Transfer Learning for Seismic Phase Picking on Different Geographic Regions

Chenyu Li[1], Lijun Zhu[1], Dongdong Yao[2], Xiaofeng Meng[3], Zhigang Peng[1], James McClellan[1], and Jake Walter[4]

[1]Georgia Institute of Technology
[2]University of Michigan
[3]Southern California Earthquake Center
[4]Oklahoma Geological Surver

November 22, 2022

Abstract

Machine learning algorithms have become a powerful tool in different areas of seismology, such as phase picking/earthquake detection, earthquake early warning and focal mechanism determination. Previously convolutional neural networks (CNN) have been applied to continuous seismic waveform recordings to perform efficient phase picking and event detection with good accuracy [Zhu et al., 2018]. However, the off-line training of current CNN requires at least a few thousands of accurately picked seismic phases, which makes it difficult to be applied to regions without sufficient picked phases. In this work, we will validate the transfer learning among different geographic regions. Our tests show that the phase picker trained on manually-labeled data acquired from Sichuan, China following the 2008 M7.9 Wenchuan earthquake [Zhu et al., 2018] works equally well on the continuous waveform acquired from Oklahoma, US [Zhu et al., 2018]. Specifically, using the CNN trained on the Wenchuan dataset, together with 895 local/regional catalog events recorded in central Oklahoma, we refine part of the networks to pick the arrival times of the local seismicity in Oklahoma. The refined CNN results are compatible with the matched filter results using the same catalog events as templates. Our next step is to extend our test to waveforms from different tectonic regions to demonstrate the generality of CNN-based phase picker. We plan to further use a New Zealand seismic dataset that includes more than 20 GeoNet stations in the North Island, where the matched-filter detected results are available to be compared with (Yao et al., 2018). Alternatively dataset include a subset of events in the waveform relocated catalog in Southern California. Updated results will be presented at the meeting.

Transfer Learning for Seismic Phase Picking on Different Geographic Regions

Chenyu Li[1], Lijun Zhu[2], Dongdong Yao[3], Xiaofeng Meng[4], Zhigang Peng[1], James McClellan[2], Jake Walter[5]

[1] School of Earth & Atmospheric Sciences, Georgia Tech [2] School of Electrical and Computer Engineering, Georgia Tech
[3] University of Michigan [4] University of Southern California [5] Oklahoma Geological Survey

Poster ID: S11E-0431

Introduction

Convolutional neural networks (CNN) is a powerful tool of feature extractor, which have been recently applied to seismic waveform recordings to perform phase picking, and is proved to be efficient (i.e., Zhu et al. 2017). For regions with a large amount of labeled data, it is easy to train a CNN from scratch. But how about regions with much less labeled data? Inspired by the success of applying CNN to seismic-phase picking in aftershocks of 2008 M7.9 Wenchuan earthquake in China, and the idea of transfer learning, we decide to verify **transfer learning** in seismic phase-picking and event detection. We apply CNN trained from a larger dataset to regions with smaller dataset. In this case, we slightly modify the CNN trained in Wenchuan dataset to smaller datasets in Oklahoma, New Zealand and Southern California. If **transfer learning** works well for seismic phase picking in different regions, we would have a more powerful and universal event-detection method, even in regions with limited information of known events.

Method

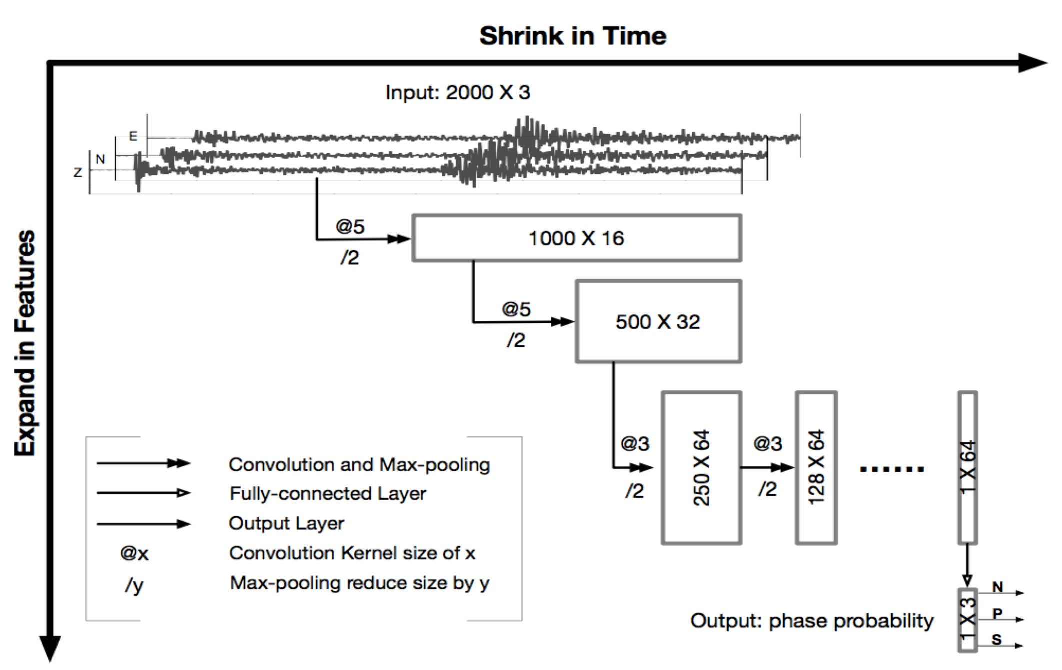

Figure 1. The CNN network structure for seismic phase detection and picking [Zhu et al. 2018, in revision].

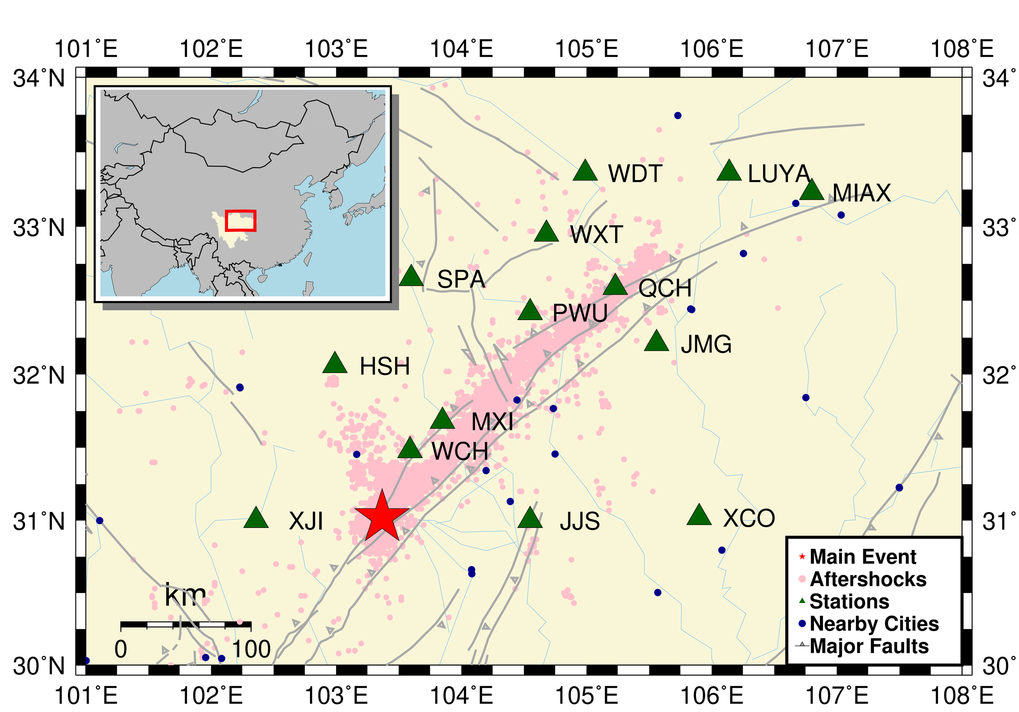

Figure 2. Location of events and stations used to train the initial CNN with aftershocks of 2008 M7.9 Wenchuan Earthquake., red star shows epicenter of mainshock.[Zhu et al., 2018, in revision].

Test 1: from Wenchuan to Oklahoma

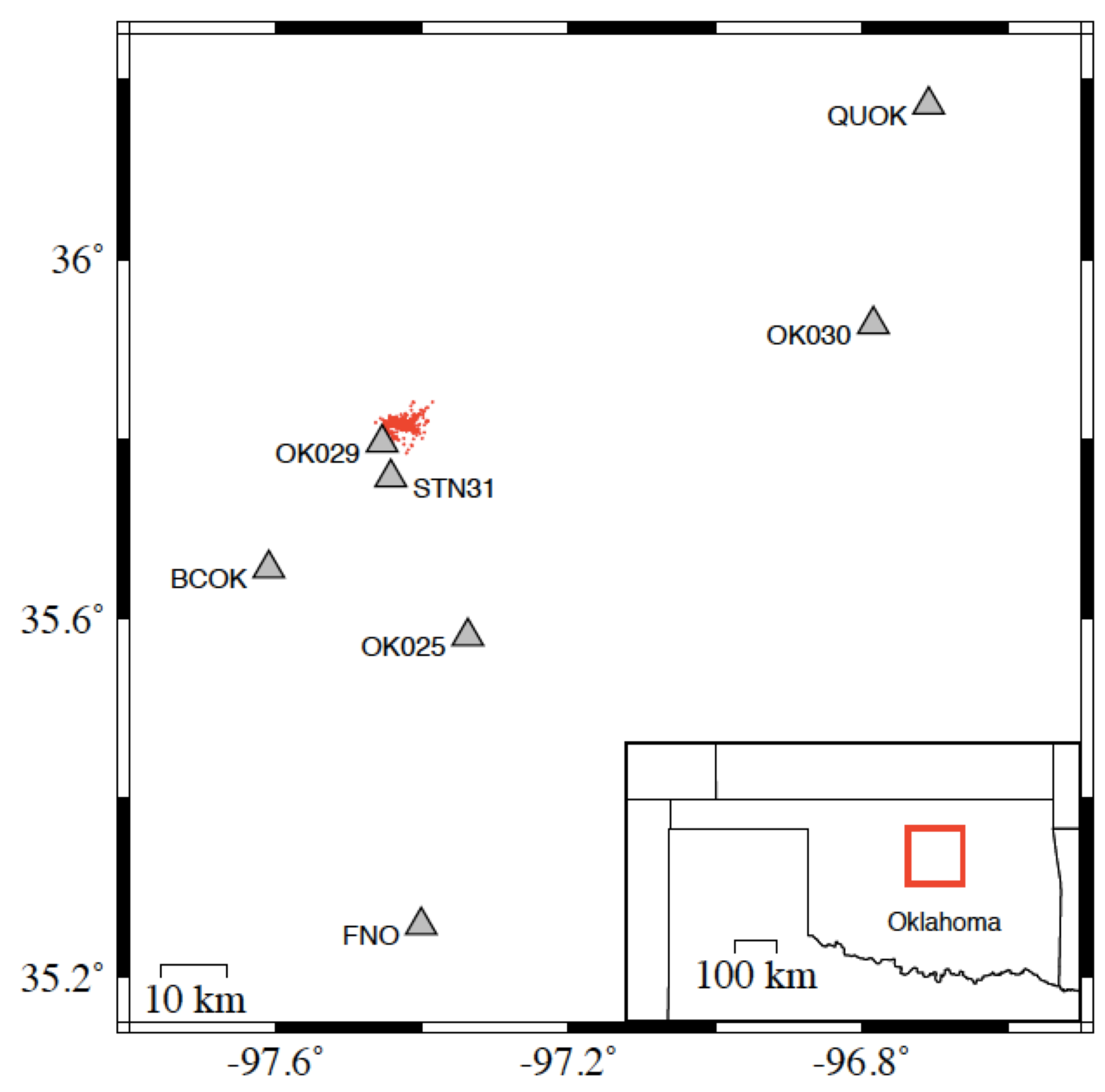

Figure 3. Map of study region in Oklahoma, central U.S. Red dots are 890 events with 11293 labelled P and S phase arrivals, gray triangles are broadband stations of US Geological Survey Networks (GS).

Scenario	Direct Transfer	Retrain FC Layer	Retrain FC+Layer 11
Accuracy(%)	79	84	92

Table 1. Classification accuracy of labeled P and S phases, when using CNN trained from Wenchuan dataset directly, or retraining the last fully-connected (FC) layer and last convolutional layer with 2500 phases.

Test 2: From Wenchuan to New Zealand

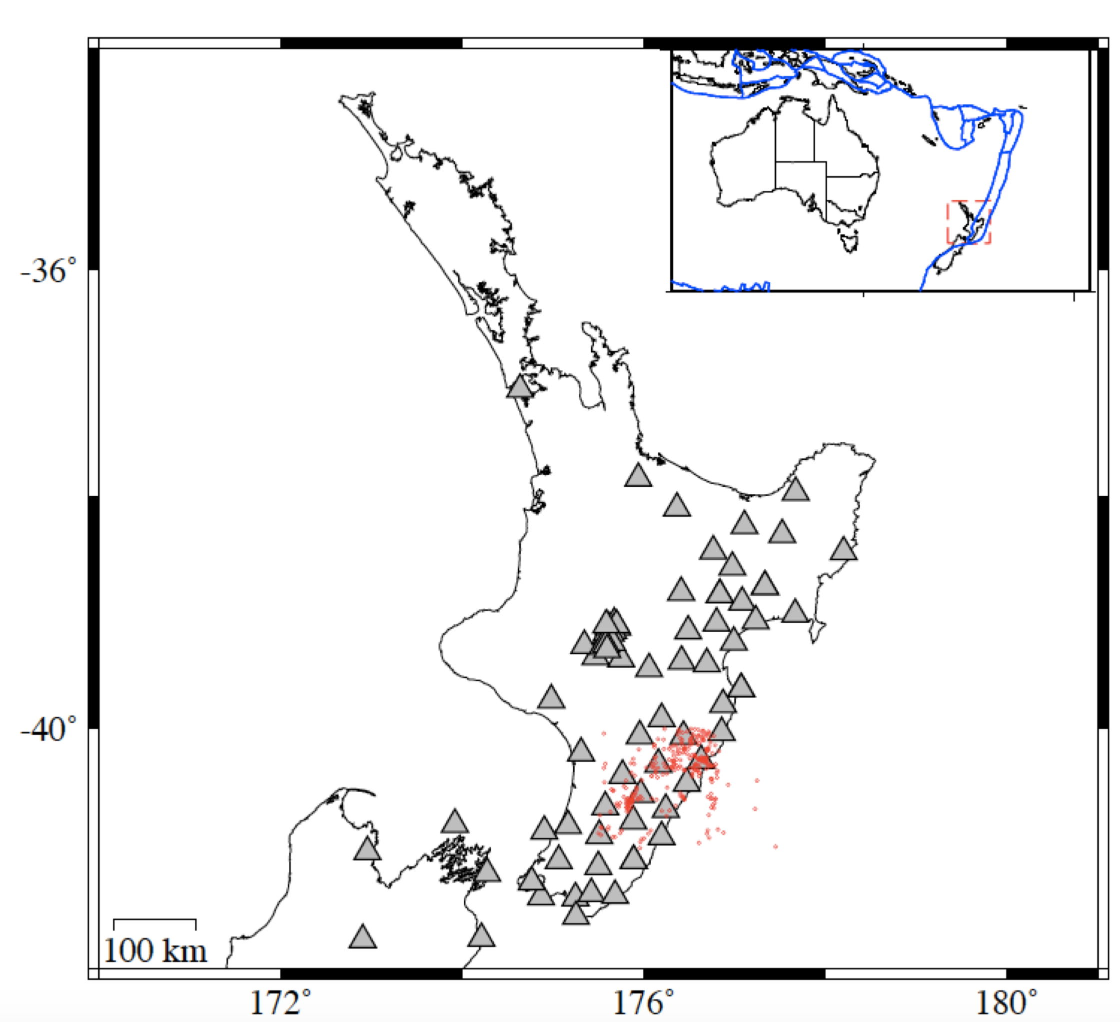

Figure 4. Map of study region in North Island of New Zealand. Red dots are 264 events with 5912 labelled P and S phase arrivals, gray triangles are 79 stations in GeoNet.

Scenario	Direct Transfer	Retrain FC Layer	Retrain FC+Layer 11
Accuracy(%)	73	85	94

Table 2. Classification accuracy of labeled P and S phases, when using CNN trained from Wenchuan dataset directly, or retraining the FC and last convolutional layer with 1000 phases.

Example

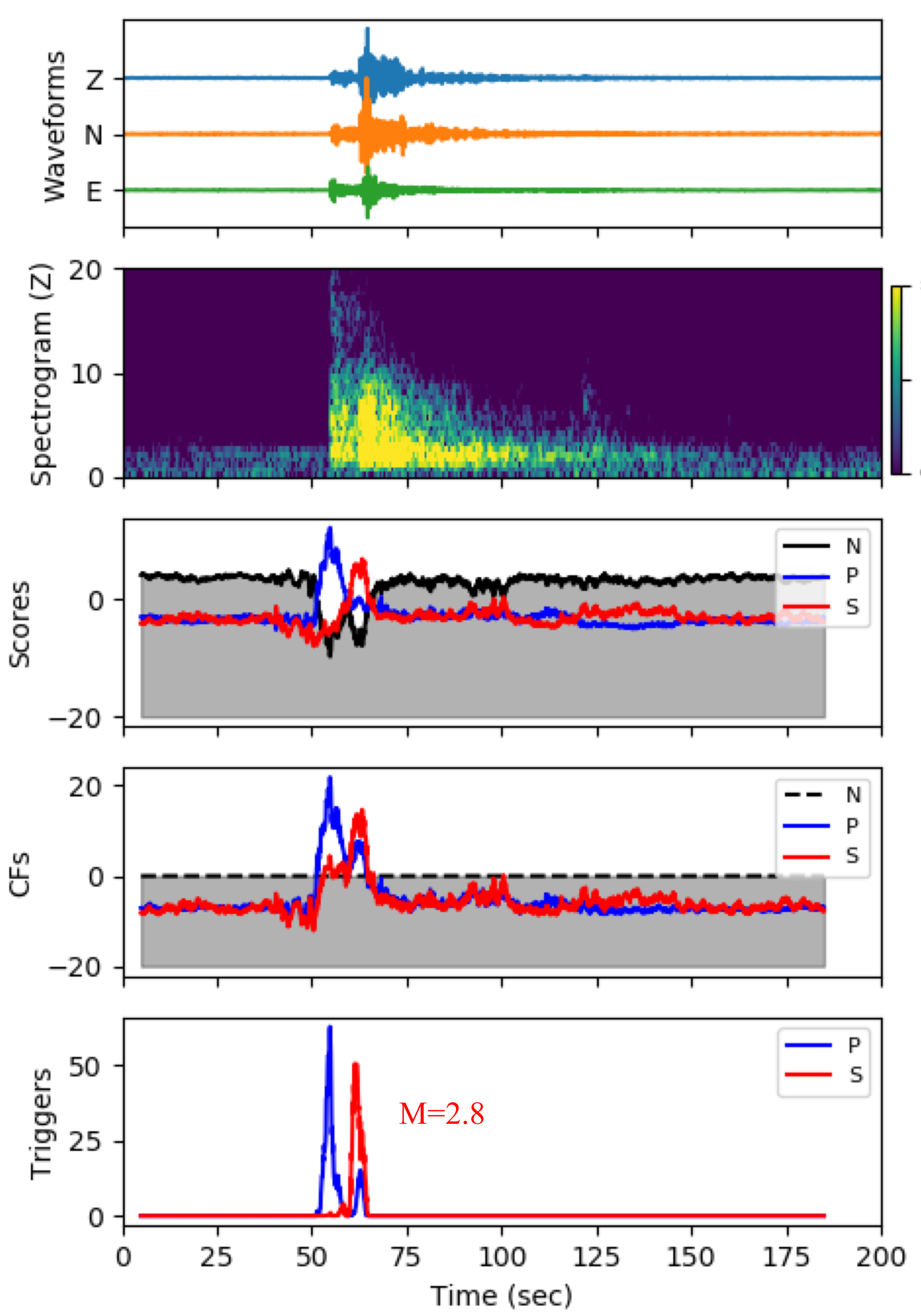

Figure 5. (a) Three-component raw waveform on station ANWS in New Zealand. (b) Spectrogram of vertical component. (c) Scores, (d) characteristic function (CF) of P-phase, S-phase and Noises. (e) Triggers. A magnitude 2.8 event on the catalog have been picked by the CNN.

Test 3: From Wenchuan to Salton Sea

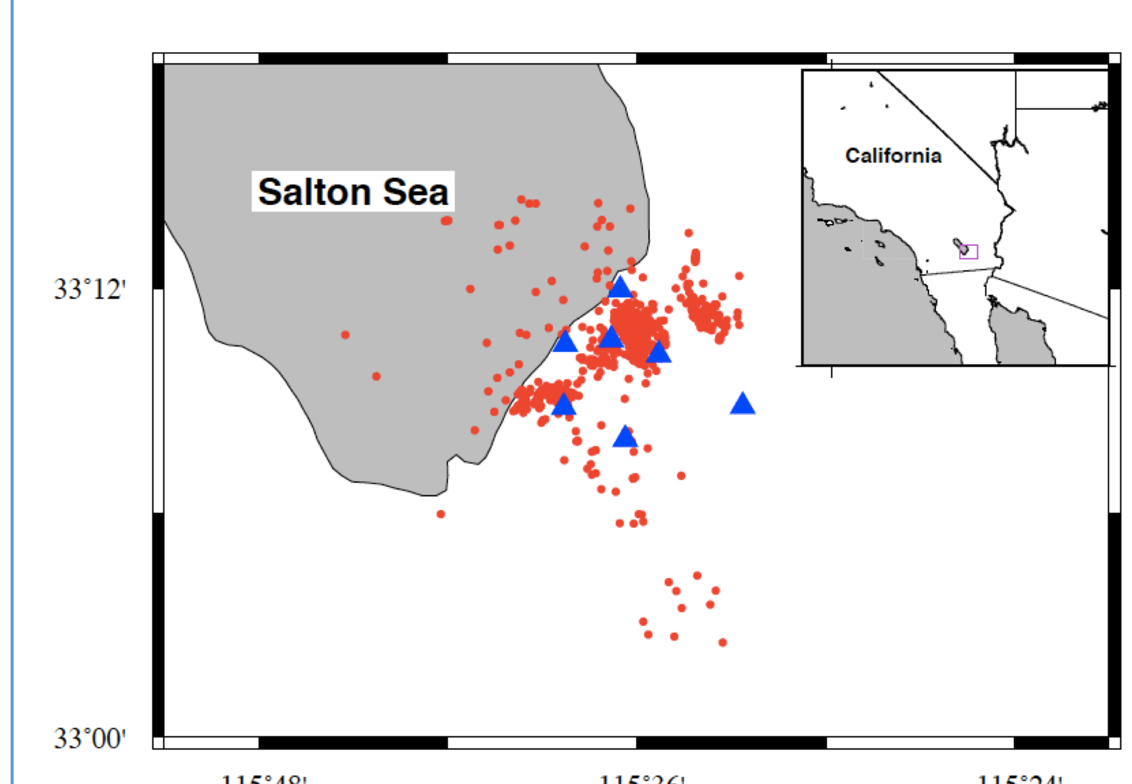

Figure 6. Map of study region in Salton Sea Geothermal Field, Southern California. Red dots are 879 events with 8939 labelled P and S phase arrivals, blue triangles are 7 stations in the EN borehole network.

Scenario	Direct Transfer	Retrain FC Layer	Retrain FC+Layer 11
Accuracy(%)	41	83	97

Table 3. Classification accuracy of labeled P and S phases, when using CNN trained from Wenchuan dataset directly, or retraining the FC and last convolutional layer with 2000 phases.

Example

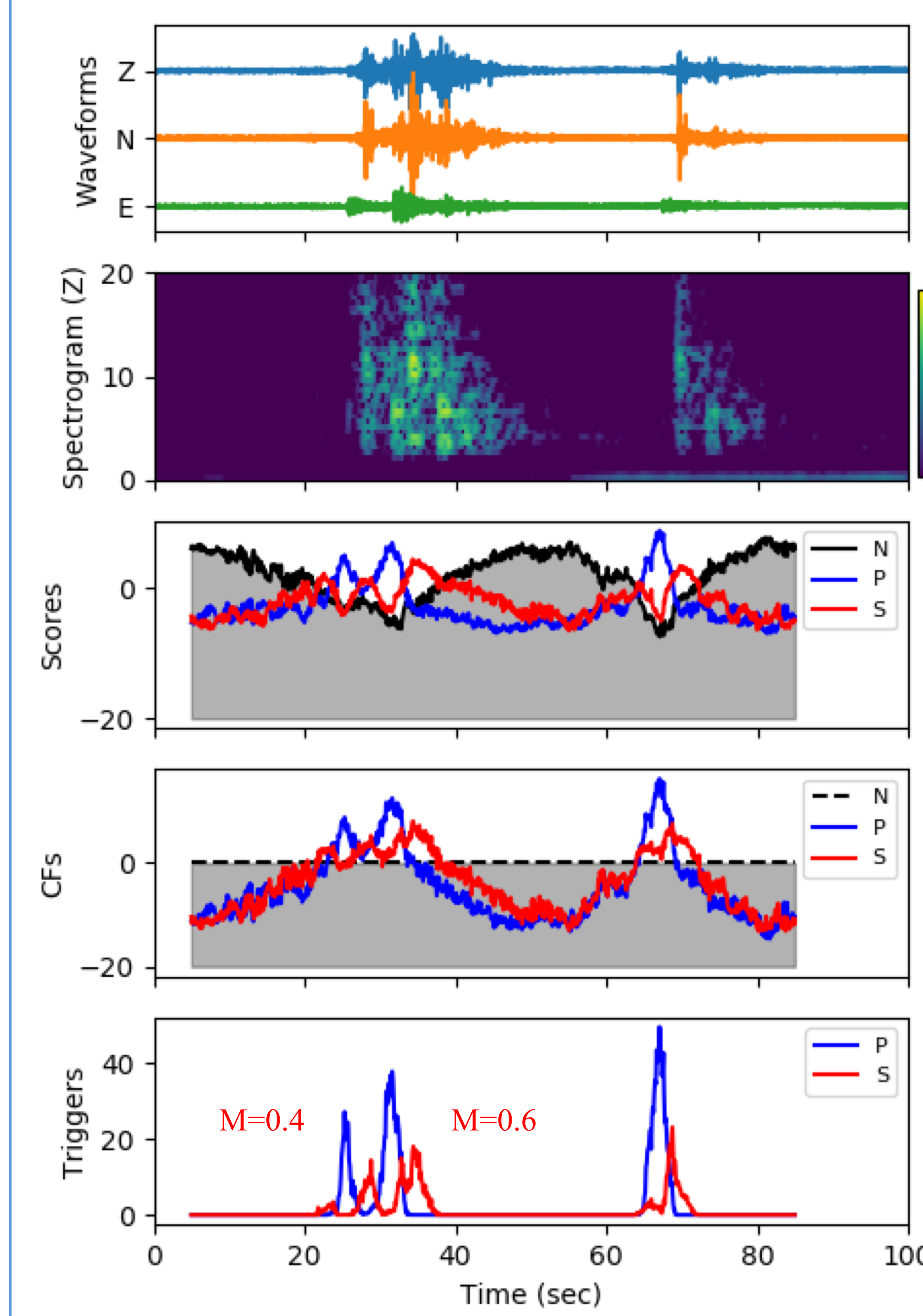

Figure 7. (a) Three-component raw waveform on station ENG in Salton Sea. (b) Spectrogram of vertical component. (c) Scores, (d) characteristic function (CF) of P-phase, S-phase and Noises. (e) Triggers. Two events on the catalog have been picked by the CNN.

Summary

- The low accuracy for direct transfer to Salton Sea is likely due to different types of station deployment (boreholeshort-period vs. surface broad-band).
- Retraining the last layer of CNN could improve the accuracy of transfer learning for all three regions.

Contact

Corresponding author: Chenyu Li; Email: lchenyu@gatech.edu