LLNL-TR-833478

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Preliminary Transfer Learning Results on Israel Data

Q. Kong, A. Price, S. Myers

April 1, 2022

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Preliminary Transfer Learning Results on Israel Data

Qingkai Kong, Amanda Price, Stephen Myers

2022-03-31

Summary

In this preliminary report, we use publicly available data recorded in Israel to test and expand upon existing machine learning models for seismic-phase detection and arrival-time measurement. We downloaded 3-years of waveform data from Geofon, and cross referenced the waveforms to Israel bulletin picks (Schardong et al., 2021). The initial results using existing models directly generated ubiquitous false detections and that obscured detections of signals that are clearly visible in the waveforms. However, after applying *transfer learning* (tuning parameters in the existing ML models using one year of the Israel-network data), the results are encouraging, i.e. ML picks agree within a few tenths of a second with bulletin picks and the number of false detections is greatly reduced. The bulletin picks are a good starting point, but they cannot be considered ground-truth. To test potential improvement in picking using ML we would like to relocate the events using the ML picks to see if the events cluster more tightly at known mine locations. However, in order to constrain event locations, we need ML picks for the whole Israeli-Jordanian network, which requires waveforms that are not publicly available.

Method

We tested three existing algorithms that were trained on the STEAD dataset (Mousavi et al., 2019), i.e. EQTransformer (EQT) (Mousavi et al., 2020), PhaseNet (PN) (Zhu & Beroza, 2019), and Generalized Seismic Phase Detection (GPD) (Ross et al., 2018). We used the models trained on this same STEAD dataset to make sure the performance can be compared. The results of these three algorithms applied directly to the data from Israel, and a representative example is shown in figure 1. We can see the results are not ideal. EQT misses the obvious P and S phases. Both PN and GPD detected the phase, but with many false detections. This is due to data from Israel having slightly different characteristics than the STEAD dataset. Since we don't have sufficient Israel data to develop a new ML detection/picking model, we use Transfer Learning (TL), which has shown great performance in many different domains (Tan et al., 2018).

Transfer learning is a method in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. In this case, the ML models are developed using a large seismic dataset with good labels (the STEAD dataset). Although STEAD has different characteristics than Israel data, the fundamental features needed to detect P-waves and S-waves and determine arrival time are identified. By using a small dataset from Israel, we can use the features that were identified using the STEAD data and fine-tune the model (instead of developing a model from scratch) based on new characteristics in Israeli data. As a first step, we use PhaseNet as the base model for transfer learning.

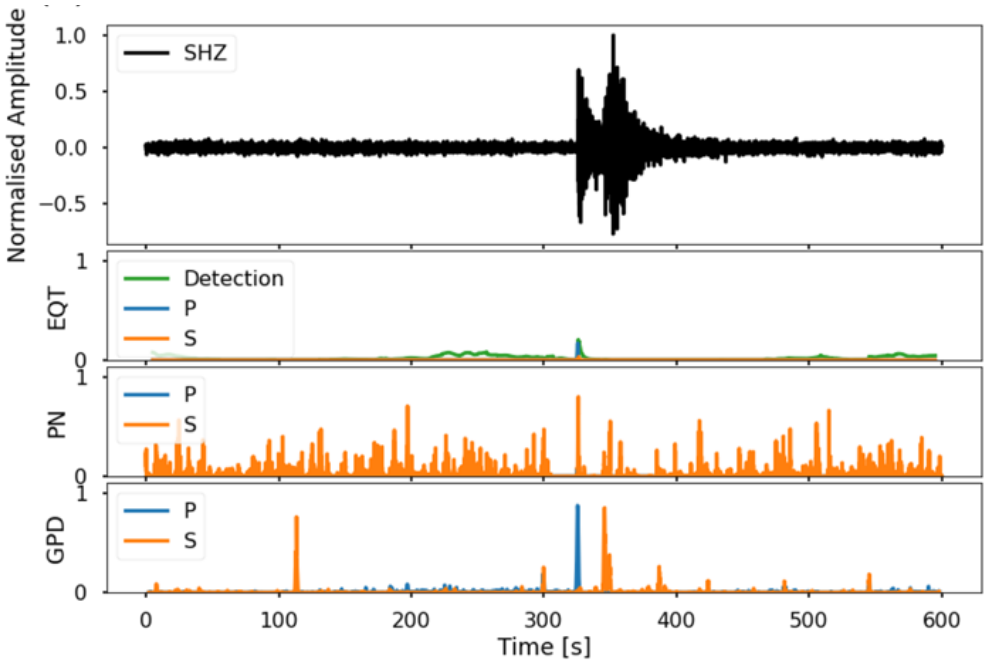

Figure 1. Test of the three different algorithms directly on data from Israel. The bottom three panels show the P (blue) and S (orange) estimated probabilities.

Data

Catalog files from the Schardon et al. (2021) study contain the origins of the events and the arrivals:

- FILE1: origins_ned_stats_max2.5_renum.out
- FILE2: arrivals.out

We downloaded data from the Israelian Broadband Seismological Network (Network code IS) for 3 years by windowing stations according to the origin of the events listed in FILE1. The beginning of the window is set to 300 s before the origin time. The end of the window is calculated using a travel-time velocity of 1 km/s and the estimated event-station distance, with a majority of window lengths being approximately 10 minutes. The downloaded data details are showing in Table 1.

Year	BH*	SH*	HH*	N stations	Total	Matched Phases
2016	11735	8073	346	31	20154	433
2017	13036	7520	716	28	21272	300
2018	49023	34007	2101	32	85131	4,848

Table 1. Details for data downloaded for 3 years. BH*, SH* and HH* columns are number of 3-components waveforms for each sensor type. N stations column shows the number of unique stations for the year. Total column shows the total number of waveforms downloaded for the year. The matched phases show the number of downloaded waveforms that have corresponding matched P and S phases in FILE2, which we can use for training model purposes.

We assembled the following dataset based on the matched phases waveforms shown in Table 1 for training and testing the model.

- Training data: 4,848 (2018)
- Validation data: 300 (2017)
- Testing data: 433 (2016)

There are multiple reasons why we only have a very small number of waveforms matched the phases (1) there are some stations missing (cannot find using the IS network code). (2) due to some of the events in FILE1 cannot find corresponding event listed in ISC. (3) also I have some quality control, such as the number of data points shorter than I cut, I will not use it.

Transfer Learning Results

Figure 2 shows one example comparing the results from our transfer learning model, as well as the original model that was trained using the STEAD dataset. We can clearly see the transfer learning-based model performs better, while the STEAD trained model missed the P wave.

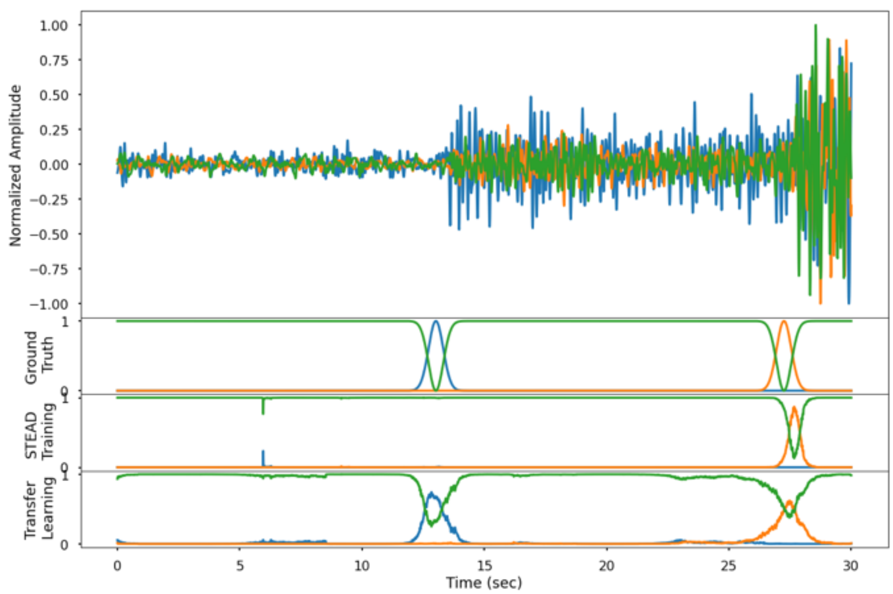

Figure 2. Example of the P and S pickings from different models. The top panel overlays the 3 components seismic data (normalized); 2nd row shows the bulletin (labeled ground truth) label from the arrival file, where blue, orange and green are probabilities for P, S and noise; 3rd row shows the estimation from the original PhaseNet data that was trained using STEAD data, and the last row shows the transfer learning result.

Application of the two models to longer waveforms is shown in Figure 3. Like Figure 2, we find that the transfer learning-based model detects and picks P and S waves more accurately and with few false detections. The model trained only using STEAD data has multiple (false) S picks at times when only noise exist. But the transfer-learning-based model does not have the multiple false picks.

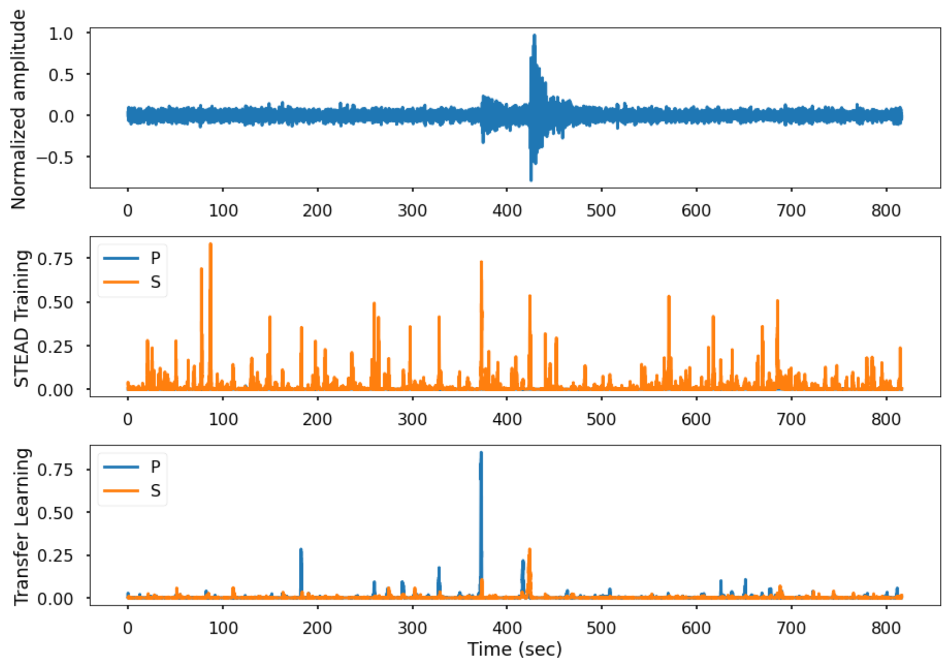

Figure 3. Model performance for a 10+ minute time window. Top: normalized vertical component time series. Middle: P and S estimated probabilities from the original model trained using the STEAD dataset. Bottom: P and S estimated probabilities from our transfer learning model trained on 4848 3-components waveforms from Israel.

	FP_1	FP_2	FN	TP	TN	Prec.	Recall
ST_P	22	45	110	189	67	0.74	0.63
TL_P	7	20	52	272	82	0.91	0.84
ST_S	16	5	74	277	61	0.93	0.79
TL_S	5	1	53	302	72	0.98	0.85

Table 2. Model performance table for 433 test detections in 2016. FP_1: False positive 1, algorithm detects phases when there is no bulletin label. FP_2: False positive 2, the estimated time is larger than 3s from the bulletin label. FN: False negative. TP: True positive. Prec.: Precision. Recall: Recall. ST_P and ST_S are the PhaseNet model trained on STEAD for P and S wave picks. TL_P and TL_S are the transfer learning model for P and S picks.

In order to systematically evaluate the model performance, we used the 2016 data that have P and S labels from the arrival files that are also listed in the ISC bulletin. Basic metrics such as false positive, false negative, true positive, true positive, as well as precision and recall values are shown in Table 2. For both P and S picks, the transfer learning outperforms the model trained only using STEAD, i.e. the precision and recall are 0.91 and 0.84 for the P picks for our transfer learning model, increased from 0.74 and 0.63 respectively, while the values for S picks are 0.98 and 0.85, increased from 0.93 and 0.79.

Figure 4 shows the time difference distribution between the two models' estimations and the labels from the arrival file for both P and S wave picks. From the mean and standard deviation values, we can see the transfer learning-based model performs much better, especially for the P wave.

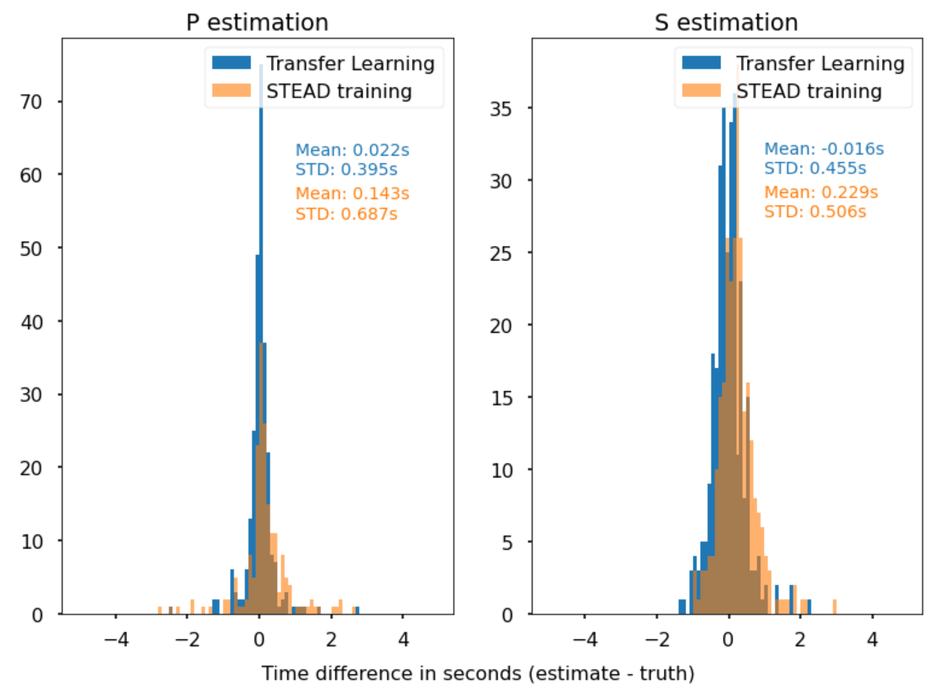

Figure 4. Time difference between the estimated picks and the arrivals from the file (assumed ground truth) on the 2016 test data. Blue histogram is for the transfer learning-based model while the orange histogram is for the model trained with STEAD data. The mean and standard deviation values are listed using the same colors.

Data request for further analysis

Initial work demonstrates the applicability of transfer learning to detection and arrival-time estimation seismic phases recorded in Israel. The publicly available PhaseNet package performs pooly, but performance is dramatically improved after transfer learning (parameter tuning). In order to further test and develop our transfer learning model, we need data from the entire Israeli and Jordanian network. A refined data set of phase labels and arrival times – if available – would also be of great utility. The data we downloaded contains a subset of stations matching the arrival files, thus only a small number of picks have corresponding waveforms. We tried to associate the estimated picks and relocate events but accurate locations are not possible with the sparse network. Thus we request the following for further development if it is possible:

1. Continuous waveforms for the Israel-Jordan network for 1 or 2 years
2. Good quality labels, i.e. manually picked phases. This is not high priority, the data we currently have can be used to develop a good model. But if we have this data, it can help us to (a) develop better models with more training data, (b) more accurate evaluation against these labels.

References

Mousavi, S. M., Sheng, Y., Zhu, W., & Beroza, G. C. (2019). STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI. *IEEE Access*, *7*, 179464–179476. https://doi.org/10.1109/ACCESS.2019.2947848

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, *11*(1), 3952. https://doi.org/10.1038/s41467-020-17591-w

Ross, Z. E., Meier, M., Hauksson, E., & Heaton, T. H. (2018). Generalized Seismic Phase Detection with Deep Learning. *Bulletin of the Seismological Society of America*, *108*(5A), 2894–2901. https://doi.org/10.1785/0120180080

Schardong L., Y. Ben-Horin Y., A. Ziv, S. C. Myers, H. Wust-Bloch, Y. Radzyner (2021). Revised Seismic Bulletin for Israel, *Seis. Res. Lett*, 92 (4): 2668–2678. https://doi.org/10.1785/0220200422.

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A Survey on Deep Transfer Learning. In V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, & I. Maglogiannis (Eds.), *Artificial Neural Networks and Machine Learning – ICANN 2018* (pp. 270–279). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-01424-7_27

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, *216*(1), 261–273. https://doi.org/10.1093/gji/ggy423

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract Number DE-AC52-07NA27344. This is LLNL Contribution Number LLNL-TR-XXXXX.