Systematic comparison of different machine learning based earthquake detection methods

Ruide Ao

ABSTRACT

The overarching goal of earthquake detection methods is to estimate the timing of seismic phase arrivals at one or more recording stations. Earthquake detection is a fundamental problem in seismology, and therefore it is very important. It is the foundation for related techniques such as earthquake relocation and tomography. However, there are many challenges in detecting earthquakes, such as noise in the seismic data or S waves emerging inside the coda of P waves. Traditionally, earthquake detection has been performed through manual, human inspection of seismic waveforms. This manual analysis is often conducted with high precision but is limited by big data and also has man-made error. Some automation techniques can reduce these problems, for example techniques that include short-term average/long-term average (STA/LTA), template matching and autocorrelation. All of these techniques have their own drawbacks, and pre-processing data is usually necessary.

Recent developments in earthquake detection methods include state-of-art approaches using machine learning, such as PhaseNet and EQTranformer. These methods are trained on large volumes of quality-controlled, labelled P and S seismic data and have demonstrated promising results compared with traditional methods. The input for each are three-component seismic waveforms without the waveform pre-processing mentioned above, and the output is the probability of P or S waves at each time point. These methods can greatly improve temporal accuracy in the detection problem. Additionally, they reduce the human labor to pre-process data and set thresholds for different windows of a seismic record. We test the two methods using a new data set called the University of Utah Event Bulletin (UUEB), where local and regional earthquake arrivals in Utah have been closely inspected by human analysts, providing a means of ground-truthing the machine learning detection algorithms. This data set includes seismic waveforms recorded on broadband and short-period seismometers, accelerometers and other geophysical sensors. Results show that PhaseNet is good at picking P waves but not S waves and is prone to false detections. EQtransformer has nearly equal detection results both for P and S but much fewer detection results overall compared with PhaseNet on the data. Our results also highlight limitations in "ground truth", human-analyzed detections, demonstrating that neither traditional or state-of-the-art techniques can accomplish the earthquake detection task error-free.

Co-Advisor: Daniel Trugman

Co-Advisor: Steve Grand