Machine Learning Based Convolutional Neural Network in Earthquake Detection and Classification and its Application in Southern California

Report for SCEC Award #18165

Submitted May 3rd, 2019

Investigators: Zhigang Peng (Georgia Institute of Technology)

I. Project Overview

A. Abstract

In the box below, describe the project objectives, methodology, and results obtained and their significance. If this work is a continuation of a multi-year SCEC-funded project, please include major research findings for all previous years in the abstract. (Maximum 250 words.)

We developed a convolutional neural network (CNN) based phase detection and picking approach named CPIC (Zhu et al., 2019). The method was first applied on a one-month aftershock sequence of 2008 Mw7.9 Wenchuan earthquake. 20-sec-long waveform frames are cut from 5 seconds before and 15 seconds after the picked P or S arrival times. Trained only on ~40,000 such frames, our CNN-based phase identification classifier (CPIC) achieve 97.6% classification accuracy on unseen 20,000 picked frames. More importantly, the CPIC approach is generally applicable to many seismic active regions, such as southern California, Oklahoma, and New Zealand. Benchmarked on the SCSN dataset (4.8M picked arrival times) released by the Caltech researchers, the CPIC model achieves similar high accuracy (99.5%) with a significantly simpler model and faster execution time. Finally, when tested on a small dataset from a different region (Oklahoma, US), CPIC achieves 97% accuracy after fine-tuning only the fully connected layer of the model. This result suggests that the CPIC developed in this study can be used to identify and pick P/S arrivals in other regions with no or minimum labeled phases.

B. SCEC Annual Science Highlights

Each year, the Science Planning Committee reviews and summarizes SCEC research accomplishments, and presents the results to the SCEC community and funding agencies. Rank (in order of preference) the sections in which you would like your project results to appear. Choose up to 3 working groups from below and re-order them according to your preference ranking.

Seismology
Computational Science
Mining Seismic Wavefields

C. Exemplary Figure

Select one figure from your project report that best exemplifies the significance of the results. The figure may be used in the SCEC Annual Science Highlights and chosen for the cover of the Annual Meeting Proceedings Volume. In the box below, enter the figure number from the project report, figure caption and figure credits.

Figure 2. Detection example on 15-minute recording on 14 stations with three catalog events for the Wenchuan dataset. Only vertical components are plotted. Blue and green curves show the probabilities of P and S phases. Red and magenta bars indicate the catalog P and S arrivals. Origin times of three catalog events are marked by the dashed vertical lines along with their magnitudes.

D. SCEC Science Priorities

In the box below, please list (in rank order) the SCEC priorities this project has achieved. See *https://www.scec.org/research/priorities* for list of SCEC research priorities. *For example: 6a, 6b, 6c*

3a, 3b, 3c

E. Intellectual Merit

How does the project contribute to the overall intellectual merit of SCEC? *For example: How does the research contribute to advancing knowledge and understanding in the field and, more specifically, SCEC research objectives? To what extent has the activity developed creative and original concepts?*

This project develops a machine learning based tool for seismic phase picking and detection that requires a small to moderate amount of training data. It is also applicable to other regions. An improved phase picking and event detection could result in many small-magnitude earthquakes being detected/located, which can help to improve our understanding of subsurface fault structures, large earthquake nucleation and earthquake interaction at nearby and long-range distances.

F. Broader Impacts

How does the project contribute to the broader impacts of SCEC as a whole? *For example: How well has the activity promoted or supported teaching, training, and learning at your institution or across SCEC? If your project included a SCEC intern, what was his/her contribution? How has your project broadened the participation of underrepresented groups? To what extent has the project enhanced the infrastructure for research and education (e.g., facilities, instrumentation, networks, and partnerships)? What are some possible benefits of the activity to society?*

This project supported collaborations of two GT students. Lijun Zhu from School of Electronic and Computer Engineering (ECE), and he is expected to defend his Ph.D. thesis in July 2019. This work is a major component of his Ph.D. thesis. Chenyu Li is a 5th year graduate student from School of Earth and Atmospheric Sciences (EAS). She is expected to graduate in summer 2020. We are in the process of releasing the related package and test dataset online at https://github.com/lijunzh/yews

G. Project Publications

All publications and presentations of the work funded must be entered in the SCEC Publications database. Log in at *http://www.scec.org/user/login* and select the Publications button to enter the SCEC Pubications System. Please either (a) update a publication record you previously submitted or (b) add new publication record(s) as needed. If you have any problems, please email *web@scec.org* for assistance.

II. Technical Report

The technical report should describe the project objectives, methodology, and results obtained and their significance. If this work is a continuation of a multi-year SCEC-funded project, please include major research findings for all previous years in the report. (Maximum 5 pages, 1-3 figures with captions, references and publications do not count against limit.)

A. Automatic Phase Arrival Picking based on CNN classifier

The increasing volume of seismic data from long-term continuous monitoring motivates the development of algorithms based on convolutional neural network (CNN) for faster and more reliable phase detection and picking (e.g., Ross et al., 2018; Zhu and Beroza, 2019). We have developed a CNN-based Phase-Identification Classifier (CPIC) designed for phase detection and picking on small to medium-sized training datasets (Zhu et al., 2019). When trained on 30,146 labeled phases and applied to one-month of continuous recordings during the aftershock sequences of the 2008 MW 7.9 Wenchuan Earthquake in Sichuan, China, CPIC detects 97.5% of the manually picked phases (Table 1) in the standard catalog and predicts their arrival times with a five-times improvement over the ObsPy (Beyreuther et al., 2010) AR-picker (Figure 1). In addition to the catalog phases manually picked by analysts, CPIC finds more phases for existing events and new events missed in the catalog (Figure 2). Among those additional detections, some are confirmed by a matched filter method while others require further investigation.

The core of the CPIC framework is the CNN classifier shown in Figure 3. An off-line training process optimizes the parameters of the CNN-based classifier iteratively over the labeled dataset. The trained classifier is then used during on-line processing for both phase detection and picking. The CNN classifier contains 11 convolutional layers along with one fully-connected layer, which processes many labeled windows known to contain P or S phases, or noise only. A Softmax function is used to normalize the probabilities in the output layer:

$$q_{i(x)} \ = \frac{e^{\{z_i\}}}{(e^\wedge\{z_0(x)\} + e^\wedge\{z_1(x)\} + e^\wedge\{z_2(x)\})}, \qquad (1)$$

where *i*=0, 1, 2 represents noise, P, and S classes, respectively. $z_i(x)$ is the unnormalized output of the last fully-connected layer for the ith class. Rectified linear unit (ReLU) is used as the activation function to introduce non-linearity to the CNN model. Data dimension is reduced at each layer via the max-pooling scheme. The CNN model is trained using the cross-entropy loss between true probability distribution *p* and the estimated distribution *q* as following:

$$H(p,q) = -\sum_x p(x) \log q(x). \qquad (2)$$

Adam algorithm is applied in the training process to optimize the model parameters, which converges nicely after 100 epochs (Figure 4b).

B. Comparison with recent works

One advantage of the CPIC method is the small requirement of the training dataset size. When compared with state-of-art methods, such as Generalized Phase Detector GPD (Ross et al. 2018) and PhaseNet (Zhu and Beroza, 2019), the CPIC method achieves comparable accuracies while only relies on a much smaller training set. Both of the aforementioned methods used training sets with more than 1 million samples. However, the CPIC model converges on training set as small as several thousand samples. Thus, the CPIC model can be applied to a broader range of datasets with fewer picked arrivals. The simpler model also results in fewer model parameters, which demands less computation power.

The CPIC model is further validated on the SCSN dataset released by SCEDC/Caltech (http://scedc.caltech.edu/research-tools/deeplearning.html). The SCSN dataset consists of 4.7 million three-component four-second-long waveforms of P phase, S phase, and noise. Trained on 80% randomly selected waveforms, the CPIC model reaches 98.6% accuracy on the rest 20% unseen waveforms. Although the validation accuracy is slightly lower than the benchmarked 99% accuracy from Ross et al. 2019, we avoid having the overfitting problem observed in that model which the loss function diverge after two epochs (Figure 4a). The CPIC model maintains a steady accuracy and loss event after 200 epochs.

Using the same benchmarked picking method (ObsPy ARPicker), we achieve a more significant improvement in picking accuracy over the benchmark method even though we have a lower absolute accuracy due to the more challenging dataset we used. We believe that the CPIC approach is more suitable for regions with fewer catalog arrival times and more challenging noise conditions.

C. Applications of CPIC on additional regions

The CPIC model is not only useful to the region it was trained on, but also shines some insight on the general phase arrival picking problem. We apply the trained CPIC model on the aftershock waveforms from Wenchuan to the induced earthquake waveforms in Oklahoma, U.S. Summarized in Table 2, the original Wenchuan CPIC model already reaches a high 87.5 % overall accuracy. The stations having similar source-receiver distances (OK025 and OK029) results in over 90% accuracy while the station further away (OK030) only reaches 69.9%. After fine-tuning the CPIC model using only 2,000 labeled waveforms, the overall accuracy of the CPIC model reach 97%. This transfer-learning example demonstrates that the CPIC model has high potential for generalizing to broader applications on different monitoring regions.

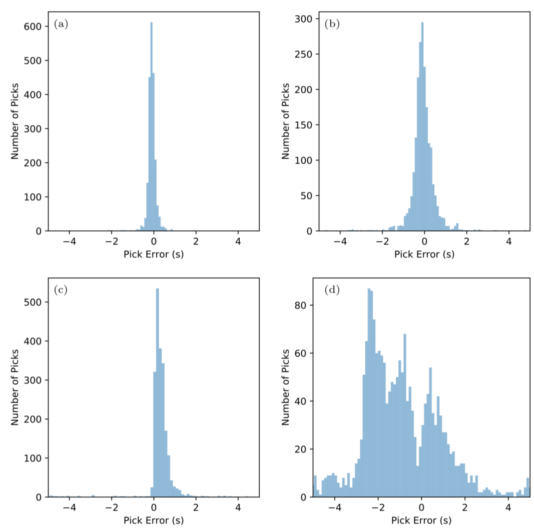

Figure 1. The distributions of picking errors of CPIC (upper panels) and ObsPy AR picker (lower panels) on the validation dataset. After Zhu et al. (2019).

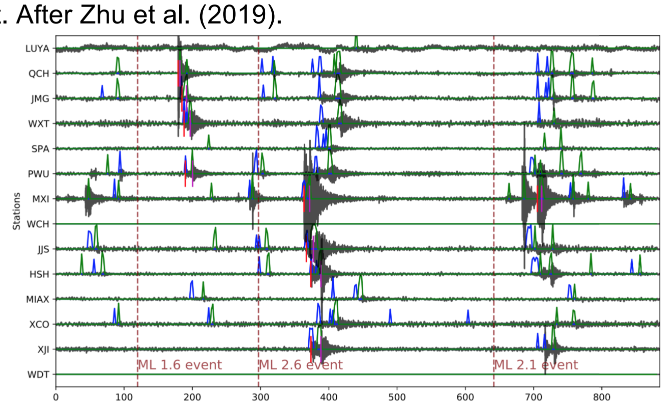

Figure 2. Detection example on 15-minute recording on 14 stations with three catalog events for the Wenchuan dataset. Only vertical components are plotted. Blue and green curves show the probabilities of P and S phases. Red and magenta bars indicate the catalog P and S arrivals. Origin times of three catalog events are marked by the dashed vertical lines along with their magnitudes. After Zhu et al. (2019).

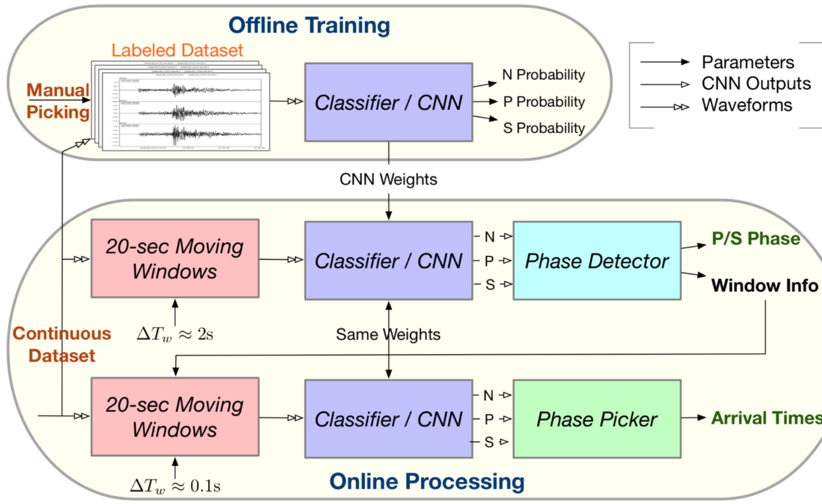

Figure 3. CNN-based Phase-Identification Classifier (CPIC) flow chart. Inputs are three-component seismograms recorded at a single station, labeled in red. Outputs are P-wave, S-wave or noise window probabilities, and picked arrival times for P and S phases, shown in green. After Zhu et al. (2019).

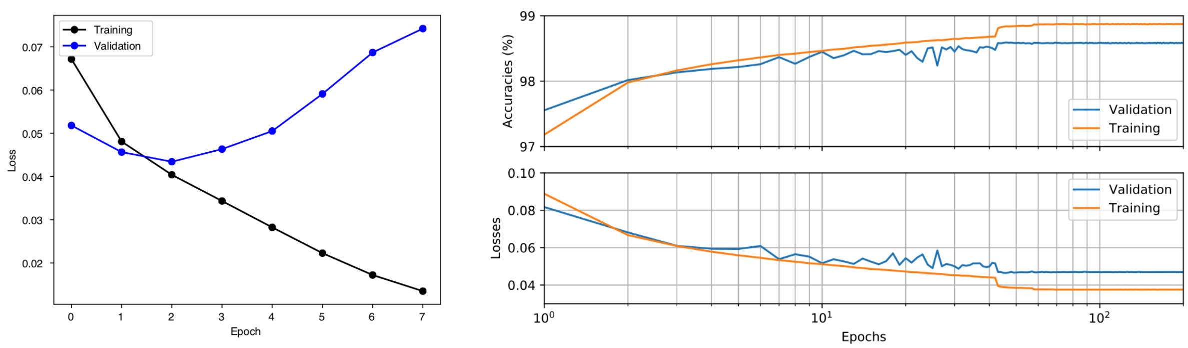

Figure 4(Left). Cross-entropy loss during training and validation of the GPD model on SCSN dataset (Ross et al., 2018). (b) Accuracies and losses during training and validation of the CPIC model on SCSN dataset.

		Detector			
		Noise	P-wave	S-wave	Total
Catalog	Noise	5,946	97	113	6,156
	P-wave	22	2,930	10	2,962
	S-wave	59	6	2,873	2,938
	Total	6,027	3,033	2,996	12,056

Table 1. Confusion matrix for phase classification on the validation dataset which is the latest 20% of the labeled phases.

Station	OK025	OK029	OK030	All
Original (%)	95.7	92.2	69.9	87.5
Fine-tuned (%)	98.8	96.2	94.2	97.0

Table 2. CPIC accuracy when testing on a three-station seismic dataset in OK, USA. The first row shows the performance of directly applying CPIC as trained on the Wenchuan, China dataset, while the second row shows the enhanced accuracy after fine-tuning CPIC on 2,000 training samples from the Oklahoma region.

D. Student Support and Involvement

This project provided 3-month support (1.5-month each) for graduate students Chenyu Li and Lijun Zhu at Georgia Tech. Li is a 5th year Ph.D. student in School of Earth and Atmospheric Sciences (EAS). She has been using the matched filter detection to study spatial-temporal changes of seismicity in SSGF following nearby and large distant earthquakes as funded by SCEC (Li et al., 2017), as well as detecting aftershocks following a magnitude 7.5 intermediate-depth earthquake in Hindu Kush region (Li et al., 2018a). Recently she was also involved in applying the CPIC to other regions (Li et al., 2018b). Zhu is a 6th year Ph.D. student in the School of Electrical and Computer Engineering (ECE) at Georgia Tech. He has been working on phase picking using CNN since 2017 (Zhu et al., 2017, 2018). The proposed work will be part of their Ph.D. thesis.

E. Acknowledgement

We utilized the PyTorch deep-learning neural network package (Paszke et al., 2017) and ObsPy package (Beyreuther et al., 2010). The seismic data utilized in this study is obtained during the 2017 "Aftershock Detection Artificial-Intelligence Contest" (Fang et al., 2017). This research is jointly supported by National Science Foundation (NSF award EAR-1818611) and Southern California Earthquake Center (SCEC award 18165). This is SCEC contribution #9046.

F. References (with publications/meeting abstracts supported by the grant marked in bold)

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: a python toolbox for seismology. Seismological Research Letters, 81, 530–533.

Fang, L., Wu, Z., & Song, K. (2017). SeismOlympics. Seismological Research Letters, 88 , 1429.Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., & Lerer, A. (2017). Automatic differentiation in PyTorch. In Advances in Neural Information Processing Systems 30,

Li, C., Peng, Z., Zhang, C. & Meng, X. (2017). Temporal Changes in Seismicity Velocities in Salton and Seismic Sea Geothermal Field, SSA Annual Meeting 2017.

Li, C., Peng, Z., Yao, D., Guo, H., Zhan, Z. & Zhang, H. (2018a). Aboundant aftershock Sequences of the 2015 Mw7.5 Hindu Kush Intermediate-Depth Earthquake, *Geophys. J. Int.*, 213, 1121-1134, doi:10.1093/ggy016.

Li, C., Zhu, L., Yao, D., Meng, X., Peng, Z., McClellan, J.H., & Walter, J.I. (2018b), Transfer learning for seismic phase picking on different geographic regions, abstract S11E-0431 submitted to the annual American Geophysical Union Fall meeting, Washington DC, 10-14 Dec.

Ross, Z. E., Meier, M., Hauksson, E., & Heaton, T. H. (2018a). Generalized seismic phase detection with deep learning (short note). Bulletin of the Seis- mological Society of America, 108, 2894–2901.

Zhu, L., Li Z., Li, C., Wang, B., Chen, C., McClellan, J.H., & Peng, Z. (2017), Machine-Learning Inspired Seismic Phase Detection for Aftershocks of the 2008 MW7.9 Wenchuan Earthquake. In AGU Fall Meeting Abstracts

Zhu, L., Peng, Z., & McClellan, J.H. (2018). Deep learning for seismic event detection of earthquake aftershocks. In 2018 52nd Asilomar Conference on Signals, Systems, and Computers (pp. 1121–1125).

Zhu, L., Peng., Z., McCllean, J.H., Li, C., Yao, D., Li, Z., & Fang, L. (2019). Deep learning for seismic phase detection and picking in the aftershock zone of 2008 M_w7.9 Wenchuan Earthquake, *Phys. Earth Planet. Int.*, accepted, https://arxiv.org/abs/1901.06396.

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. Geophysical Journal International, 216, 261–273.