Implementing EQTransformer and creating a simpler model for earthquake detection in Indian subcontinent

Saptarshi Datta
School of Physical Sciences
NISER, HBNI
Bhubaneswar
saptarshi.datta@niser.ac.in

Vanshaj Vidyan
School of Physical Sciences
NISER, HBNI
Bhubaneswar
vanshaj.vidyan@niser.ac.in

Abstract

This work explores the application of the EQTransformer model for the detection of earthquakes worldwide and India in particular; we put special emphasis on in-plate and cluster earthquakes. The model implements a attentive learning followed by a deep learning model to detect the *P* and *S waves* from seismic waveform data recorded at different stations and channels. We report the results obtained from the EQT applications on labelled datasets, along with technical studies done for future work on this project.

1 Introduction

Earthquakes are detected from seismic waveform data collected at geological stations throughout the world. These waveforms are the vibrations (displacement) of the earth's crust noted as a function of time at three coordinates called the *North (N)*, *East (E)*, and *Vertical (Z)*. Since the 1930s, trained analysts have identified such earthquake events by visual inspections of a time series plot of the waveform and notice initial onset times, waveform shapes, and arrival order to differentiate between P and S waves. While most of the seismic data have been digitised and made publicly available in other countries, India has been lagging behind; scientific personnel still sit through thousands of hours of data and visually identify and label individual earthquake events. Despite advancements in ML and AI models for this purpose, most of them are largely trained on plate-edge earthquakes and never on Indian datasets.

As the first part of our semester project, we aimed to implement the acclaimed deep learning model EQTransformer [2]. It has been trained on the international database Stanford Earthquake Database (STEAD) and claims to be performing well on earthquakes from all parts of the world. Since there are some *features* whose values vary at different parts of the world and EQT has not been trained on Indian datasets, we plan to check its efficiency on labelled Indian datasets. The motivation for our project comes from the fact that this will greatly ease the work of PhD students in India and will facilitate better focus on the physics analysis behind earthquakes. Our particular contribution is towards the study of cluster earthquakes which are harder to detect as several hundreds of earthquakes occur in a few minutes' interval. Such events at Palghar, Maharashtra, are of interest to us, and our work will allow for a better understanding of the earthquake mechanics there.

2 EQTransformer

The model has two units: the convolutional unit [1] and the Long-Short Term Memory unit. Both are neural network models here used for two purposes. The first one (CNN) is used for the detection of a possible zone in the input waveform where an earthquake is present (a blanket detection). Then that particular section is input into the attention model [4] first to detect the *P-wave* and then the section

Preprint. Under review.

Table 1 Detection performance.

Model	*Pr*	*Re*	*F1*	*Training data*	*Training size*	*Ref.*
EQTransformer	**1.0**	**1.0**	**1.0**	Global	1.2M	This study
CRED	**1.0**	0.96	0.98	Global	1.2M	7
DetNet	**1.0**	0.89	0.94	China	30K	5
Yews	0.84	0.85	0.85	Taiwan	1.4M	4
STA/LTA	0.91	**1.0**	0.95	—	—	11

Pr, Re, and F1 are precision, recall, and F1-score respectively. EQTransformer and CRED have been trained on STEAD dataset while DetNet and Yews results are based on pre-trained models on different datasets. Recursive STA/LTA algorithm is used here.
Bold values represent the best performance.

Figure 1: Detection performance of EQTransformer [2]

after it for *S-wave*. The model has been trained, validated and tested on portions of STEAD. They used 1M earthquake and 300K noise waveforms net associated with about 450 K earthquakes. The data was randomly split into 85% for training, 5% for validation and 10% for testing. The training was finished in $\mathcal{O}(89)$ hours when validation loss did not improve over 12 consecutive epochs.

The trained model was applied to just one-third of the Tottori earthquakes' dataset and detected twice the manual detections in the overall waveform dataset. As the Tottori data from Japan was not in the STEAD, the EQT was claimed to be of universal use. They demonstrated that EQT can detect earthquakes from different geographical locations with different magnitudes, wave shapes, and wave durations. This is justified by the very deep nature of the model.

3 Report

This section summarizes the work done till the mid-semester.

3.1 Targets Achieved

1. Extensive literature review completed, focused on two aspects:
   - Domain knowledge: Learning the principal *features* used for the manual annotation of seismic waveforms.
   - ML knowledge: Learning the logic behind the codes in the EQTransformer package. Further an overview study on other acclaimed and efficient approaches to earthquake detection. [3]
2. Numerous dependency clashes were resolved in the installation of the EQTransformer package.
3. EQTransformer applied successfully on international data on the Lingaraj system, positive results successfully replicated.
4. We are in the process of obtaining the novel labelled Indian dataset on the Palghar cluster in-plate earthquakes from Dr. Pathikrit's Lab at the *School of Earth and Planetary Sciences, NISER*.

3.2 Methodology

In a nutshell, the EQTransformer package works in the following manner. It is illustrated with code snippets followed by the results (as per our workflow).

First the code snippet:

```
import os
json_basepath = os.path.join(os.getcwd(),"json/station_list.json")

from EQTransformer.utils.downloader import makeStationList

makeStationList(json_path=json_basepath, client_list=["SCEDC"],
    min_lat=35.50, max_lat=35.60, min_lon=-117.80, max_lon=-117.40,
    start_time="2019-09-01 00:00:00.00", end_time="2019-09-03
    00:00:00.00", channel_list=["HH[ZNE]", "HH[Z21]", "BH[ZNE]"],
    filter_network=["SY"], filter_station=[])
```

```
%%%%%%%%%%
from EQTransformer.utils.downloader import downloadMseeds

downloadMseeds(client_list=["SCEDC", "IRIS"], stations_json=
    json_basepath, output_dir="downloads_mseeds", min_lat=35.50,
    max_lat=35.60, min_lon=-117.80, max_lon=-117.40, start_time="
    2019-09-01 00:00:00.00", end_time="2019-09-03 00:00:00.00",
    chunk_size=1, channel_list=[], n_processor=2)
```

This creates a list of seismic stations between the desired longitudes and latitudes and specifies the desired range of download *(start and end time)*. Also, the target directory is specified, and a .mseed file is downloaded.

For instance, a *.mseed* file can be visualised in a waveform structure which looks like (only z-axis data is plotted here, as an example): Then,

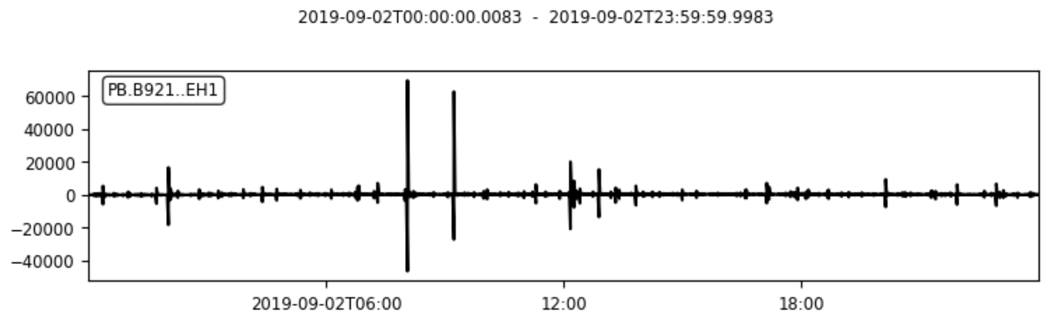

Figure 2: Waveform data (z-axis) downloaded from B921 seismic station in California

```
from EQTransformer.utils.hdf5_maker import preprocessor

preprocessor(preproc_dir="preproc", mseed_dir='downloads_mseeds',
    stations_json=json_basepath, overlap=0.3, n_processor=2)
```

The above snippet divides the total downloaded (continuous) data from *.mseed* formats into 60-second strips in *.hdf5* formats with a sampling rate of 100 Hz, so it can be fed into the earthquake transformer model. Finally, the code snippet:

```
from EQTransformer.core.predictor import predictor

predictor(input_dir= 'downloads_mseeds_processed_hdfs', input_model='
    EqT_model.h5', output_dir='detections', detection_threshold=0.3,
    P_threshold=0.1, S_threshold=0.1, number_of_plots=100, plot_mode='
    time')
```

This uses the predictor module in the EQTransformer and feeds in the preprocessed *.hdf5* files. Resultantly, the EQT model outputs the detections made in the 60-second waveform strip in the form of both a *.csv* file as well as a .png file, both of which are shown below in Figs 3 and 4. Finally, the model also outputs the final result for the full duration of the input data (here, one day) in the form of a *.txt* file, which is shared below in Fig 5.

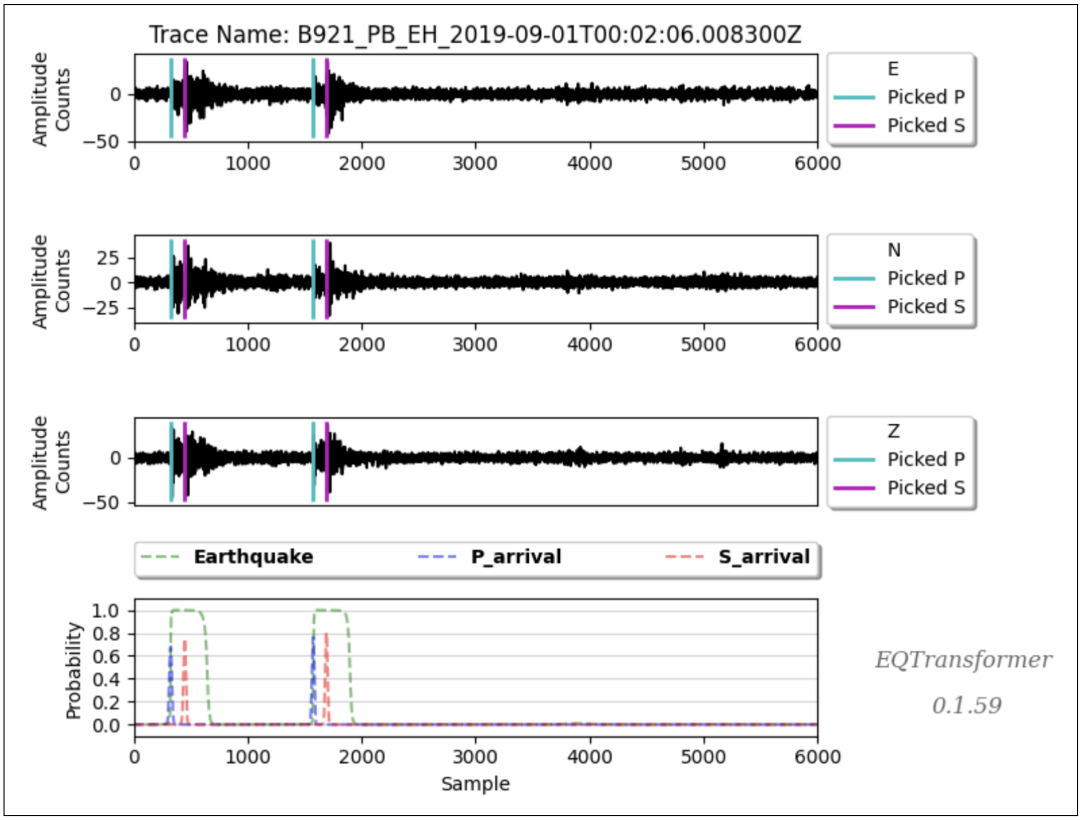

Figure 3: A one-minute long sample with EQT-annotated detections

```
B921_PB_EH_2019-09-01T00:02:48.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:03:27.488300,2019-09-01
00:03:30.388300,0.96,,2019-09-01 00:03:27.488300,0.87,,30.3,2019-09-01 00:03:28.698300,0.85,,7.1
B921_PB_EH_2019-09-01T00:03:30.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:04:10.568300,2019-09-01
00:04:14.138300,0.93,,2019-09-01 00:04:10.678300,0.57,,3.7,2019-09-01 00:04:12.068300,0.82,,8.5
B921_PB_EH_2019-09-01T00:04:54.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:05:08.688300,2019-09-01
00:05:20.028300,0.76,,2019-09-01 00:05:09.008300,0.81,,0.7,2019-09-01 00:05:13.238300,0.77,,12.3
B921_PB_EH_2019-09-01T00:05:36.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:06:18.048300,2019-09-01
00:06:23.408300,0.96,,2019-09-01 00:06:18.048300,0.77,,30.4,2019-09-01 00:06:20.218300,0.63,,10.1
B921_PB_EH_2019-09-01T00:06:18.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:06:20.308300,2019-09-01
00:06:21.048300,0.43,,2019-09-01 00:06:20.018300,0.1,,10.9,2019-09-01 00:06:20.358300,0.33,,9.7
B921_PB_EH_2019-09-01T00:07:00.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:07:14.088300,2019-09-01
00:07:21.698300,0.48,,2019-09-01 00:07:17.008300,0.13,,1.7,2019-09-01 00:07:17.438300,0.12,,-0.2
B921_PB_EH_2019-09-01T00:07:00.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:07:29.048300,2019-09-01
00:07:36.848300,0.54,,2019-09-01 00:07:32.298300,0.18,,1.7,2019-09-01 00:07:32.688300,0.19,,-0.2
B921_PB_EH_2019-09-01T00:07:00.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:07:54.298300,2019-09-01
00:07:59.998300,0.68,,,,,1.7,2019-09-01 00:07:58.228300,0.41,,-0.2
B921_PB_EH_2019-09-01T00:07:42.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:08:08.118300,2019-09-01
00:08:15.598300,0.74,,2019-09-01 00:08:08.058300,0.34,,0.1,2019-09-01 00:08:10.508300,0.41,,4.9
B921_PB_EH_2019-09-01T00:07:42.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:08:27.708300,2019-09-01
00:08:40.098300,0.7,,2019-09-01 00:08:27.778300,0.58,,0.1,2019-09-01 00:08:29.728300,0.7,,4.9
B921_PB_EH_2019-09-01T00:08:24.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:09:03.288300,2019-09-01
00:09:14.918300,0.75,,2019-09-01 00:09:03.448300,0.22,,21.7,2019-09-01 00:09:05.718300,0.64,,13.2
B921_PB_EH_2019-09-01T00:09:48.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:10:01.088300,2019-09-01
00:10:20.098300,0.85,,2019-09-01 00:10:01.008300,0.92,,8.8,2019-09-01 00:10:02.528300,0.13,,1.5
B921_PB_EH_2019-09-01T00:09:48.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:10:40.048300,2019-09-01
00:10:47.998300,0.77,,,,,8.8,2019-09-01 00:10:45.778300,0.29,,1.5
B921_PB_EH_2019-09-01T00:10:30.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:10:44.288300,2019-09-01
00:10:48.238300,0.83,,2019-09-01 00:10:45.008300,0.76,,6.2,2019-09-01 00:10:45.918300,0.7,,5.6
B921_PB_EH_2019-09-01T00:10:30.008300Z,PB,B921,EH,35.5865,-117.4622,694.5,2019-09-01 00:10:54.648300,2019-09-01
00:10:58.628300,0.83,,2019-09-01 00:10:54.698300,0.34,,6.2,2019-09-01 00:10:56.158300,0.7,,5.6
```

Figure 4: A snippet of the .csv output for the station

3.3 Future Plan

Going forward, our work plan will be as follows. First, we will test the prediction performance of EQTransformer on Palghar data (around 1200 labelled) and other earthquake datasets obtained from Dr. Pathikrit's collaboration with NGRI. Next, we will explore possible ways to increase our data set

```
================== Overal Info ==============================
date of report: 2024-03-04 18:59:42.311121
input_hdf5: downloads_mseeds_processed_hdfs/B921.hdf5
input_csv: downloads_mseeds_processed_hdfs/B921.csv
input_model: EqT_model.h5
output_dir: /home/group4/detections/B921_outputs
================== Prediction Parameters =======================
finished the prediction in:  0 hours and 1 minutes and 12.28 seconds
detected: 5282 events.
writting_probability_outputs: False
loss_types: ['binary_crossentropy', 'binary_crossentropy', 'binary_crossentropy']
loss_weights: [0.03, 0.4, 0.58]
batch_size: 500
================== Other Parameters =========================
normalization_mode: std
estimate uncertainty: False
number of Monte Carlo sampling: 5
detection_threshold: 0.3
P_threshold: 0.1
S_threshold: 0.1
number_of_plots: 100
use_multiprocessing: True
gpuid: None
gpu_limit: None
keepPS: True
spLimit: 60 seconds
```

Figure 5: The output .txt file for the specified one-day duration in the B921 station

of Indian earthquakes by going for plate-edge earthquakes of the Himalayan region (tentative) or by oversampling algorithms supplied with existing labelled datasets. In case that becomes unfeasible, we will take earthquake datasets from international geographical locations which are *'geophysically similar'* to that of Indian subcontinent tectonics. Further, we will perform hyperparameter tuning on the assorted Indian (and/or India-like) dataset to improve the EQT model and check its prediction accuracy. Then we will pick a simpler ML model for waveform identification and build a rudimentary earthquake detector.

References

[1] S. M. Mousavi and G. C. Beroza. A machine-learning approach for earthquake magnitude estimation. *Geophysical Research Letters*, 47(1), 2019.

[2] S.M. Mousavi, W.L. Ellsworth, and W. Zhu. Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, 11, 2020.

[3] Saad and Chen. Earthquake detection and p-wave arrival time picking using capsule neural network, 2021.

[4] A Vaswani et. al. Attention is all you need. Aug 2017.