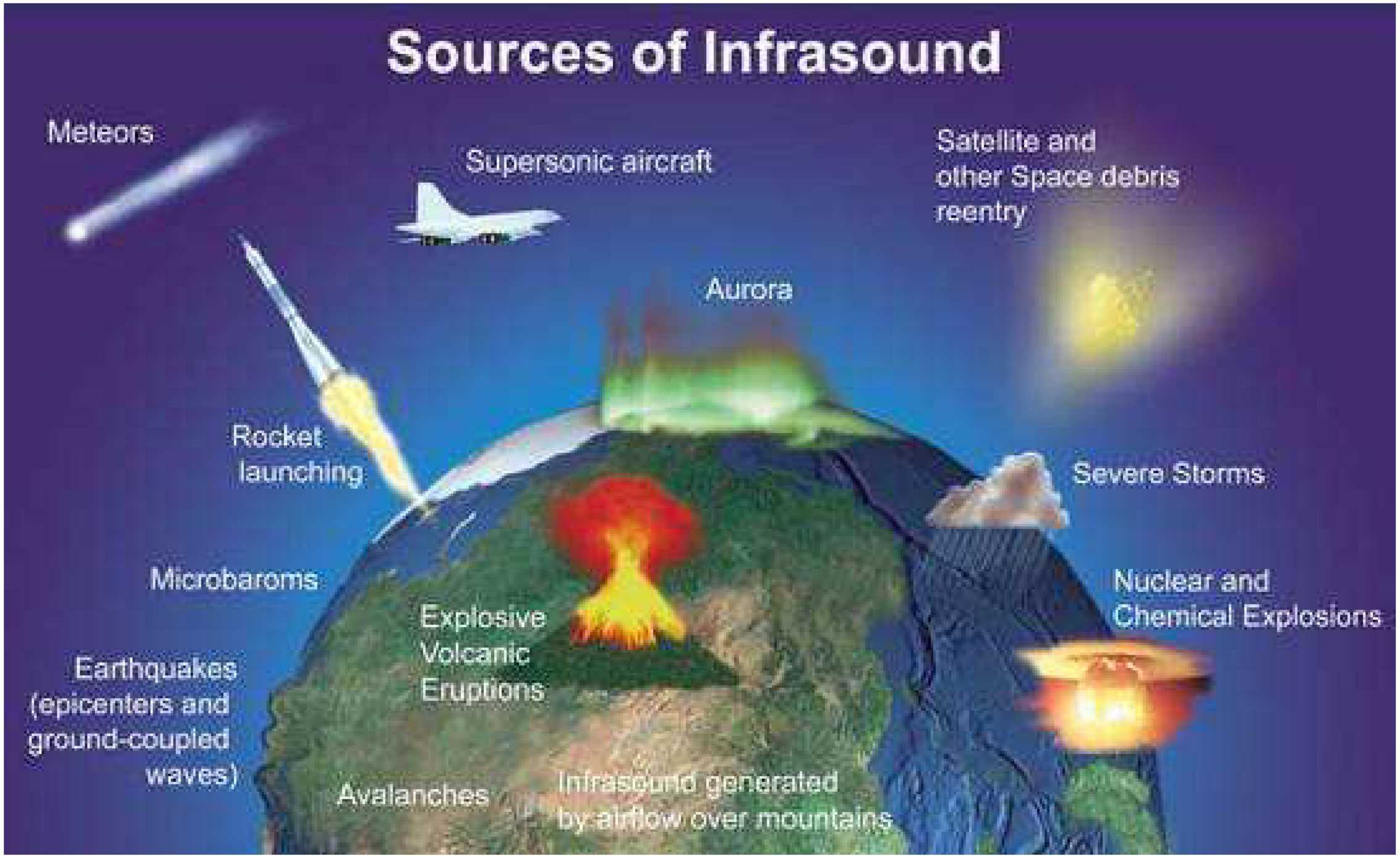
Bittner et al., 2010



Hedlin et al., 2010

What is Infrasound?

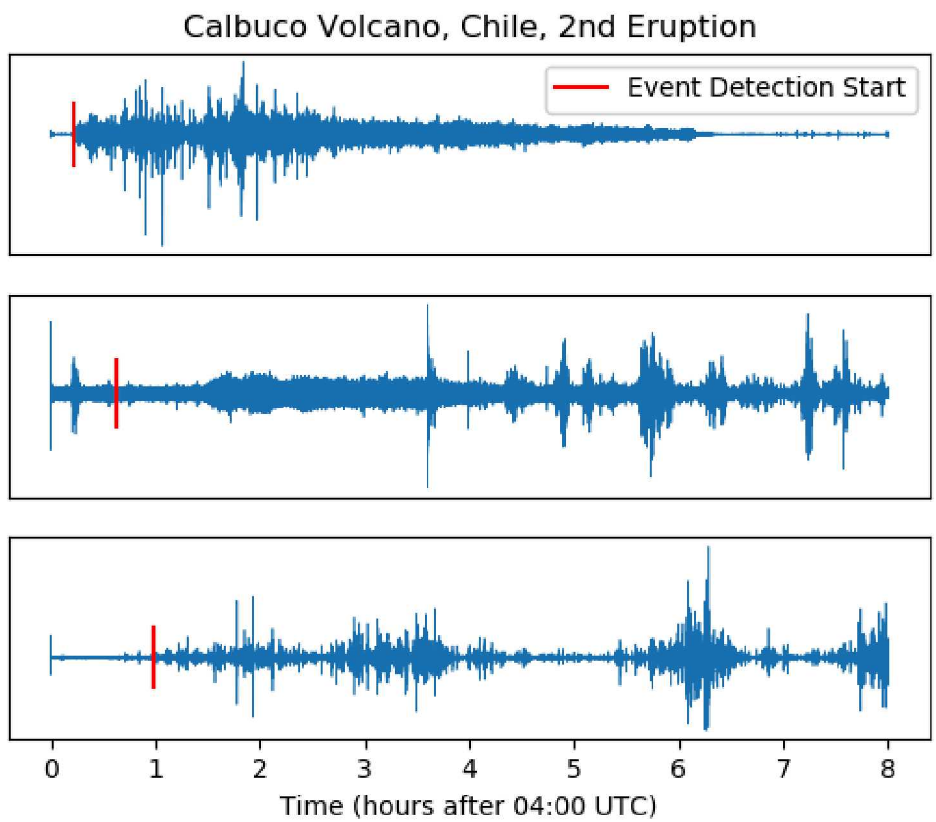
Generated by a variety of natural and anthropogenic sources.



https://neuroresearchproject.com/2013/02/19/1289/

The Infrasound Event Classification Problem

- Nearly impossible to classify by eye
- Analysts currently require ground truth for classification



Can we use machine learning to exploit similarities that cannot be seen by eye and use these to provide real time classifications?

A Global, Labeled, Infrasound Event Catalog

Labeled infrasound event catalog

- IMS catalog with local, regional, and global events

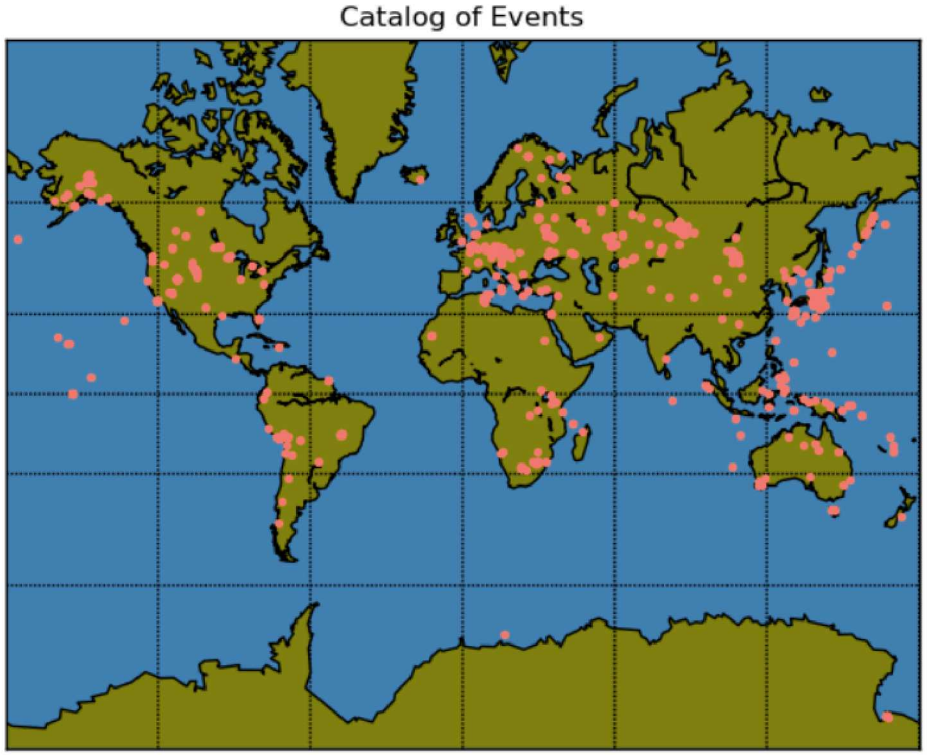
750 events

- Recorded at multiple stations
- Each station consists of at least 3 infrasound microphones
- Some events (volcanos and bolides) have multiple subevents

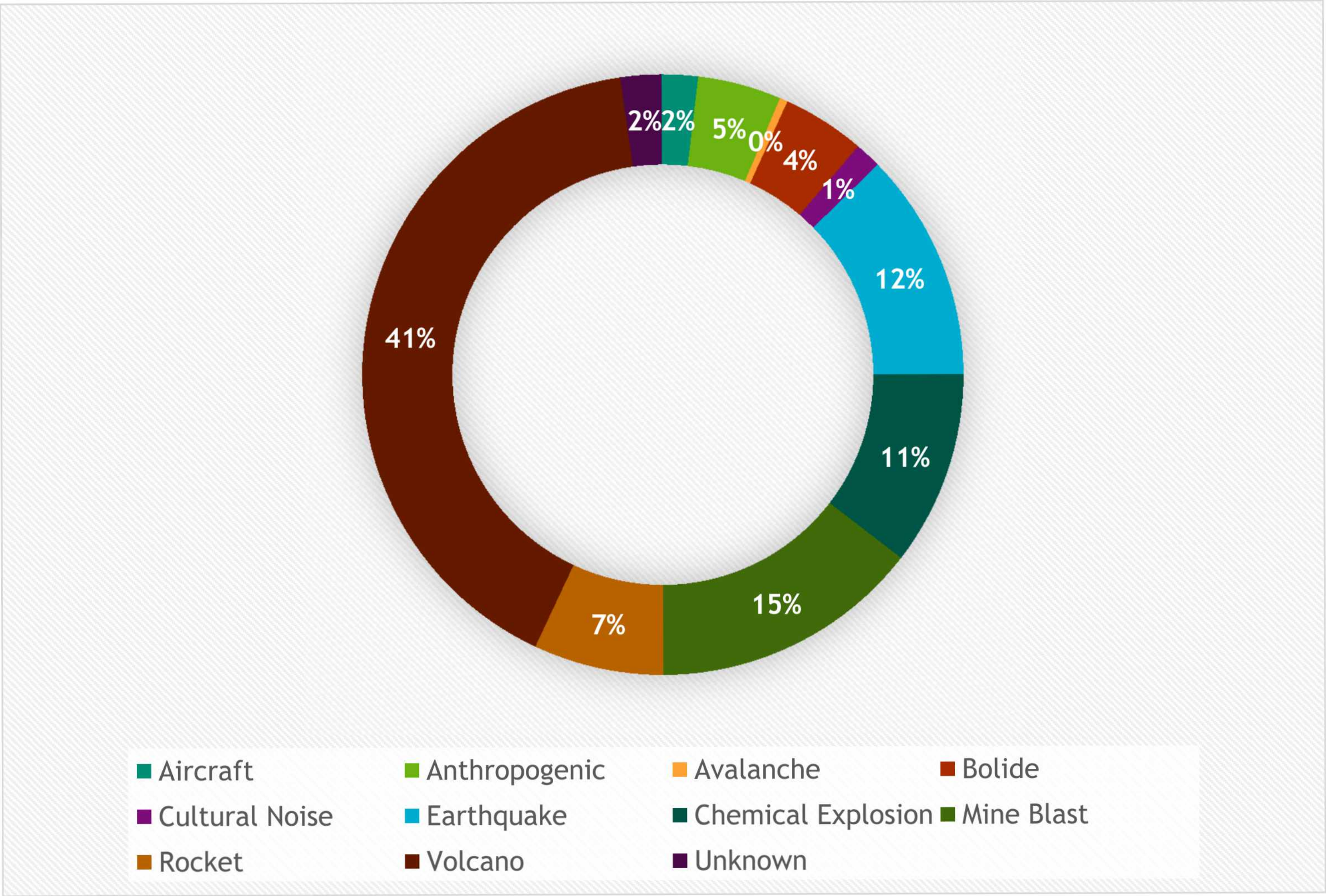
Variety of events

- Aircraft, Avalanches, Bolides, Cultural Noise, Earthquakes, Chemical Explosions, Anthropogenic Activity, Mine Blasts, Rockets, and Volcanoes

Total of **> 36,000** waveforms



Event Class Imbalance



We welcome suggestions on how to deal with this problem!

Two Approaches to a Solution

Method 1: SVM

Current feature extraction method: Spectral Entropy (Li et al., 2016)

- Wavelet Singular Spectrum Entropy
- Wavelet Power Spectrum Entropy
- Wavelet Energy Spectrum Entropy

Will experiment with more "physical" features in the future

Method 2: CNN

- Trains faster and can be more accurate than fully connected NN
- Testing in seismic domain indicates that performs nearly as well as RNN (LSTM) for sequences of similar length, but is more compact
- Potentially more accurate than SVM baseline because it identifies meaningful features that we are unaware of
- But less transparent (without directly encoding the physics, it is less clear which signal characteristics are most important for prediction

Preliminary Results - SVM

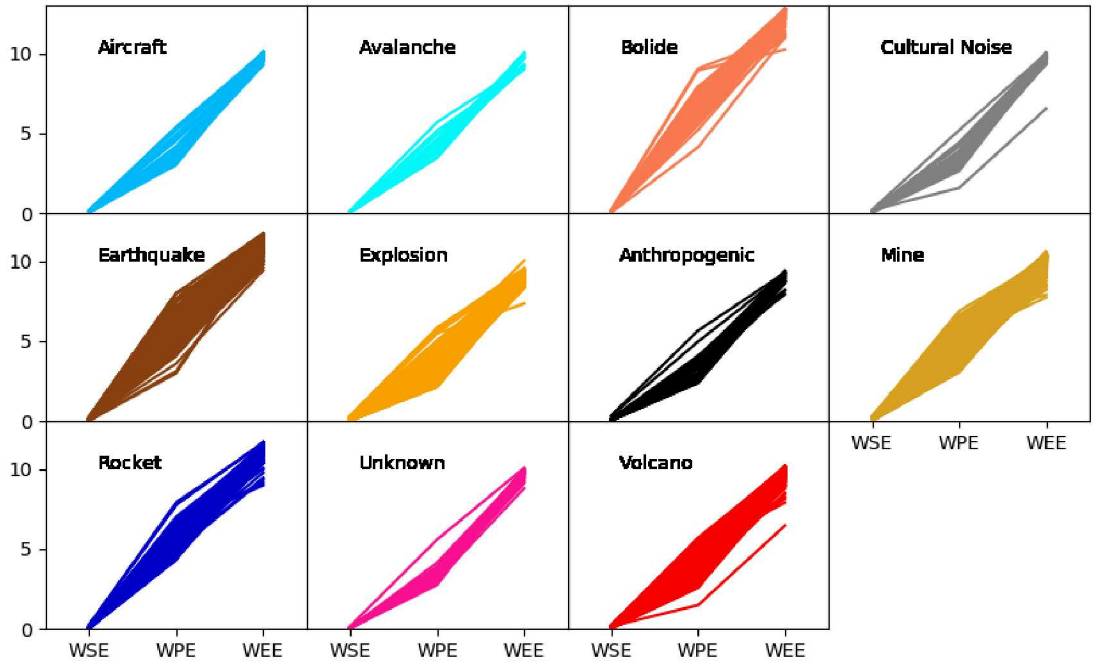
Completed feature extraction using Spectral Entropy (Li et al., 2016)

- 1,948 feature vectors/labels (consists of mean waveforms instead of all waveforms)

First-pass SVM shows an accuracy of 70%

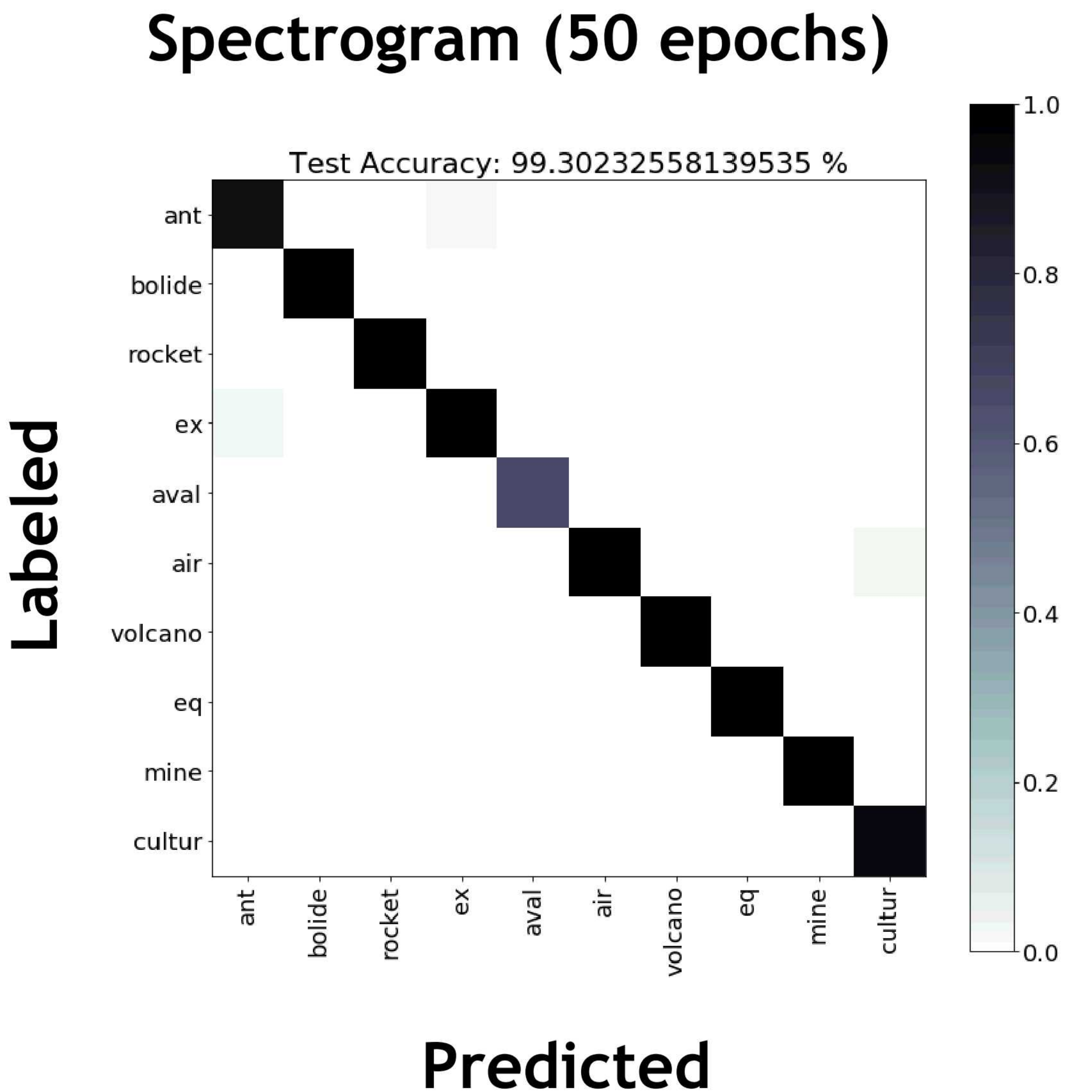
- Average of 10-fold cross validation

Note: This analysis was done with the previous event catalog version



Preliminary Results - CNN

- Exclude Unknown, use first arrival from each event.
- Samples are trimmed or padded to a length of 9000 (~1 std of average length)
- Exclude ~15k volcano signals (use first 3k randomly sampled class examples)
- Transform raw, filtered timeseries to spectrogram
- Achieve 85-99% accuracy depending on architecture and hyperparameters, but generalization is poor



Future Work

Feature extraction on newly updated event catalog

- Preliminary results from a previous catalog version

Try using more "physical" features for SVM

- Max, mean amplitude
- Dominant frequency
- Signal envelope

Achieve better generalization with CNN

- Use average signals across each array- less data but better performance?
- Dynamic length input sequences – not trimmed/padded
- Would RNN be a better choice for infrasound signals?

Test regions of input most important for prediction

- Saliency mapping- can we use model predictions to help analysts understand which aspects of the signal input are most diagnostic per class?

References

Bittner, M., Hoppner, K., Pilger, C., and Schmidt, C. (2010). Mesopause temperature perturbations caused by infrasonic waves as a potential indicator for the detection of tsunamis and other geo-hazards. *Nat. hazards Earth Syst. Sci.*, 10, 1431-1442. doi:10.5194/nhess-10-1431-2010

Hedlin, M. A. H., Walker, K. T. (2012). A study of infrasonic anisotropy and multipathing in the atmosphere using seismic networks. *Phil. Trans. Royal Soc.*, 371(1984). doi:10.1098/rsta.2011.0542

Pan, H., and Jiang, H. (2016). A Deep Learning Based Fast Image Saliency Detection Algorithm. *CoRR*, abs/1602.00577.