Nuclear Explosion Monitoring with Machine Learning
A LSTM Approach to Seismic Event Discrimination
Steven A. Magana-Zook

DSI DATA SCIENCE INSTITUTE

Lawrence Livermore National Laboratory

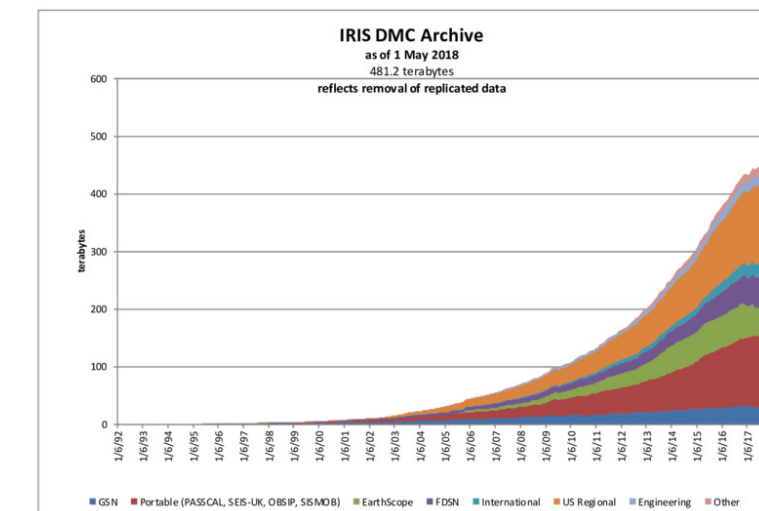UNIVERSITY OF CALIFORNIA National Laboratories

Background

- Nuclear explosion monitoring is 24/7 mission, that requires quick answers to events happening anywhere in the world.

- Incoming data streams are increasing exponentially, year over year, stressing existing monitoring infrastructure and staff demands .

- North Korea has conducted several recent tests, with new regional anxieties being stoked presently.
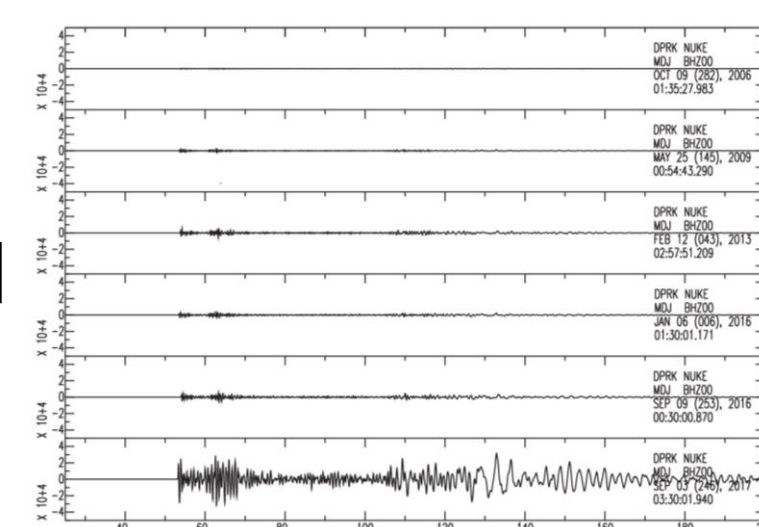
- Known test sites, such as the DPRK's Punggye-ri, hide the scale of the problem. Monitoring agencies must discriminate events occurring anywhere in the world _including_ at publicly known test sites.
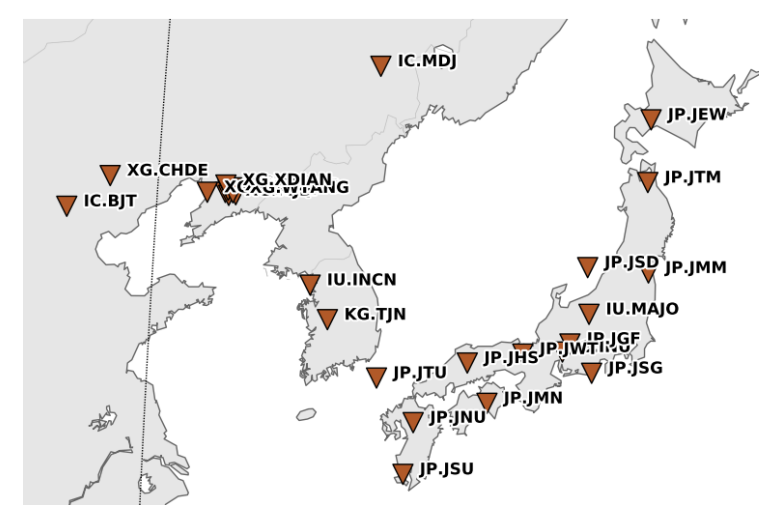
- LSTM was chosen as an algorithm to explore given its success with image recognition. By using spectrograms as a type of image, we hoped to tap into LSTM's powerful modeling capabilities.
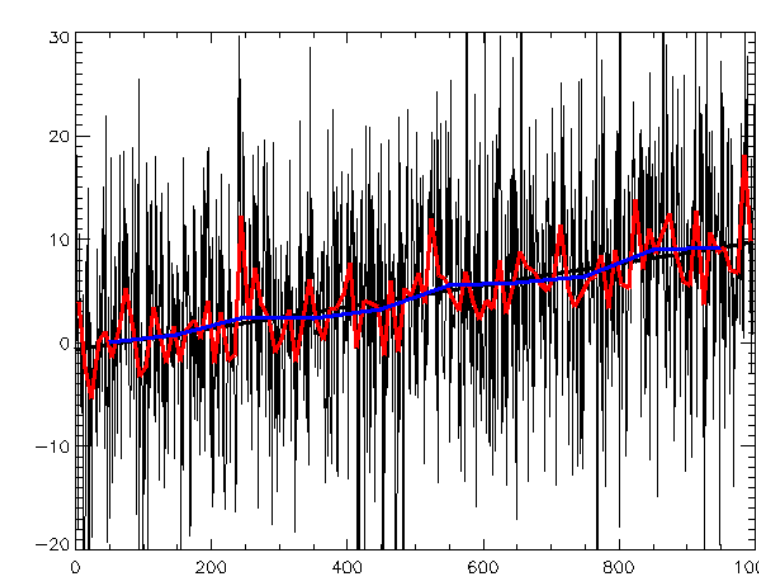
Objectives

- Investigate the problem space of explosion monitoring at scale – particularly through the use of machine learning methods.
- Determine what existing anomaly detection and classification methodologies can be transferred to the seismic domain?
- Develop a scalable solution wherein researchers can filter out routine signals needing no further analysis.
- Identify solutions to remediate the issues of quality control and data skew.

Data and Methods

Explosion Sources

Earthquake Sources
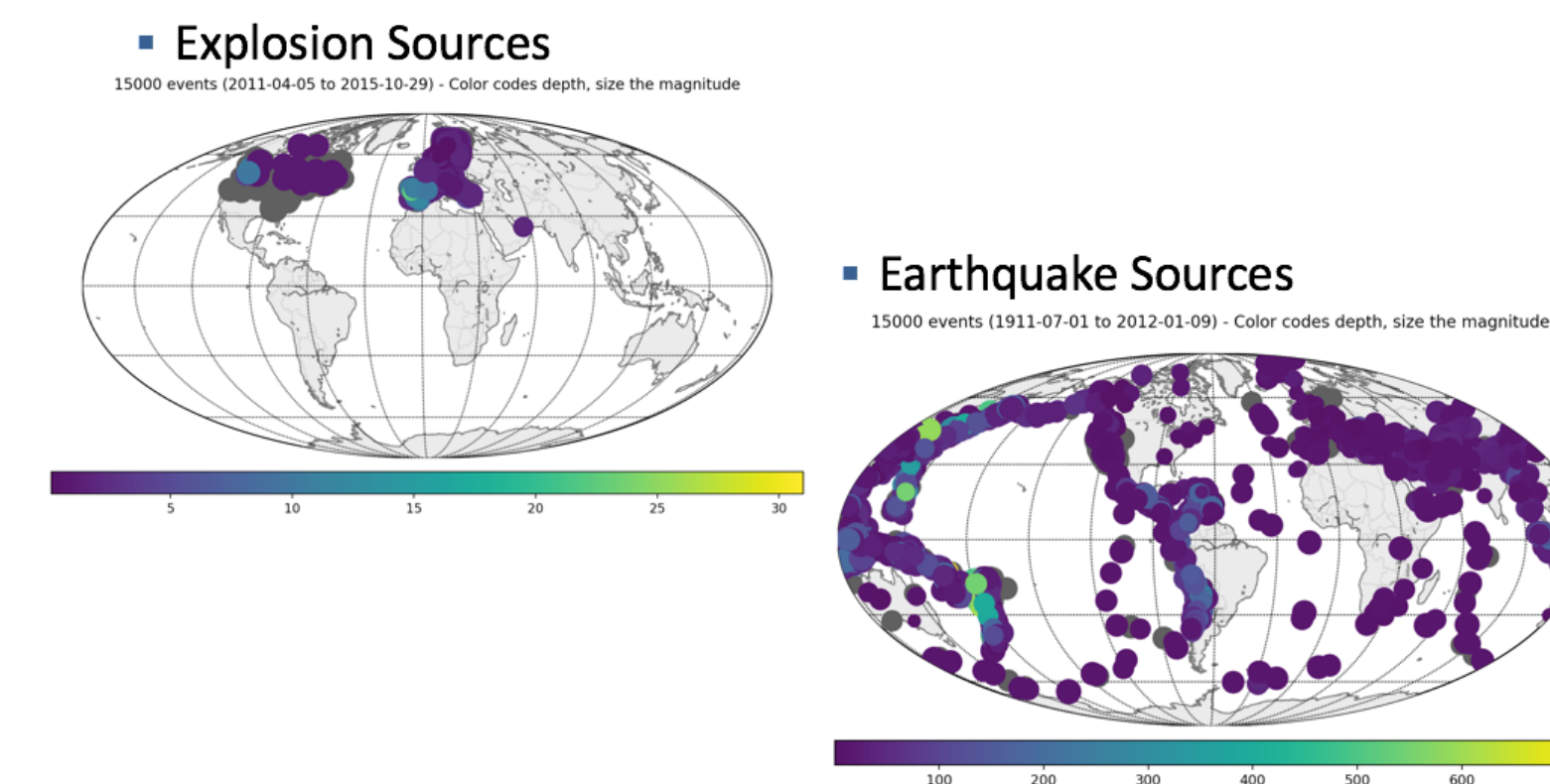
1. For each event in the global event catalog:
   1. For each station within 10-degrees of the event:
      1. Calculate p and s-wave travel times
      2. Download waveforms cut to 15 seconds before p-onset and 15 seconds after s-wave onset.
      3. Write waveform to labeled directory (explosion, earthquake) depending on event type.

Earthquake   Explosion
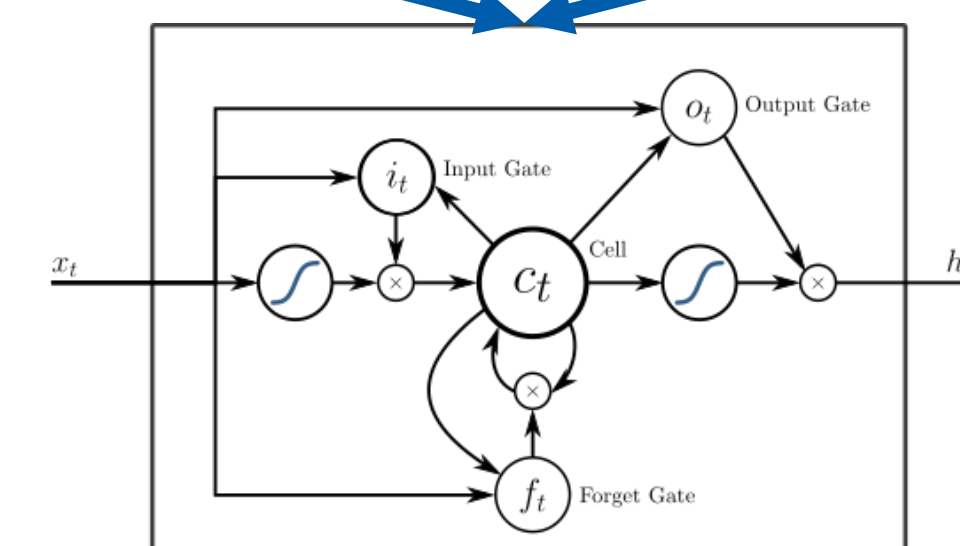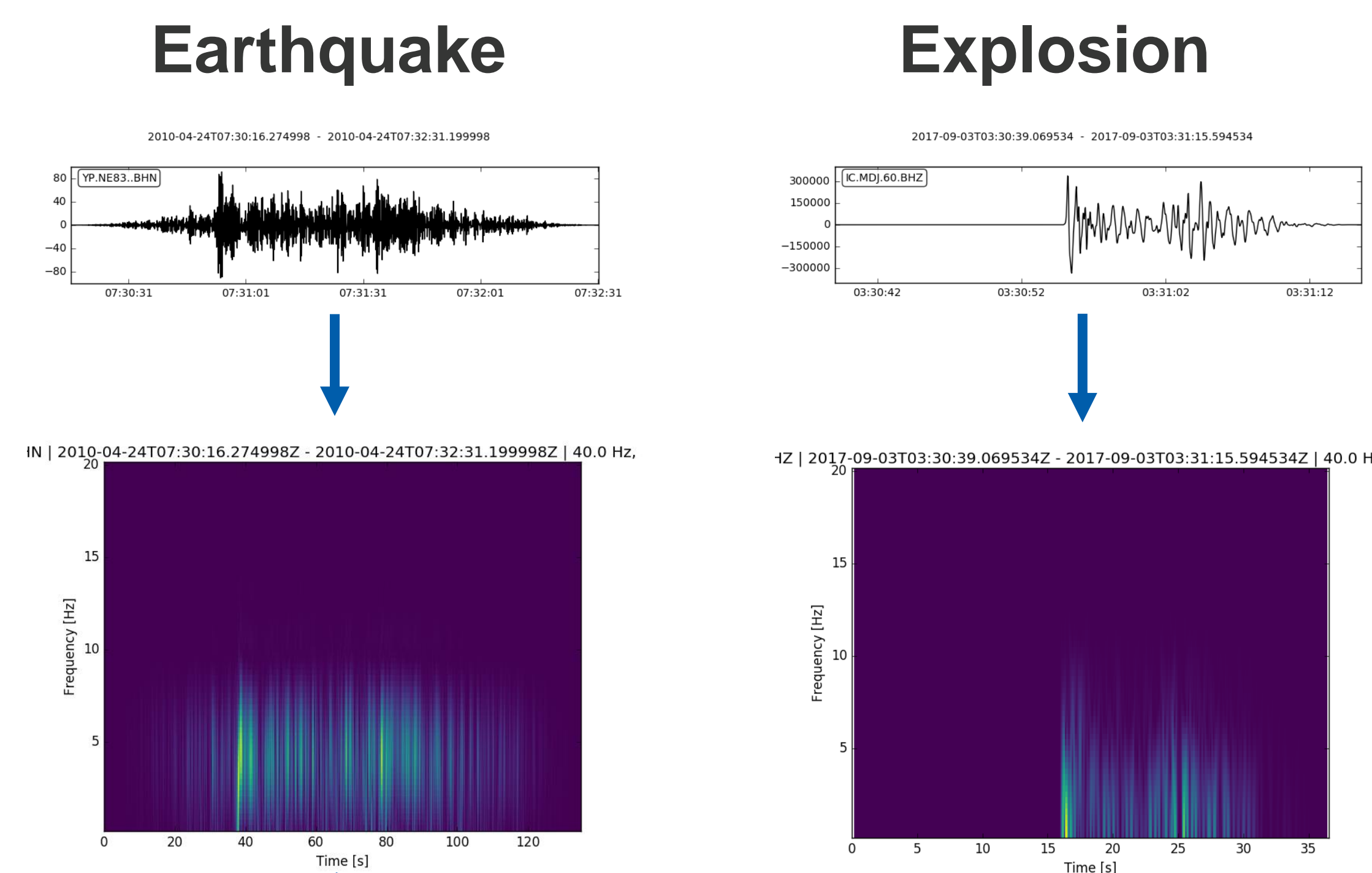
Long Short-term Memory (LSTM) Model

Seismograms that pass quality checks are converted to their spectrogram representations and used for the training and evaluation of a LSTM model.

Results

Confusion matrix
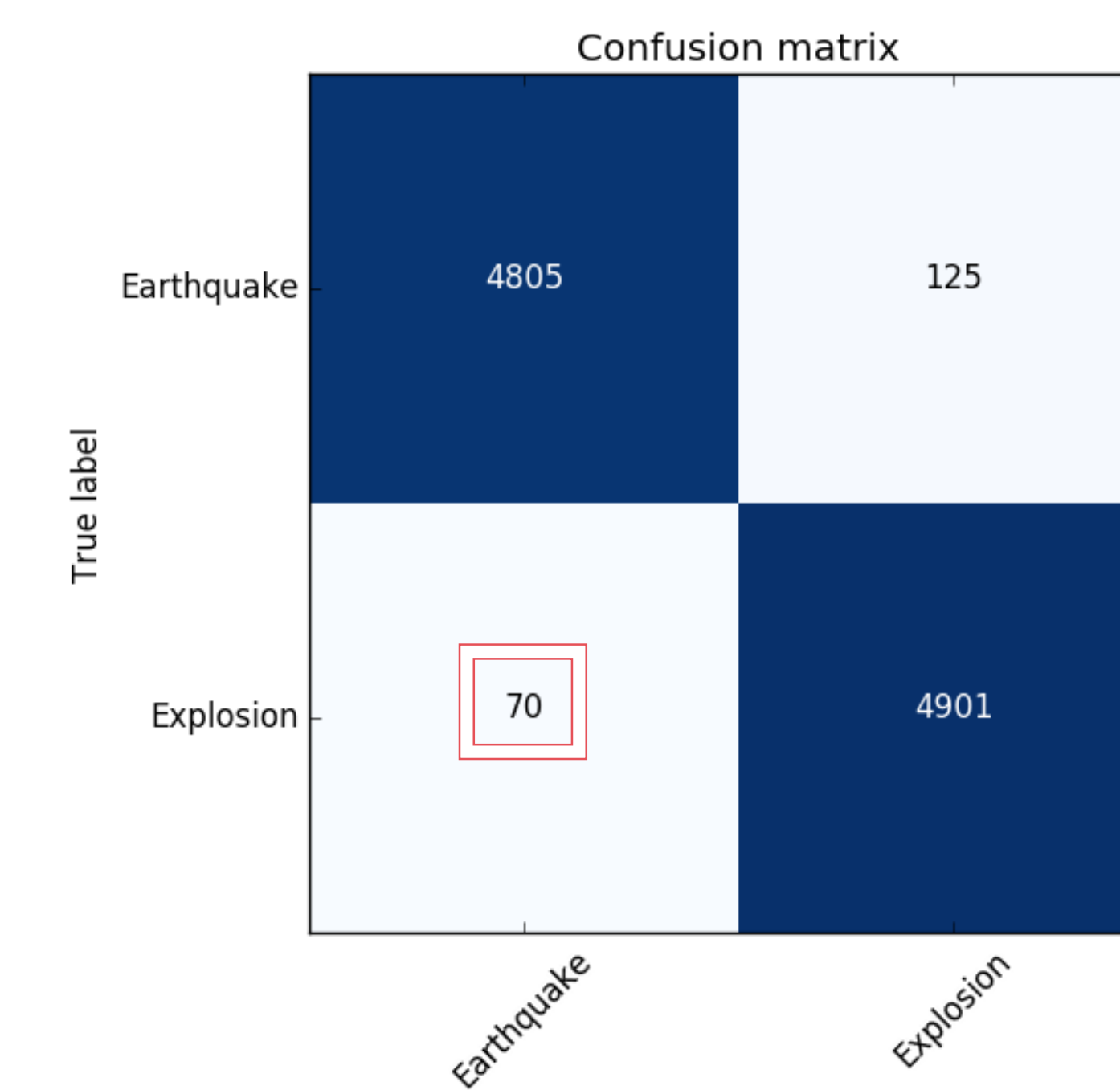
- Accuracy: 98%
- Precision: 98.56%
- Recall: 97.46%

While the standard classification metrics appear satisfactory, the labeling of explosions as earthquakes is problematic. For wider acceptance of this approach, additional research needs to be conducted to drive down this percentage.

CONCLUSIONS

- We created a machine learning model that can provide event discrimination at scale.
- Our model was created from a diverse global waveform dataset making it more robust than local and regional machine learning models.
- The resulting model does well on events of interest, however its accuracy is not high enough to handle millions of waveforms a day where false positives are mere annoyances, but false negatives are unacceptable.
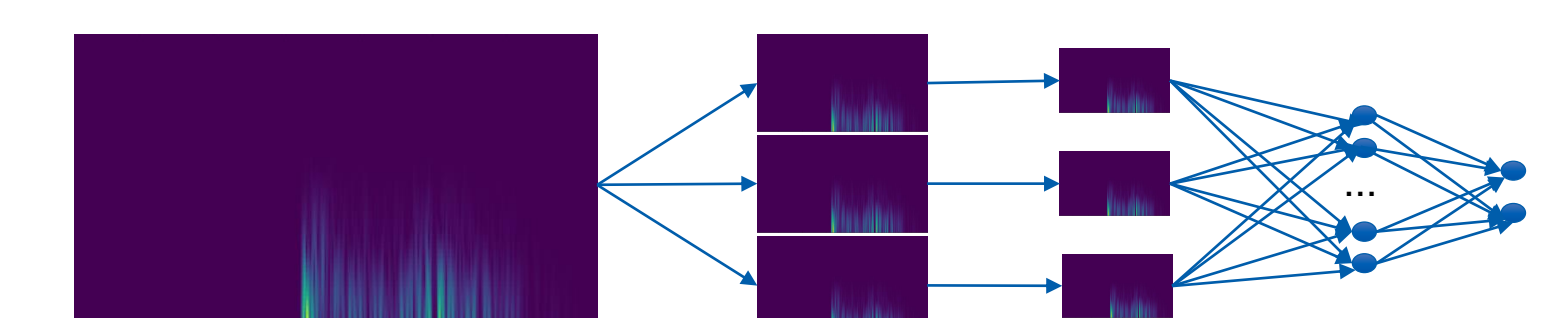
Future Work

Quality Control
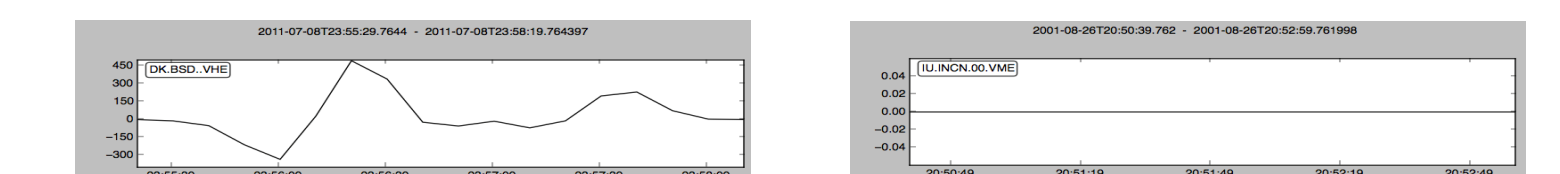
Data Augmentation

Convolutional Neural Networks

Machine learning is a tool that can solve the problem of nuclear explosion monitoring at scale.