An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network

Dario Jozinović
Anthony Lomax and Alberto Michelini (2019)
Introduction

- There is an increased need for rapid earthquake detection and characterization which is currently mainly provided by real-time analysis of seismogram waveforms using empirical and physics rule-based procedures.
- Extracting information from seismograms requires a lot of processing using rule-based seismological procedures.
- Development of ML techniques lead to application of ML to wide variety of waveform analysis problems.
- One of the most general ML approaches is the CNN which can be successfully applied to waveform analysis problems (Perol et al., 2018.)
**ConvNetQuake - Perol et al. (2018)**

- ConvNetQuake, a CNN for earthquake detection and location for earthquakes that originate from a single region
- Data: noise and local events of two stations in Oklahoma, USA
- Goal: noise/event and rough geographic area classification using single-station 3C data

<table>
<thead>
<tr>
<th></th>
<th>Autocorrelation</th>
<th>FAST</th>
<th>ConvNetQuake (this study)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>100%</td>
<td>88.1%</td>
<td>94.8%</td>
</tr>
<tr>
<td>Recall</td>
<td>77.5%</td>
<td>80.1%</td>
<td>100%</td>
</tr>
<tr>
<td>Event location</td>
<td>NA</td>
<td>NA</td>
<td>74.6%</td>
</tr>
<tr>
<td>accuracy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported runtime</td>
<td>9 days, 13 hours</td>
<td>48 min</td>
<td>1 min, 1 s</td>
</tr>
</tbody>
</table>
CNN comparison with classical methods

- The CNN can operate directly on the waveforms, with little pre-processing and without feature extraction such as energy detection, time-series transformation or frequency-domain analysis.
- The CNN architecture is shift invariant and so not sensitive to or dependent on the time position of features such as P and S wave arrivals in the waveform.
- Unlike standard regression, the CNN is not limited by assumed and simplified mathematical relations between quantities.
- The CNN does not make explicit use of existing knowledge and the question remains does it learn the general physics-based principles or it remains just a high-dimensional regression.
Work goals

- We extend the work of Perol et al. to detect and determine the location of global earthquakes at any distance over a large range of magnitudes and depths.
- The inputs are single-station, 50 seconds, 3 component 20 Hz waveforms from all the stations within the regional MedNet network.
- The outputs are classification event/noise and binned, probabilistic estimates of the distance, azimuth, depth and magnitude of an event.
- The goal is to examine how well the network can detect and characterize earthquakes and how well can it generalize to unseen data (regions).
Data

- Downloaded noise and event data for MedNet stations using FDSN web services
- Minimal pre-processing: trim event waveforms to start 5 sec prior to P arrival, normalize to global maximum of all 3 traces and store the normalization value so it can be appended to the input
The network

- The network ConvNetQuake_INGV is a CNN with 9 convolutional layers, each with 32 channels and half the number of features as the preceding layer, and two fully connected layers at the end.
- The final, fully connected layer, contains an event node, 50 distance nodes, 20 magnitude nodes, 20 depth nodes and 36 azimuth nodes.
- The ReLu activation function is used in CNN layers and softmax function is applied to the output fully connected layer class scores to obtain a properly normalized probability distribution over classification bins.
- For training, ConvNetQuake_INGV uses an L2-regularized cross-entropy loss (misfit) function and the Adam Optimizer algorithm.
 CovNetQuake_INGV Architecture

input data: 3 component waveforms

1001 samples

3 channels

convolution 1

501 features

conv 2

251 features

conv 3

126 features

conv 4

63 features

conv 5

32 features

conv 6

16 features

conv 7

8 features

conv 8

4 features

conv 9

2 features

flattening (reshape)

32 channels

65 features

fully connected 1

output class scores: 1 “no event”, 50 distance, 20 magnitude, 20 depth, 36 azimuth
Training

- For training: 15200 event and 10724 noise waveforms starting from year 2010 to year 2018; Validation: 1773 event and 1198 noise waveforms, same timespan
Testing

- Testing has been done on 1003 event and 621 noise events from 2009
Detection results

- detection accuracy (correct detections / total number of waveforms): 0.87
- detection precision (number of correct event detections / total number of predicted event detections): 0.97
- recall (number of correct event detections / total number of events): 0.81
- F1 score (2 * precision * recall / (precision * recall)): 0.88.
Results

- Test results are similar to, but notably more scattered than the validation results; validation events include many aftershock, swarm and other events with similar waveforms; suggests difficulty in generalizing.

Predicted vs True for validation data set classification

Predicted vs True for test data set classification
Analysis of epicentral errors

- Figure a): increase in distance error with distance; sharp and strong increase in maximum error beyond about 3.5° indicates reduced distance accuracy when there is no S arrival in the waveform window
- Figure b): Mild reduction of distance error with increasing training event density around true epicenter
Summary

- CNN for single station broadband seismograms over a large span of event types, distances, magnitudes and depths
- Performs very well in detecting events whose waveforms are dissimilar to any in the training data
- Event characterization does not show the same performance, with a tendency to overfit to areas with high event density, and with lot of outliers; there is, however, some generalization of the determination of the event parameters
- Not a practical monitoring tool, but allows investigation into rapid detection and characterization for short, single-station waveforms
Outlook

- Incorporate new machine learning techniques (CNN + LSTM...)
- Use additional metadata about the stations ($V_s 30$, $f_0$...)
- Use two or more stations
- Move to ground-motion prediction
Thank you for your attention
Citations

- FDSN web services at http://cnt.rm.ingv.it/webservices_and_software