London School of Economics and Political Science

Xixiang Hu, Zhiyu Li, Qixuan Wu, Wen Xiang, Jiangchen Zhao[1]

[1]Department of Statistics, London School of Economics and Political Science

Introduction

An earthquake early warning system limits the impact of future earthquakes on communities living nearby. The system detects an earthquake at its very beginning and rapidly issues alerts to users in its path. The alert outpaces strong earthquake shaking and may provide critical time to take basic protective actions such as seeking cover or exiting a building.
A seismic station is a device that records ground motion. In particular Grillo devices record the acceleration of the ground motion. The device uses a similar (albeit more precise) accelerometer sensor than you would find in consumer electronics such as smartphones. The device records the ground acceleration in three components - two horizontal ones and one vertical (x, y, z).

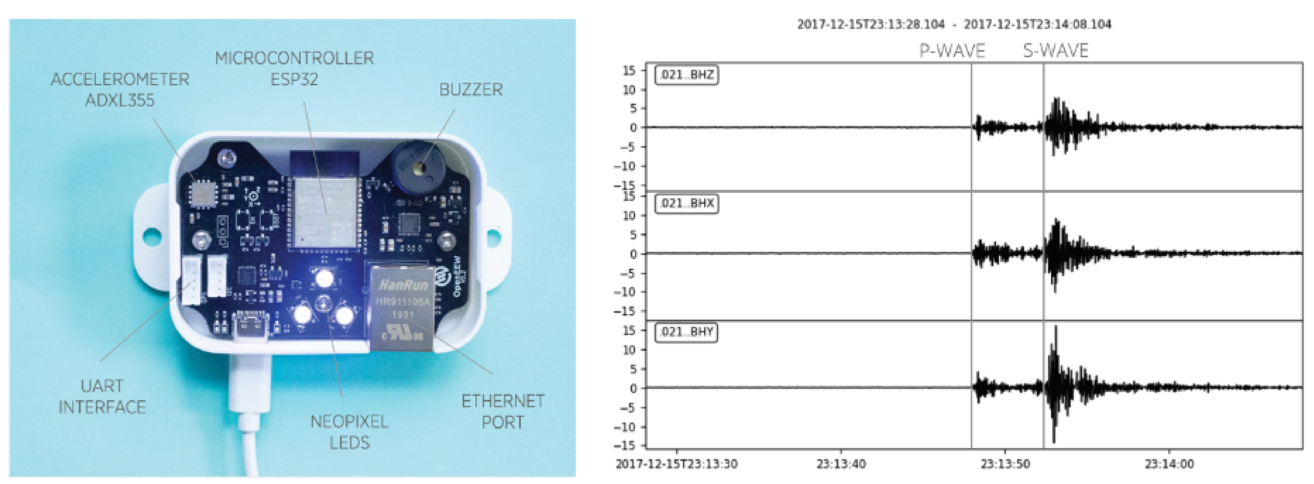

Fig. 1: Device and signal.

As mentioned above, once an earthquake occurs, it generates two kinds of waves that penetrate the Earth as body waves. Each wave has a characteristic speed and style of motion.

- P-wave / Primary wave: The primary wave the first seismic wave detected by seismographs due to high velocity in the rock environment between 5 and 8 km/s. The wave is usually relatively low-amplitude, does not carry much of the earthquake energy, and thus does not cause a significant damage.
- S-wave / Secondary wave: The secondary wave travels slower than the P-wave (3 to 5 km/s). However, it carries more energy and causes more damage than the P-wave.

EEW systems need to reliably detect the P-waves as this will allow for faster detections and therefore provide more time for the end users to take action. There is a number of traditional seismological algorithms that are usually based on rapid increase of [amplitude of ground motion](https://www.esgsolutions.com/technical-resources/microseismic-knowledgebase/event-detection-and-triggering). They are simple and effective, yet, they lack the ability to distinguish between signals created by earthquakes and other kind of disturbances (slamming doors, passing trucks..). In recent years, seismologists and data scientist have started to utilize neural-networks-based algorithms, hoping that those will be able to reduce the number of false positive detections.

Dataset

A network of about 20 Grillo instruments has been monitoring the southwest coast of Mexico since 2017. Grillo recorded segments with and without earthquake P-waves (a.k.a. signal and noise). And the task will be to train a model that can distinguish between them.

- Signal: Roughly 1,223 earthquake records from Grillo stations. Each record comes from a single seismic station, has 3 independent channels (x, y, z) and is 2000 data samples long. The records are centered at earthquake P-wave, which starts always at data sample 1,000.
- Noise: Randomly selected noise segments from the data and non-earthquake disturbances. There are 17,850 noise segments. Noise segments are also 2,000 data samples long.

Methodology

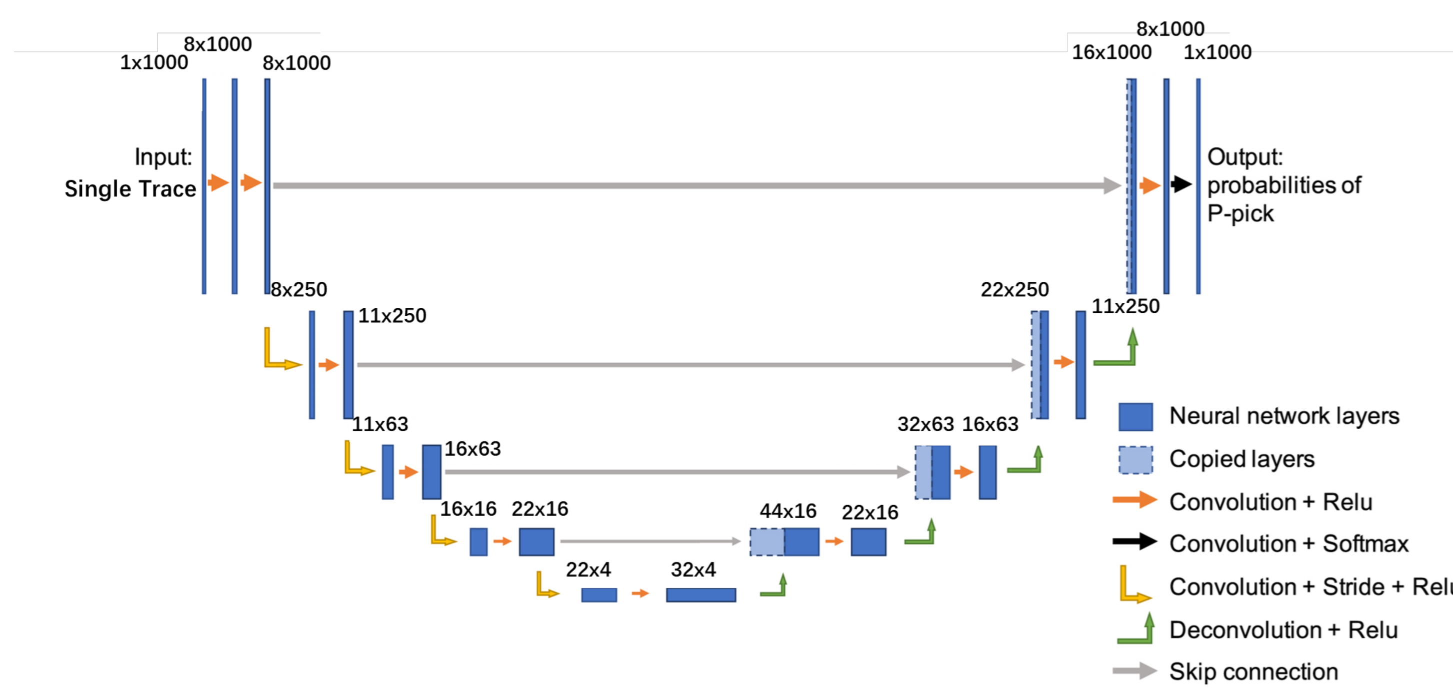

Fig. 2: Neural Network

The architecture of PhaseNet is modified from U-Net (Ronneberger et al. 2015) to deal with 1-D time-series data. The mapping to our problem is to localize the properties of our time-series into three classes: P pick, S pick and noise. The inputs are three-component seismograms of known earthquakes. The outputs are probability distributions of P wave, S wave and noise. In our experiments, the input and output sequences contain 1000 data points for each component (32 s long, sampled at 31.25 Hz). The P and S arrival times are extracted from the first time the result value passing through the threshold.

Focal loss is an improved version of cross-entropy which can handle the imbalance of classes. The formula of focal loss is:

$$FL(p_t) = -\alpha_t(1-p_t)^{\gamma}log(p_t)$$

γ can be used to adjust the penalty to the misclassification of hard class. The higher γ is, the heavier the penality will be. After several experiments, we find that $\gamma = 3$ gives the best performance in this case.

Result

Evaluation Metrics:
$F1score = 2 * (Precision * Recall)/(Precision + Recall)$ where
$Precision = (TruePositives)/(TruePositives + FalsePositives)$
$Recall = (TruePositives)/(TruePositives + FalseNegatives)$

	Noise	Earthquake
Noise	1778	2
Earthquake	4	106

$$F1Score = 0.9383886255924171$$

Conclusion

We have built a training data set using manually picked P and S arrival times from the data provided by Grillo. We have built our network based on PhaseNet, a deep neural network algorithm that uses three component waveform data to predict the probability distributions of P pick, S pick and noise.

References

[1] Weiqiang Zhu, Gregory C Beroza, PhaseNet: a deep-neural-network-based seismic arrival-time picking method, Geophysical Journal International, Volume 216, Issue 1, January 2019, Pages 261–273, https://doi.org/10.1093/gji/ggy423
[2] Jannes Münchmeyer, Jack Woollam, Andreas Rietbrock, Frederik Tilmann, Dietrich Lange, Thomas Bornstein, Tobias Diehl, Carlo Giunchi, Florian Haslinger, Dario Jozinović, Alberto Michelini, Joachim Saul, Hugo Soto. (2021). SeisBench - A Toolbox for Machine Learning in Seismology (v0.1.7). Zenodo. https://doi.org/10.5281/zenodo.5568813