
Doctoral Dissertations

Student Theses and Dissertations

Fall 2022

**APPLICATION OF MACHINE LEARNING IN GEOPHYSICS:
RANKING TELESEISMIC SHEAR WAVE SPLITTING
MEASUREMENTS AND CLASSIFYING DIFFERENT TYPES OF
EARTHQUAKES**

Yanwei Zhang
Missouri University of Science and Technology

Follow this and additional works at: https://scholarsmine.mst.edu/doctoral_dissertations



Part of the [Geology Commons](#), and the [Geophysics and Seismology Commons](#)

Department: Geosciences and Geological and Petroleum Engineering

Recommended Citation

Zhang, Yanwei, "APPLICATION OF MACHINE LEARNING IN GEOPHYSICS: RANKING TELESEISMIC SHEAR WAVE SPLITTING MEASUREMENTS AND CLASSIFYING DIFFERENT TYPES OF EARTHQUAKES" (2022). *Doctoral Dissertations*. 3245.
https://scholarsmine.mst.edu/doctoral_dissertations/3245

This thesis is brought to you by Scholars' Mine, a service of the Missouri S&T Library and Learning Resources. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

APPLICATION OF MACHINE LEARNING IN GEOPHYSICS: RANKING
TELESEISMIC SHEAR WAVE SPLITTING MEASUREMENTS AND
CLASSIFYING DIFFERENT TYPES OF EARTHQUAKES

by

YANWEI ZHANG

A DISSERTATION

Presented to the Graduate Faculty of the
MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

GEOLOGY AND GEOPHYSICS

2022

Approved by:

Stephen S. Gao, Advisor
Kelly H. Liu
Ryan Smith
Wenqing Hu
Guangzhi Zhang

© 2020

Yanwei Zhang

All Rights Reserved

PUBLICATION DISSERTATION OPTION

This dissertation consists of the following articles that has been formatted in the style used by the Missouri University of Science and Technology

Paper I: Page 3-32, has been published by Geophysical Research Letters.

Paper II: Page 33-45, is intended for submission to Geophysical Journal International.

ABSTRACT

During the past decades, applications of Machine Learning have been explosively developed to solve various academic and industrial problems, and over-human performance has been shown in diverse areas. In geophysical research, Machine Learning, especially Convolutional Neural Network (CNN), has been applied in numerous studies and demonstrated considerable potential. In this study, we applied CNN to solve two geophysical problems, ranking teleseismic shear splitting (SWS) measurements and classifying different types of earthquakes.

For ranking teleseismic SWS measurements, we utilized a CNN-based method to automatically select reliable SWS measurements. The CNN was trained by human-verified teleseismic SWS measurements and tested using synthetic SWS measurements. Application of the trained CNN to broadband seismic data recorded in south-central Alaska reveals that CNN classifies 98.1% of human-selected measurements as acceptable and revealed ~30% additional measurements.

For classifying different types of earthquakes, we utilized a CNN to classify natural earthquakes, mine collapses, and explosions using seismic waveforms recorded by 287 stations in Shandong Province, China. Cross-validation is employed to scan the whole dataset, and the measurements with different labels between human and the CNN are manually assessed and kept, corrected, or abandoned in the dataset. Testing with the corrected dataset, the classification accuracies of the three types of events increase from 97.3% to 99.2% for earthquakes, from 84.9% to 95.8% for mine collapses, and from 93.6% to 98.1% for explosions.

ACKNOWLEDGMENTS

Firstly, I would like to deliver my best appreciation and regards to my advisors, Dr. Stephen S. Gao and Dr. Kelly H. Liu for their outstanding and patient education and guidance during my Ph.D. study at MST. With their advising, helping, and encouraging, I overcame various issues encountered in these 5 years. I learnt how to not only be a qualified researcher, but also face the difficulties in life and future career.

And then, I am grateful to my dissertation committee members: Dr. Ryan Smith, Dr. Wenqing Hu, and Dr. Guangzhi Zhang. They provided me with valuable suggestions and helpful support for my studies. I learnt critical knowledge from them and expanded the horizon of my academic research.

I am also thankful to all colleagues and former members in the Geophysics Group who helped me during the past 5 years. Especially, Dr. Junhao Qu and authors of Liu et al. (2014), Yang et al. (2016) and Yang et al. (2017). Without their greatly quantitative and highly qualitative accumulation of datasets, I could not start my studies in Machine Learning which is a highly data-drive technique.

Moreover, I would like to thank the supporting team of Foundry, the campus high performance computer cluster. Foundry has dramatic capacity of computation, and I could realize my different ideas in a short time with strongly supporting from Foundry.

The last but not the least, the special thanks go to my parents. I could not finish my degree without their support.

TABLE OF CONTENTS

	Page
PUBLICATION DISSERTATION OPTION	iii
ABSTRACT.....	iv
ACKNOWLEDGMENTS	v
LIST OF ILLUSTRATIONS.....	viii
NOMENCLATURE	ix
 SECTION	
1. INTRODUCTION	1
 PAPER	
I. CLASSIFICATION OF TELESEISMIC SHEAR WAVE SPLITTING MEASUREMENTS: A CONVOLUTIONAL NEURAL NETWORK APPROACH.....	3
ABSTRACT.....	3
PLAIN LANGUAGE SUMMARY	4
1. INTRODUCTION	4
2. TRAINING DATA SET AND PREPROCESSING	7
3. STRUCTURE AND TRAINING OF THE CNN.....	9
4. TESTING WITH SYNTHETIC SWS MEASUREMENTS	11
5. APPLICATION TO SWS MEASUREMENTS IN SOUTH CENTRAL ALASKA	13
6. DISCUSSION	16
6.1. COMPARISON WITH HUMAN-DETERMINED MEASUREMENTS	16

6.2. COMPARISON WITH A FULLY AUTOMATED NON-MACHINE LEARNING SWS MEASUREMENT APPROACH	18
6.3. LIMITATIONS OF THE CURRENT CNN AND SUGGESTED NEXT STEPS.....	19
7. CONCLUSION.....	20
ACKNOWLEDGEMENTS.....	20
DATA AVAILABILITY STATEMENT	21
SUPPORTING INFORMATION.....	22
REFERENCES	28
II. CONVOLUTIONAL NEURAL NETWORK CLASSIFICATION OF NATURAL EARTHQUAKES, MINE COLLAPSES AND EXPLOSIONS	33
ABSTRACT.....	33
1. INTRODUCTION	34
2. DATA AND METHOD.....	36
3. RESULTS	39
4. DISCUSSION	41
5. CONCLUSIONS.....	42
ACKNOWLEDGEMENTS.....	43
REFERENCES	43
SECTION	
2. CONCLUSION.....	46
BIBLIOGRAPHY.....	47
VITA.....	53

LIST OF ILLUSTRATIONS

PAPER I	Page
Figure 1. Distribution of seismic stations (red triangles; Liu et al., 2014; Yang et al., 2016, 2017) used for training the CNN.	8
Figure 2. Performance tests of the CNN using synthetic dataset.	12
Figure 3. SWS measurements (red bars) in south central Alaska plotted at the recording stations (blue triangles).	14
Figure 4. (a) Comparison of human-determined (blue bars; Yang et al., 2021) and CNN-selected (red bars) SWS measurements in south central Alaska. Green bars are measurements accepted by both CNN and human operators. All the measurements are plotted above the XKS ray-piercing points at 200 km deep which is the most likely depth of the anisotropic layer (Yang et al., 2021). The contour lines show the depth of the subducted Pacific slab, and the thick dashed lines separate four regions (A-D) with different patterns of splitting measurements (Yang et al., 2021). The CNN results are the same as those shown in Figure 3f. (b) Cross-plot of human and CNN determined station-averaged ϕ measurements. The black bars are the standard deviation. (c) Same as (b) but for station-averaged δt measurements. (d) and (e) are respectively the same as (b) and (c) but the CNN station averages were computed using results that were determined by CNN as acceptable but rejected by the human operators.	17
PAPER II	
Figure 1. Stations and events distribution of dataset in the study area.	36
Figure 2. The workflow of CNN in this study.	37
Figure 3. The performance of the CNN.	39
Figure 4. (a) the accuracy of CNN with corrected dataset various with number of measurements of each event. (b) the accuracy of CNN with corrected dataset various with magnitude.	41

NOMENCLATURE

Symbol	Description
ϕ	Fast Orientation
δt	Splitting Time
BAZ	Back Azimuth
CNN	Convolutional Neural Network
IRIS	Incorporated Research Institutions for Seismology
ML	Machine Learning
SNR	Signal-to-Noise Ratio
SSNC	Shandong Seismic Network Center
SWS	Shear Wave Splitting
XCC	Cross-Correlation Coefficient.

1. INTRODUCTION

In the past few decades, the outstanding performance of Machine Learning (ML) has attracted the attention of scientists, which leads to a significant increase of applications in various areas. In geophysical research, relied on the incredible ability of feature extraction, Convolutional Neural Network (CNN), one of the most common ML techniques, shows considerable potential in classification problems based on seismic waveforms (Perol et al., 2018; Zhu & Beroza, 2018; Linville et al., 2019; Zhang and Gao, 2022). In this study, we applied CNN to solve two geophysics problems: ranking teleseismic SWS measurements and classifying different types of earthquakes. Traditionally, both are manually assessed by human experts and consume large amounts of time in each project.

The dissertation is mainly composed of two parts. The first part introduces the application in ranking teleseismic SWS measurements. The measurements in training dataset are from Liu et al. (2014), Yang et al. (2016) and Yang et al. (2017) for the contiguous United States and adjacent areas. All of measurements are manually ranked based on criteria proposed by Liu and Gao (2013). After the training process, our CNN shows high performance with a synthetic dataset and high consistency with published results in south central Alaska (Yang et al., 2021). To our knowledge, this is the first time when a ML-based technique is applied to SWS analysis.

The second part of the dissertation utilizes CNN to classify natural earthquakes, mine collapses and quarry explosions. The human labeled dataset from Shandong Seismic Network Center (SSNC) is used to train and test our CNN. To minimize the influence of unreliable measurements in the dataset, a 10-fold cross validation is

employed to scan the whole dataset. The measurements with different labels between human and CNN are manually assessed and kept, corrected, or abandoned in dataset. Comparing with the original dataset, the classification accuracies of the three types of events have obviously increase and all of them are above 95%. Our results reveal that the unreliable measurements have negative effect in ML studies and indicate that the cross-validation with CNN can evaluate, correct, and enhance the dataset.

PAPER

I. CLASSIFICATION OF TELESEISMIC SHEAR WAVE SPLITTING MEASUREMENTS: A CONVOLUTIONAL NEURAL NETWORK APPROACH

Yanwei Zhang and Stephen S. Gao

Department of Geoscience and Geological and Petroleum Engineering, Missouri
University of Science and Technology, Rolla, MO 65409

ABSTRACT

SWS analysis is widely used to provide critical constraints on crustal and mantle structure and dynamic models. In order to obtain reliable splitting measurements, an essential step is to visually verify all the measurements to reject problematic measurements, a task that is increasingly time consuming due to the exponential increase in the amount of data. In this study, we utilized a CNN based method to automatically select reliable SWS measurements. The CNN was trained by human-verified teleseismic SWS measurements and tested using synthetic SWS measurements. Application of the trained CNN to broadband seismic data recorded in south central Alaska reveals that CNN classifies 98.1% of human selected measurements as acceptable and revealed ~30% additional measurements. To our knowledge, this is the first study to systematically explore the potential of a machine-learning based technique to assist with SWS analysis.

PLAIN LANGUAGE SUMMARY

One of the routinely utilized seismological techniques to delineate the internal structure and dynamics of the Earth's crust and mantle is SWS analysis. In order to obtain reliable measurements, it is necessary to perform manual verification of the numerous measurements, a task that is increasingly time consuming due to the rapid increase in the amount of seismic data recorded worldwide. In this study, by taking advantage of the recent revolutionary development of a machine-learning technique called the CNN, we systematically investigate the potential and performance of a CNN that is trained using human-labeled SWS measurements and tested using both synthetic and recorded data. The results demonstrate satisfactory performance of the CNN on both types of data. Although additional development of the CNN is needed to reach over-human performance, our tests suggest that if a human operator uses CNN-classified results for manual verification, an approximately 60% reduction of the effort will be achieved, and only about 2% of the measurements will be missed.

1. INTRODUCTION

It has long been recognized that the P-to-s converted phases from the core-mantle boundary such as SKS, PKS, and SKKS (hereafter referred to as XKS) split into orthogonally polarized fast and slow components in azimuthally anisotropic media (Ando et al., 1983; Long & Silver, 2009; Savage, 1999; Silver & Chan, 1991). The two splitting parameters, the polarization orientation of the ϕ and the δt between the two waves, reveal the orientation and splitting magnitude of the anisotropy, respectively. Over the past

several decades, the SWS analysis technique has been widely used to delineate azimuthal anisotropy in the upper mantle, where lattice preferred orientation of crystallographic axes of main constitute minerals such as olivine is the dominant cause of the observed anisotropy (Katayama & Karato, 2006; Savage, 1999; Silver, 1996; Zhang & Karato, 1995).

Several different methods have been utilized to measure the splitting parameters of the XKS phases, among which the transverse energy minimization method (Silver & Chan, 1991) is arguably the most reliable one for its stability for noisy data (Vecsey et al., 2008). In this method, a grid search procedure is applied to find the optimal pair of splitting parameters (ϕ , δt) corresponding to the minimum XKS energy on the corrected transverse component. Numerous SWS studies suggest that in order to obtain reliable splitting measurements, an essential step is to visually verify all the measurements (e.g., Liu & Gao, 2013), as demonstrated in recent studies in North America (e.g., Liu et al., 2014; Yang et al., 2016, 2017; Yang et al., 2021). Due to the ever-increasing number of stations established around the world and the resultant exponential increase in the amount of data available for SWS analysis, this laborious task is increasingly time-consuming, and therefore alternate time-efficient yet reliable approaches are needed.

In recent years, applications of ML based techniques on various scientific problems have dramatically increased, and over-human performance has been shown in diverse areas. Especially after AlphaGo showed unexpected high performance (Silver et al., 2016), ML became widely known and attracted the attention of researchers from different fields. In geophysical research, ML has been applied in numerous studies and demonstrated considerable potential. Such applications include earthquake early warning

(Li et al., 2018), earthquake detection and magnitude estimation (Lomax et al., 2019; Mousavi & Beroza, 2020; Perol et al., 2018), seismic phase picking (Dokht et al., 2019; McBrearty et al., 2019; Ross et al., 2018; Woollam et al., 2019; Zhu & Beroza, 2018), event classification (Linville et al., 2019; Titos et al., 2018), first-motion polarity determination (Ross et al., 2018), seismic denoising (Zhu et al., 2019), and earthquake prediction (Mignan & Broccardo, 2019; Rouet-Leduc et al., 2017). ML-based applications on structural seismological problems such as teleseismic tomography, receiver functions, and SWS analyses are relatively rare and are starting to takeoff (e.g., Bianco et al., 2019; Garcia et al., 2021).

In this study we design a CNN to classify the automatically determined SWS splitting parameters measured based on the set of procedures outlined in Liu and Gao (2013) into acceptable and unacceptable ones. Traditionally such kind of grouping was conducted by trained human operators and was the most time-consuming step in SWS analysis (e.g., Liu & Gao, 2013), although efforts have been made to make such processes fully automatic for both teleseismic (Link et al., 2022; Teanby et al., 2004) and local S-wave (Peng & Ben-Zion, 2004) splitting measurements with promising results. While many SWS studies group the measurements in to “good,” “fair,” “unacceptable,” and “null,” the vast majority of them use the first two categories (which are collectively called “acceptable” in this study) indistinguishably for interpretation. For the “nulls,” which are characterized by a lack of XKS energy on the original transverse component while strong XKS energy appears on the radial component (e.g., Silver & Chan, 1991), we identify them in the pre-processing stage. The CNN is trained using published SWS measurements verified by human operators and is tested using splitting measurements

obtained using synthetic data. The trained CNN is applied to data from 127 stations in south central Alaska and the results are compared with those reported in a recent study (Yang et al., 2021). To our knowledge, this is the first time when a ML-based technique is applied to SWS analysis.

2. TRAINING DATA SET AND PREPROCESSING

Our CNN is trained and verified using 86,903 published human-labeled XKS SWS measurements recorded by 1,108 stations from Liu et al. (2014), Yang et al. (2016) and Yang et al. (2017) for the contiguous United States and adjacent areas (Figure 1). This data set was measured, manually verified, and ranked based on the procedures in Liu and Gao (2013), and contains 8,117 well-defined (ranks A and B) and 78,786 not well-defined (rank C and rank N, which are for null measurements). The procedures start with an automatic SWS measuring step built based on the minimization of transverse energy method of Silver & Chan (1991). The automatically determined measurements are then auto-ranked based on the XKS signal to noise ratios on the original radial, original transverse, and the corrected transverse components (Liu et al., 2008). All the resulting measurements and their ranks are subsequently manually screened to validate the results based on the following criteria: goodness of the fitting between the fast and slow components, the linearity of the particle motion pattern, the robustness and uniqueness of the minimum value on the contour of remaining energy on the corrected transverse components, and the strength of the XKS arrival on the original radial and transverse components. If necessary, the XKS time window, the band-pass frequency, and the

automatically determined ranking are manually adjusted to exclude non-XKS arrivals in the XKS window and to improve the SNR.

The stations cover different tectonic areas of the North America (Figure 1) so that our training and verification data contain measurements from various geological backgrounds. Additionally, the data set contains stations with both azimuthally invariant and azimuthally varying individual measurements, indicating the presence of simple and complex anisotropy, respectively. The splitting measurement procedure takes the original radial and transverse components as the input data, grid-searches for the optimal splitting parameters, and generates corrected radial and transverse components. The optimal pair of splitting parameters produces a corrected transverse component that has the minimum XKS energy among all the candidate pairs of the splitting parameters. Our CNN takes all

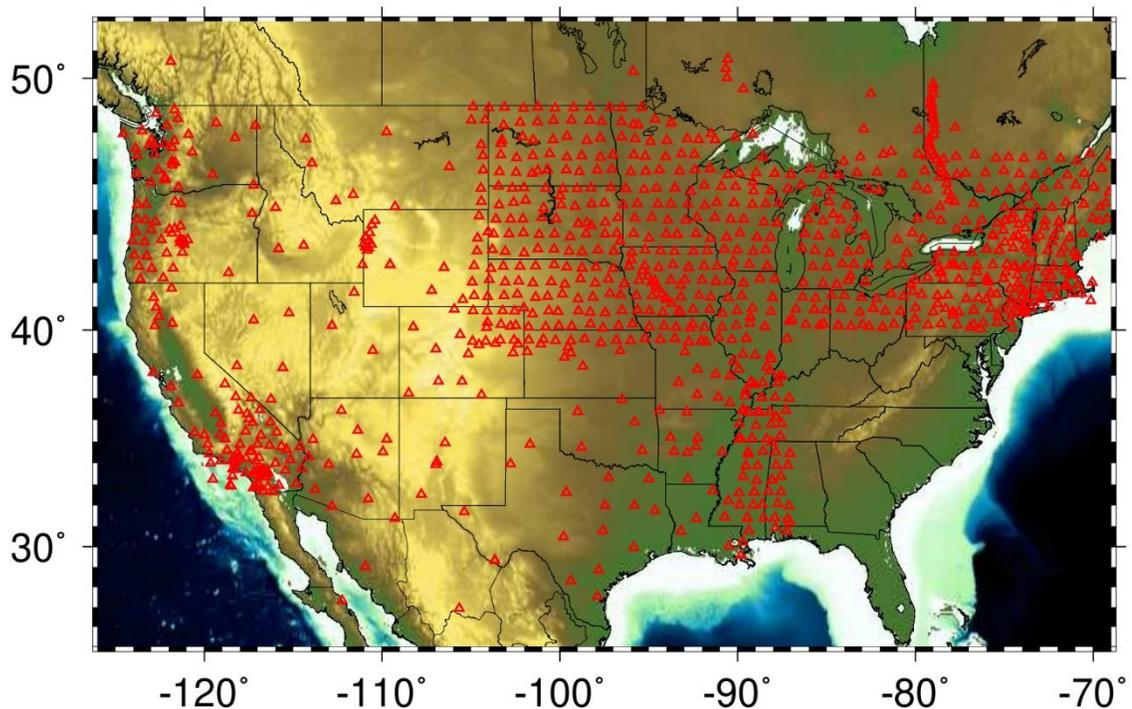


Figure 1. Distribution of seismic stations (red triangles; Liu et al., 2014; Yang et al., 2016, 2017) used for training the CNN.

the four seismograms as input data and groups the optimal pair of parameters as either acceptable (ranks A and B) or unacceptable (ranks C and N). After testing with various combinations of the length and the onset of the time window for the four seismograms, a 50 s window centered at the theoretical arrival time of the XKS phase predicted based on the IASP91 Earth model is used in the study.

Acceptable measurements are labeled as array [1 0] and unacceptable measurements are labeled as [0 1]. Because the number of acceptable measurements is significantly greater than that of the unacceptable data in the training data set, to avoid overfitting, we balance the data set by setting different class weights to the two sets of data (nine for the acceptable ones and one for the unacceptable ones), which is a common practice in similar situations (Japkowicz & Stephen, 2002). After random shuffling, 80% of the measurements are used for training and 20% for validation.

3. STRUCTURE AND TRAINING OF THE CNN

The CNN is built on Keras which is a high-level neural networks application programming interface (Gulli & Pal, 2017). Following Perol et al. (2018), we designed a CNN with eight 1-D convolutional layers followed by a full-connect layer. For each splitting measurement, the input for CNN includes four seismic waveforms each with 1,000 nodes (50 s seismic waveforms with a sampling rate of 20 Hz) in length. Rectified Linear Unit is applied as the activation function between each layer (Nair & Hinton, 2010). The output is the probability of the acceptable and unacceptable measurements (Table S1 and Figure S1 in Supporting Information S1). The probability is

given by Softmax which is a popular activation function for classification problems (Goodfellow et al., 2016). The equation can be shown as:

$$p(x)_i = \frac{e^{x_i}}{\sum_{j=1}^2 e^{x_j}} \quad (1)$$

where $j = 1, 2$ and $i = 1, 2$ represent the 2 nodes of the final layer, and $p(x)_1$ and $p(x)_2$ represent the probability of acceptable and unacceptable measurements, respectively. If the probability of the acceptable measurements is greater than a threshold, this measurement is considered as an acceptable one. Because this is a bi-class classification problem and the training dataset is balanced, the threshold probability of acceptable measurements used in this study is 0.5.

To numerically reveal the difference between the CNN-predicted and human-labeled results, the cross-entropy loss is applied as the cost function (Goodfellow et al., 2016). The equation can be shown as:

$$L = -\sum_{i=1}^n p_i \log(q_i) \quad (n = 2) \quad (2)$$

where n is 2 representing the two types of measurements (acceptable and unacceptable), p is the probability of the CNN predicted result and q is the human-labeled result. The weights of the CNN are updated to minimize the cost function Adam algorithm (Kingma & Ba, 2014) with a learning rate of 0.001 during each training iteration. In each iteration, 100 measurements are randomly selected to train the CNN and each measurement in the training data set is used for 64 times. The training history of accuracy and loss value (Figure S2 in Supporting Information S1) show a high level of similarity between the trends of the curves for the training and validation data sets, suggesting a low probability of overfitting. In addition, both curves become nearly flat at the highest epoch numbers, which suggests a low possibility of underfitting.

4. TESTING WITH SYNTHETIC SWS MEASUREMENTS

Two sets of synthetic SWS measurements are generated based on Kong et al. (2015) to test the performance of the CNN. Firstly, the radial component of a pre-splitting XKS wave is defined as

$$R(t) = A_0 \sin(2\pi ft) e^{-\alpha t} \quad (3)$$

where $A_0 = 5000$ is the amplitude of the pre-splitting XKS wave, $f = 0.125$ Hz is the frequency, and α is the decaying factor which randomly changes from 0.1 to 0.5. After penetrating the anisotropic layer, which has a fixed ϕ of 0° and a randomly assigned δt ranging from 0.5 s to 2.0 s, the shear wave splits into the fast and slow components.

$$S_f(t) = R(t) \cos(\theta) \quad (4)$$

$$S_s(t) = -R(t - \delta t) \sin(\theta) \quad (5)$$

where S_f and S_s are the fast and slow waves, and θ is the angle between ϕ and the BAZ of the event. The epicentral distances of the events are randomly assigned in the range of 90° to 120° , and the focal depths vary from 20 to 50 km. Finally, S_f and S_s are projected to the north-south and east-west components and random noise is added to generate synthetic seismograms for SWS measurements. The SNR is defined as $\max(R(t))/\max(N(t))$, where $N(t)$ is the trace of the random noise.

It can be demonstrated mathematically that the energy on the original transverse component and the reliability of the resulting splitting parameters are dependent on θ , which is the angular difference between the BAZ and the fast orientation (Silver & Chan, 1991). In the modulo- 90° domain, close-to-null measurements, which are characterized by hardly observable XKS energy on the original transverse component, are dominant if

θ is less than 15° or greater than 75° . The relationship between θ and the reliability of the measurements, as well as the influence of θ on CNN's ability to correctly separate the acceptable measurements from the unacceptable ones, can be quantified using synthetic data.

For this purpose, we produce 72 groups of synthetic seismograms with SNR ranging from 4 to 10. Each group has 1000 measurements. The BAZ (which equals to θ because the fast orientation is set as 0° in the model) of the n -th group is $(n-1)*5^\circ$. The other parameters, including the SNR, used for generating the synthetic seismograms in Equations 3-5 are the same among the groups. The same data processing procedures used to generate the uniform SWS database for North America (Liu et al., 2014; Yang et al., 2016, 2017) are applied to the synthetic waveforms to automatically determine the splitting measurements, which are then grouped by the trained CNN into acceptable and unacceptable categories. The results suggest that the number of CNN-accepted

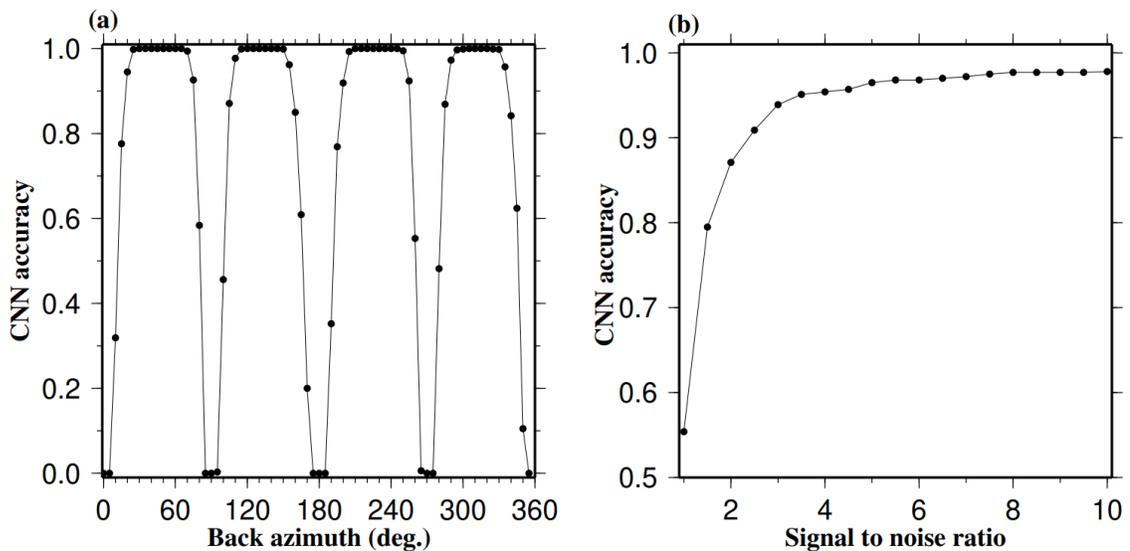


Figure 2. Performance tests of the CNN using synthetic dataset. (a) 72 groups of synthetic SWS measurements with different back azimuths. (b) 20 groups of synthetic SWS measurements with different signal-to-noise ratios.

measurements reduces rapidly when the BAZ approaches 0° and 90° in the modulo- 90° domain (Figure 2a), due to the weak XKS energy (relative to that of the noise) on the original transverse component. Such measurements are either ranked as C or N by human operators, depending on the noise level and the preference of the operators. When the BAZ is $\geq 15^\circ$ from the fast or slow orientations in the modulo- 90° domain, clear XKS energy is present on the original transverse component, and consequently the trained CNN successfully identified almost all the measurements with a rate of success $> 99\%$ (Figure 2a).

To test the performance of the CNN on data with different SNR values, 20 groups of synthetic SWS measurements are generated, each with 1000 measurements. The SNR of each group varies from 1 to 10 with an interval of 0.5. Based on the results of the first test (Figure 2a), the BAZ of each of the events used for the test is at least 15° from the fast or slow orientation. The results show that the success rate of CNN is over 90% for $\text{SNR} \geq 2.5$ and over 97% when $\text{SNR} \geq 6.5$ (Figure 2b).

5. APPLICATION TO SWS MEASUREMENTS IN SOUTH CENTRAL ALASKA

We next apply the trained CNN to broadband seismic data in south central Alaska recorded by 127 stations with variable recording period from 1988 to October 2019. The procedures to request (from the Data Management Center of the Incorporated Research Institutions for Seismology) and preprocess the XKS data follow those described in Liu and Gao (2013) and are identical to those used by Yang et al. (2021). In total 19,960 pairs of splitting parameters are obtained at 127 Stations (Figure 3a). The SNR based auto-

ranking procedure of Liu et al. (2008) ranked 6,314 measurements from 127 stations as potentially acceptable (Figure 3b) and 13,646 measurements as unacceptable. Note that the Liu et al. (2008) approach was intentionally designed for excluding only the

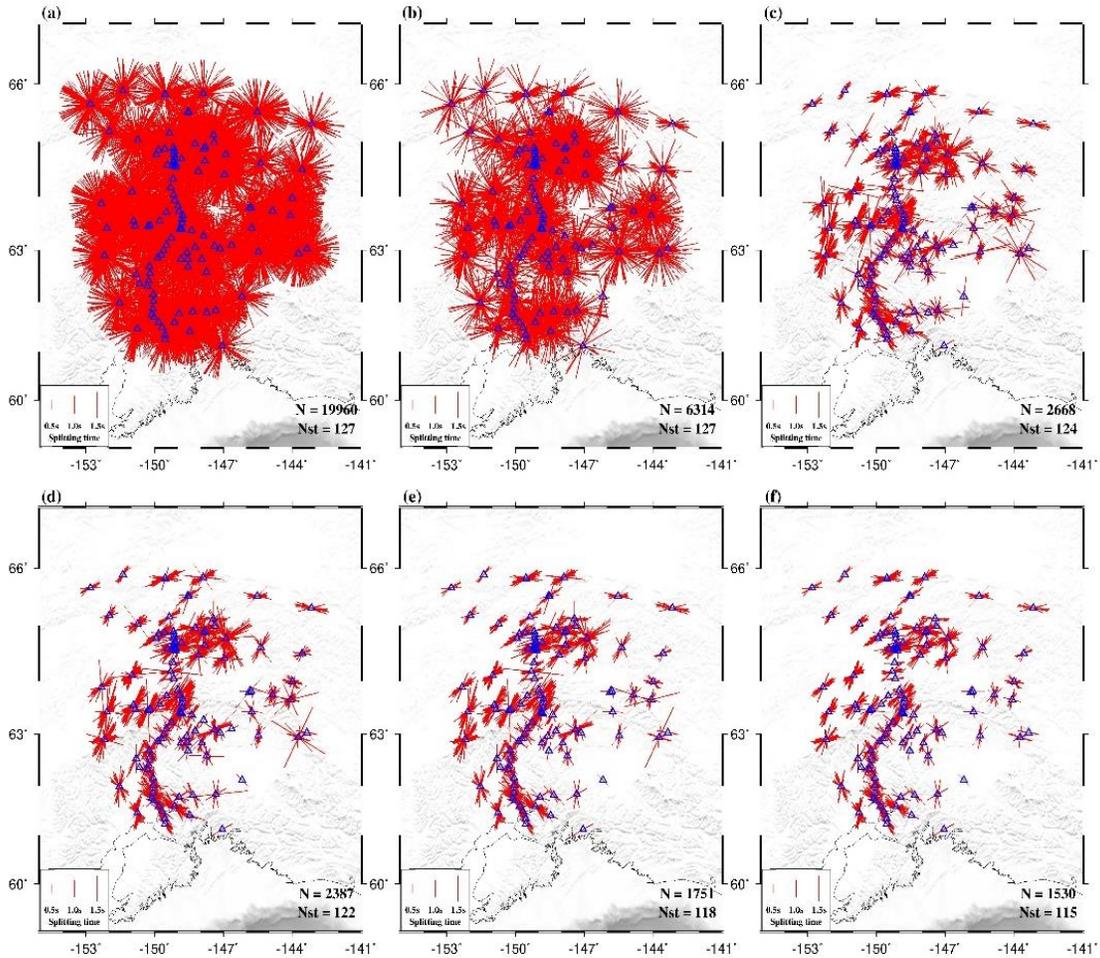


Figure 3. SWS measurements (red bars) in south central Alaska plotted at the recording stations (blue triangles). The orientation of the bars represents the fast orientation, and the length is proportional to the splitting time. (a) All the measurements recorded by stations in the study area. (b) Results of auto-ranking based on the approach of Liu et al. (2008) which was designed as a pre-screening step to reduce human workload in the subsequent manual screening step. (c) Results of CNN with a threshold of 0.5. (d) Same as (c) by after removing measurements for which the angular difference between the BAZ and the fast or slow wave polarization orientations is smaller than 15° . (e) Same as (d) but after removing measurements with standard deviation of $\phi > 15^\circ$ and standard deviation of $\delta t > 1.5$ s. (f) Same as (e) but after removing measurements with $\delta t > 2.0$ s. The number of measurements (N) and the number of stations (Nst) are shown in the lower right corner of each plot.

measurements that are impossible to be acceptable, for the purpose of reducing the amount of effort for the subsequent step of visual screening and at the same time minimizing the risk of missing potentially acceptable measurements.

The trained CNN is then applied to classify the 6,314 pairs of splitting parameters (Figure 3b). When a threshold value of 0.5 for the probability to be acceptable is used, a total of 2668 pairs from 124 stations are determined as acceptable (Figure 3c). We apply three set of conditions to refine the CNN-selected results. These conditions are necessary to exclude false positives caused by limitations in the current CNN, as described in Section 6.2 below. First, because the synthetic tests show that a measurement cannot be reliably classified by CNN (and human operators, as discussed below) when the difference between the BAZ and the fast orientation is less than $\sim 15^\circ$ (Figure 2a), we excluded such near-null measurements, and the remaining data set contains 2387 pairs of measurements from 122 stations (Figure 3d). Second, measurements with large standard deviations ($\geq 15^\circ$ for ϕ or ≥ 1.5 s for δt) are excluded, resulting in a total of 1,751 pairs of measurements from 118 stations (Figure 3e). Third, because measurements with large splitting times are rarely found in SWS studies in Alaska and elsewhere and are frequently associated with erroneously determined splitting parameters, we remove the 221 pairs (or 12.6%) of measurements with splitting times ≥ 2.0 s, leaving 1,530 pairs of measurements from 115 stations in the final data set (Figure 3f).

Understandably, as more conditions are applied to the CNN classified measurements, the number of remaining measurements reduces, while the consistency among the measurements at the stations increases (Figures 3c–3f). Note that due to the existence of complex anisotropy especially in the central and southern portions of the

study area (Yang et al., 2021), variability among the measurements at the same stations is present at some of the stations. Such a variability does not necessarily indicate inaccurately determined results but are mostly the result of azimuthal and piercing-point variations of the splitting parameters. The former is usually an indicator of complex anisotropy (Rümpker & Silver, 1998; Silver & Savage, 1994), and the latter is the result of a station being located at the boundary of two or more regions with different characteristics of anisotropy (Alsina & Snieder, 1995; Jia et al., 2021).

6. DISCUSSION

6.1. COMPARISON WITH HUMAN-DETERMINED MEASUREMENTS

After manually verifying the automatically ranked measurements (Figure 3b), Yang et al. (2021) obtained 971 measurements from 106 stations, among which 952 (98.1%) are within the 2668 measurements classified as acceptable by CNN before the application of the three conditions (Figure 3c). This suggests that if a human operator uses the CNN accepted results (Figure 3c) rather than results from the SNR based ranking system (Figure 3b) as the starting point for manual verification, an approximately 60% reduction in the number of measurements to be verified will be achieved, and only less than 2% of the measurements will be missed.

To objectively compare the human (Yang et al., 2021) and CNN determined final results (Figure 3f), the same set of three conditions applied to the CNN-accepted data are applied to the 971 human-determined measurements, and the remaining human-determined data set contains 865 measurements from 102 stations (Figure 4a), among

which 816 from 100 stations are in the final data set that CNN determined (Figure 3f). In other words, CNN missed merely 49 (5.7%) of the human determined measurements.

Most of the missed measurements are in Area B (Figure 4a) where the interaction of two flow systems with nearly orthogonal directions leads to weak anisotropy with small splitting times (Yang et al., 2021).

In spite of the fact that the station averaged splitting parameters from the CNN and human determined results show a high similarity with a cross-correlation coefficient (XCC) of 0.9631 for ϕ and 0.7947 for δt (Figures 4b and 4c), the number of

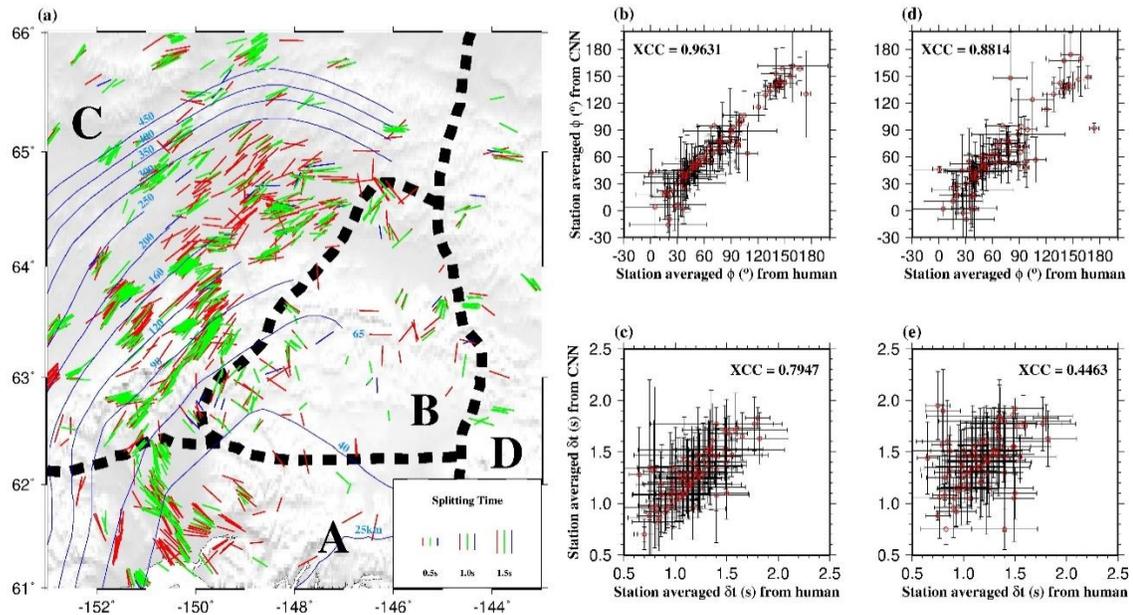


Figure 4. (a) Comparison of human-determined (blue bars; Yang et al., 2021) and CNN-selected (red bars) SWS measurements in south central Alaska. Green bars are measurements accepted by both CNN and human operators. All the measurements are plotted above the XKS ray-piercing points at 200 km deep which is the most likely depth of the anisotropic layer (Yang et al., 2021). The contour lines show the depth of the subducted Pacific slab, and the thick dashed lines separate four regions (A-D) with different patterns of splitting measurements (Yang et al., 2021). The CNN results are the same as those shown in Figure 3f. (b) Cross-plot of human and CNN determined station-averaged ϕ measurements. The black bars are the standard deviation. (c) Same as (b) but for station-averaged δt measurements. (d) and (e) are respectively the same as (b) and (c) but the CNN station averages were computed using results that were determined by CNN as acceptable but rejected by the human operators.

measurements in the final CNN-determined data set (Figure 3f) is about twice as many as that in the human-determined data set. This difference might be caused by the fact that the human operators applied a stricter set of standards when verifying the measurements. To test this possibility, we computed station averages using only the CNN determined measurements that were not selected by the human operators. The results (Figures 4d and 4e) show a reduced similarity between the two sets of data, with reduced XCC values of 0.8814 for ϕ and 0.4463 for δt . The most obvious explanation for this reduced similarity is that some measurements with a marginal quality were classified as acceptable by CNN but were rejected by the human operators.

6.2. COMPARISON WITH A FULLY AUTOMATED NON-MACHINE LEARNING SWS MEASUREMENT APPROACH

Several non-machine learning methods have been proposed to measure SWS parameters in a completely automated manner (e.g., Link et al., 2022; Teanby et al., 2004). Here, we choose the latest one, SplitRacerAUTO (Link et al., 2022), to compare with the CNN-based approach proposed in this study. The MATLAB-based SplitRacerAUTO can automatically select the XKS time window and categorize the splitting measurements. The same data set in Alaska is applied to test the performance of SplitRacerAUTO against our ML-based approach. To be consistent with the parameters used in our preprocessing step, we used the frequency range of 0.04–0.5 Hz for the band-pass filter in SplitRacerAUTO, and kept all the other parameters the same as the default values. The results show that this method accepts 950 measurements from 110 stations, and 467 of them (48.1%) are contained in results of Yang et al. (2021). After applying the

same constraints that produced Figure 3f, there are 586 measurements remained and 404 (46.7%) of them are in Yang et al. (2021). In other words, the automated procedure missed 53.3% of the human-determined measurements, while our CNN-based approach merely missed 5.7% of them. The results of the automatically determined splitting parameters and results after various constraints are plotted in Figure S4 of Supporting Information S1 using the same style as Figure 3 for easy comparison. Similarly, comparisons of human and SplitRacerAUTO determined measurements similar to Figure 4 are plotted in Figure S5 of Supporting Information S1. Comparing Figures 4 and S5 in Supporting Information S1, it is clear that the CNN-based approach resulted in a significantly greater number of measurements than SplitRacerAUTO, especially in areas with weaker anisotropy such as Area B in Figures 4 and S5 in Supporting Information S1.

6.3. LIMITATIONS OF THE CURRENT CNN AND SUGGESTED NEXT STEPS

In spite of the satisfactory performance of our CNN on both synthetic and real data, a major drawback of the current CNN is that it does not have the capability to adjust the data processing parameters including the beginning and end times of the XKS window and the bandpass filtering frequencies. For a small portion of the measurements, such adjustments are required in order to obtain reliable results. For instance, if the epicentral distance is smaller than 90° , part of the S wave energy can be included in the XKS window, leading to unreliable results. Therefore, it is necessary to design and train a CNN that can automatically recognize the necessity and make such adjustments. One of the approaches is to design a separate pre-processing CNN for picking the arrival time of the XKS arrival, similar to those designed for picking the onset time of P or S waves

from local events (Zhu & Beroza, 2018). Additional work is needed to find the optimal ending time of the XKS window, and to detect the dominant frequency range of the noise and perform band pass filtering to enhance the SNR when strong noise is present.

Alternatively, the optimal XKS window can be determined during the pre-processing stage using non-CNN based approaches such as the time-frequency spectrum technique recently proposed by Link et al. (2022).

7. CONCLUSION

In this study, we have established a CNN to automatically classify teleseismic SWS measurements. The CNN is trained by published human-labeled datasets and tested using synthetic SWS measurements to evaluate its performance against different levels of noise and its dependence on the difference between the fast orientation and the back-azimuth of the events. When the SNR is greater than 6.5, more than 97% of the non-null synthetic measurements can be correctly accepted by the CNN. Application of the CNN to data from south central Alaska shows that it can classify almost all human-accepted measurements (98.1%) as acceptable when a threshold probability of 0.5 is used. The study suggests a high potential for CNN-based methods to significantly improve the efficiency of measuring SWS parameters.

ACKNOWLEDGEMENTS

The facilities of IRIS Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms and related metadata used in this study. IRIS

Data Services are funded through the Seismological Facilities for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Support Agreement EAR-1851048. We thank Frederik Link, an anonymous reviewer, and Editor Daoyuan Sun for constructive reviews that significantly improved the manuscript. The study was partially supported by the U.S. National Science Foundation under awards 1830644 and 1919789 to S.G.

DATA AVAILABILITY STATEMENT

The Python codes to train and test the CNN have been uploaded to GitHub under the address <https://github.com/YW-Zhang94/CNN-SWS.git>, and the seismic data used in the study (in SAC-Seismic Analysis Codes format with a total size of about 1.3 GB) can be found under https://web.mst.edu/~yzcd4/21a_CNN_SWS_data/. All the waveform data used in the study are openly accessible from the IRIS Data Management Center (<https://ds.iris.edu/ds/nodes/dmc/>, last accessed March 2019), under the main network codes of AK (<https://doi.org/10.7914/SN/AK>), AT (<https://doi.org/10.7914/SN/AT>), DW (<https://doi.org/10.7914/SN/DW>), IM (International Miscellaneous Stations), IU (<https://doi.org/10.7914/SN/IU>), TA (<https://doi.org/10.7914/SN/TA>), XE (https://doi.org/10.7914/SN/XE_1999), XR (https://doi.org/10.7914/SN/XR_2004), XV (https://doi.org/10.7914/SN/XV_2014), YE (https://doi.org/10.7914/SN/YE_2011), and YV (https://doi.org/10.7914/SN/YV_2006).

SUPPORTING INFORMATION

Table S1. Structure of the convolutional neural network used in the study

	Length	Depth	Kernel size	Stride	Activation function
Input	1000	4	-	-	-
Conv-1D	500	32	3	2	ReLU
Conv-1D	125	32	3	2	ReLU
Conv-1D	62	32	3	2	ReLU
Conv-1D	31	32	3	2	ReLU
Conv-1D	15	32	3	2	ReLU
Conv-1D	7	32	3	2	ReLU
Conv-1D	4	32	3	2	ReLU
Conv-1D	2	32	3	2	ReLU
Flatten	64	1	-	-	-
Output	2	1	-	-	Softmax

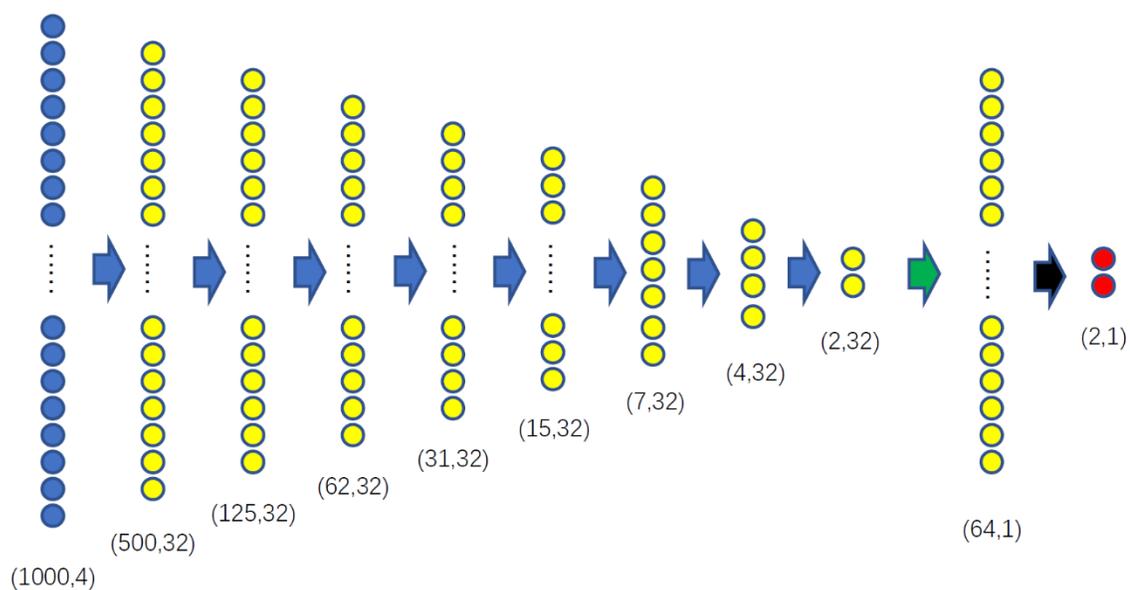


Figure S1. The structure of the convolutional neural network used in the study. The blue arrows are convolutional layers with ReLU, the green arrow represents the flattened layer, and the black arrow is a full-connect layer with Softmax. The blue dots are input data, yellow dots are nodes in the network, and the red dots are the output of the network. The length and depth are noted at the bottom of the columns of dots.

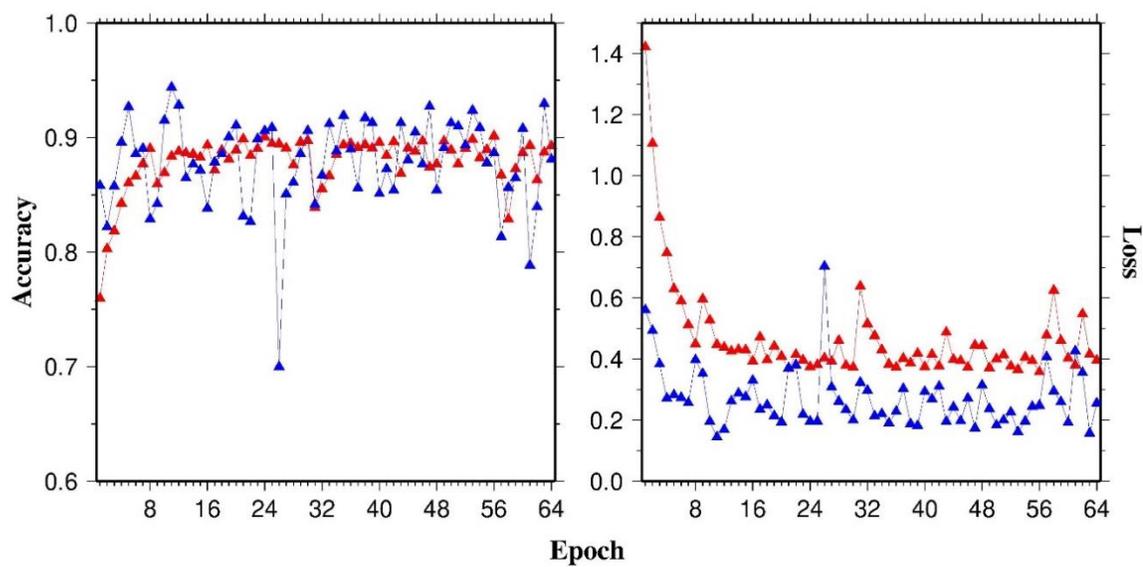


Figure S2. Accuracy and Loss in each epoch during training CNN. The red curves and triangles are based on training dataset. The blue curves and triangles are based on validation dataset.

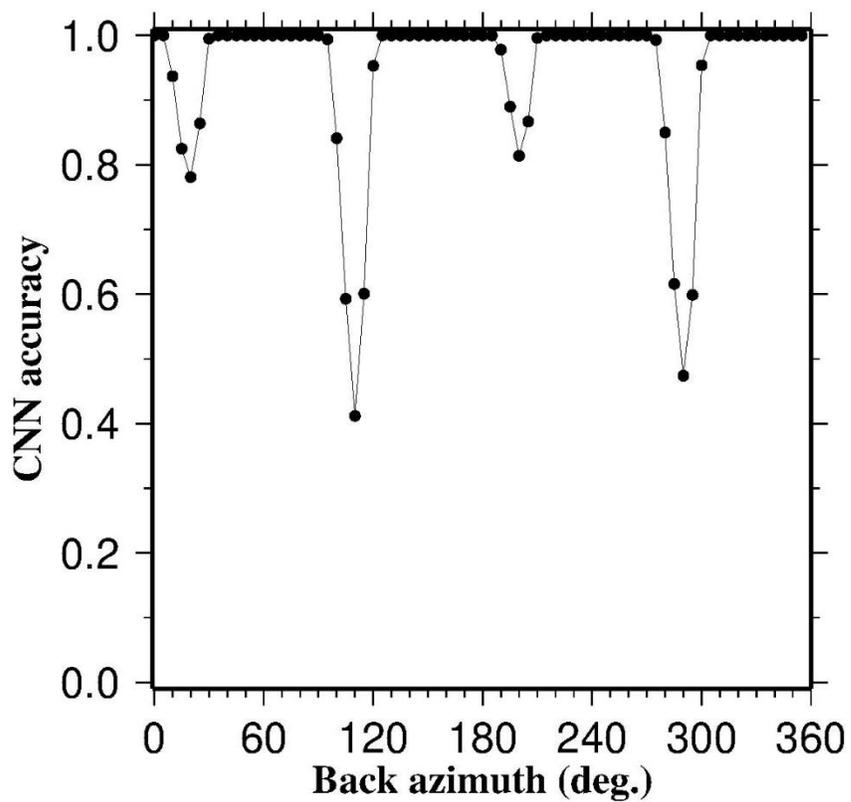


Figure S3. Performance tests of the CNN using synthetic dataset based on 2 anisotropic layers.

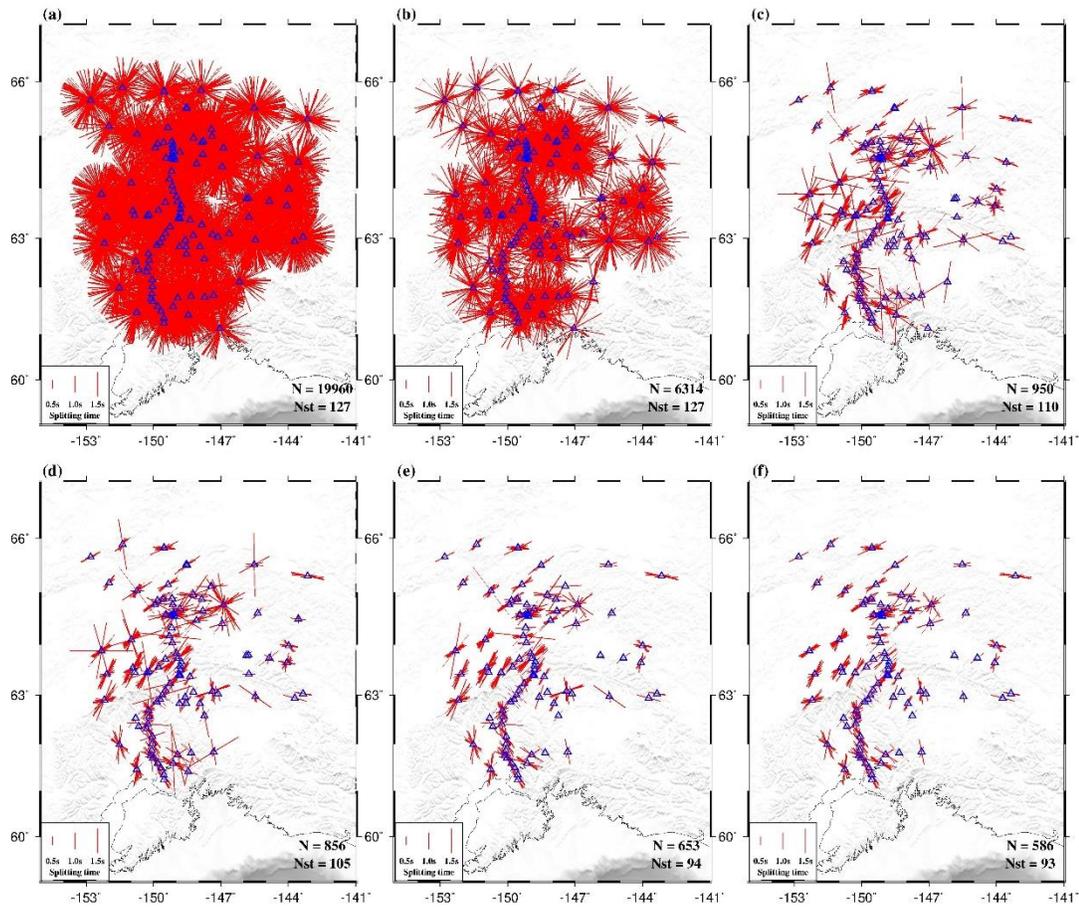


Figure S4. SWS measurements (red bars) in south central Alaska plotted at the recording stations (blue triangles). The orientation of the bars represents the fast orientation, and the length is proportional to the splitting time. (a) All the measurements recorded by stations in the study area. (b) Results of auto-ranking based on the approach of Liu et al. (2008) which was designed as a pre-screening step to reduce human workload in the subsequent manual screening step. (c) Results of auto-checking method from Link et al. (2022). (d) Same as (c) but after removing measurements for which the angular difference between the BAZ and the fast or slow wave polarization orientations is smaller than 15° . (e) Same as (d) but after removing measurements with standard deviation of $\phi > 15^\circ$ and standard deviation of $\delta t > 1.5$ s. (f) Same as (e) but after removing measurements with $\delta t > 2.0$ s. The number of measurements (N) and the number of stations (Nst) are shown in the lower right corner of each plot.

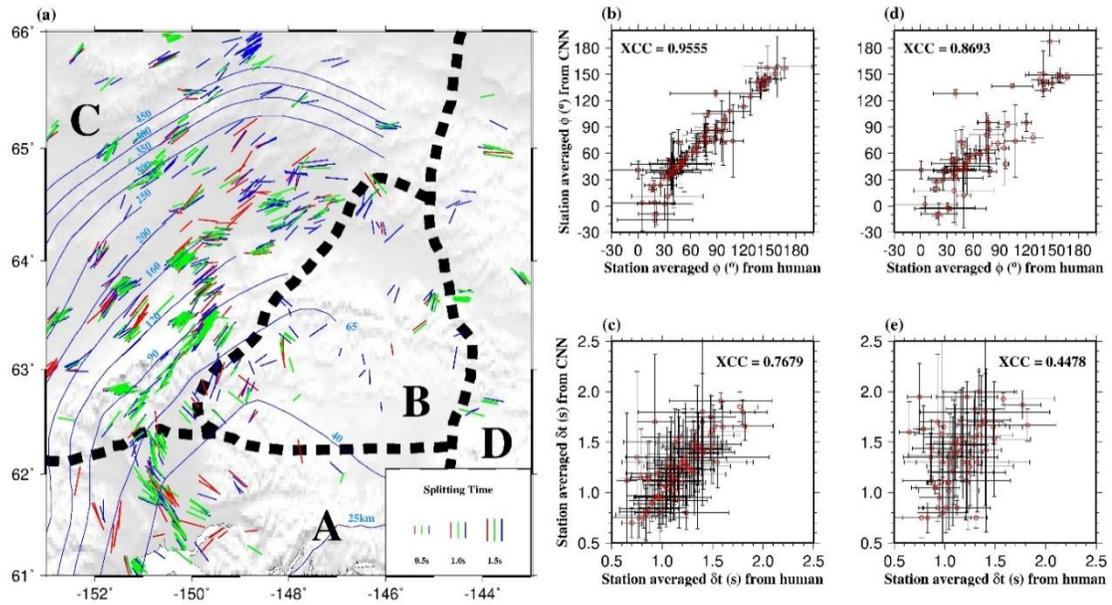


Figure S5. (a) Comparison of human-determined (blue bars; Yang et al., 2021) SWS measurements and results from auto-checking method (red bars; Link et al., 2022) in south central Alaska. Green bars are measurements accepted by both auto-checking method and human operators. All the measurements are plotted above the XKS ray-piercing points at 200 km deep which is the most likely depth of the anisotropic layer (Yang et al., 2021). The contour lines show the depth of the subducted Pacific slab, and the thick dashed lines separate four regions (A-D) with different patterns of splitting measurements (Yang et al., 2021). The auto-checking method results are the same as those shown in Figure 3f. (b) Cross-plot of human and auto-checking determined station-averaged ϕ measurements. The black bars are the standard deviation. (c) Same as (b) but for station-averaged δt measurements. (d) and (e) are respectively the same as (b) and (c) but the station averages from auto-checking method were computed using results that were determined by auto-checking method as acceptable but rejected by the human operators.

REFERENCES

- Alsina, D., & Snieder, R. (1995). Small-scale sublithospheric continental mantle deformation: Constraints from SKS splitting observations. *Geophysical Journal International*, 123(2), 431–448. <https://doi.org/10.1111/j.1365-246X.1995.tb06864.x>
- Ando, M., Ishikawa, Y., & Yamazaki, F. (1983). Shear wave polarization anisotropy in the upper mantle beneath Honshu, Japan. *Journal of Geophysical Research*, 88(B7), 5850–5864. <https://doi.org/10.1029/JB088iB07p05850>
- Bianco, M. J., Gerstoft, P., Olsen, K. B., & Lin, F. C. (2019). High-resolution seismic tomography of Long Beach, CA using machine learning. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-50381-z>
- Dokht, R. M., Kao, H., Visser, R., & Smith, B. (2019). Seismic event and phase detection using time–frequency representation and convolutional neural networks. *Seismological Research Letters*, 90(2A), 481–490. <https://doi.org/10.1785/0220180308>
- Garcia, J. A., Waszek, L., Tauzin, B., & Schmerr, N. (2021). Automatic identification of mantle seismic phases using a Convolutional Neural Network. *Geophysical Research Letters*, 48(18), e2020GL091658. <https://doi.org/10.1029/2020GL091658>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Machine learning basics. *Deep learning*, 1(7), 98–164.
- Gulli, A., & Pal, S. (2017). *Deep learning with Keras*. Packt Publishing Ltd.
- Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis*, 6(5), 429–449. <https://doi.org/10.3233/IDA-2002-6504>
- Jia, Y., Liu, K. H., Kong, F., Liu, L., & Gao, S. S. (2021). A systematic investigation of piercing-point-dependent seismic azimuthal anisotropy. *Geophysical Journal International*, 227(3), 1496–1511. <https://doi.org/10.1093/gji/ggab285>
- Katayama, I., & Karato, S. I. (2006). Effect of temperature on the B-to C-type olivine fabric transition and implication for flow pattern in subduction zones. *Physics of the Earth and Planetary Interiors*, 157(1–2), 33–45. <https://doi.org/10.1016/j.pepi.2006.03.005>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

- Kong, F., Gao, S. S., & Liu, K. H. (2015). A systematic comparison of the transverse energy minimization and splitting intensity techniques for measuring shear-wave splitting parameters. *Bulletin of the Seismological Society of America*, 105(1), 230–239. <https://doi.org/10.1785/0120140108>
- Li, Z., Meier, M. A., Hauksson, E., Zhan, Z., & Andrews, J. (2018). Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophysical Research Letters*, 45(10), 4773–4779. <https://doi.org/10.1029/2018GL077870>
- Link, F., Reiss, M. C., & Rumpker, G. (2022). An automatized XKS-splitting procedure for large data sets: Extension package for SplitRacer and application to the USArray. *Computers & Geosciences*, 158, 104961. <https://doi.org/10.1016/j.cageo.2021.104961>
- Linville, L., Pankow, K., & Draelos, T. (2019). Deep learning models augment analyst decisions for event discrimination. *Geophysical Research Letters*, 46(7), 3643–3651. <https://doi.org/10.1029/2018GL081119>
- Liu, K. H., Elsheikh, A., Lemnifi, A., Purevsuren, U., Ray, M., Refayee, H., et al. (2014). A uniform database of teleseismic shear wave splitting measurements for the western and central United States. *Geochemistry, Geophysics, Geosystems*, 15(5), 2075–2085. <https://doi.org/10.1002/2014GC005267>
- Liu, K. H., & Gao, S. S. (2013). Making reliable shear-wave splitting measurements. *Bulletin of the Seismological Society of America*, 103(5), 2680–2693. <https://doi.org/10.1785/0120120355>
- Liu, K. H., Gao, S. S., Gao, Y., & Wu, J. (2008). Shear wave splitting and mantle flow associated with the deflected Pacific slab beneath northeast Asia. *Journal of Geophysical Research*, 113, B01305. <https://doi.org/10.1029/2007JB005178>
- Lomax, A., Michelini, A., & Jozinović, D. (2019). An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network. *Seismological Research Letters*, 90(2A), 517–529. <https://doi.org/10.1785/0220180311>
- Long, M. D., & Silver, P. G. (2009). Shear wave splitting and mantle anisotropy: Measurements, interpretations, and new directions. *Surveys in Geophysics*, 30(4), 407–461. <https://doi.org/10.1007/s10712-009-9075-1>
- McBrearty, I. W., Delorey, A. A., & Johnson, P. A. (2019). Pairwise association of seismic arrivals with convolutional neural networks. *Seismological Research Letters*, 90(2A), 503–509. <https://doi.org/10.1785/0220180326>

- Mignan, A., & Broccardo, M. (2019). One neuron versus deep learning in aftershock prediction. *Nature*, 574(7776), E1–E3. <https://doi.org/10.1038/s41586-018-0438-y>
- Mousavi, S. M., & Beroza, G. C. (2020). A machine-learning approach for earthquake magnitude estimation. *Geophysical Research Letters*, 47(1), e2019GL085976. <https://doi.org/10.1029/2019GL085976>
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th international conference on international conference on machine learning*, 807–814.
- Peng, Z., & Ben-Zion, Y. (2004). Systematic analysis of crustal anisotropy along the Karadere–Duzce branch of the North Anatolian fault. *Geophysical Journal International*, 159, 253–274. <https://doi.org/10.1111/j.1365-246X.2004.02379.x>
- Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2), e1700578. <https://doi.org/10.1126/sciadv.1700578>
- Ross, Z. E., Meier, M. A., & Hauksson, E. (2018). P wave arrival picking and first-motion polarity determination with deep learning. *Journal of Geophysical Research: Solid Earth*, 123, 5120–5129. <https://doi.org/10.1029/2017JB015251>
- Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes. *Geophysical Research Letters*, 44(18), 9276–9282. <https://doi.org/10.1002/2017GL074677>
- Rümpker, G., & Silver, P. G. (1998). Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy. *Geophysical Journal International*, 135(3), 790–800. <https://doi.org/10.1046/j.1365-246X.1998.00660.x>
- Savage, M. K. (1999). Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting? *Reviews of Geophysics*, 37(1), 65–106. <https://doi.org/10.1029/98RG02075>
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484. <https://doi.org/10.1038/nature16961>
- Silver, P. G. (1996). Seismic anisotropy beneath the continents: Probing the depths of geology. *Annual Review of Earth and Planetary Sciences*, 24(1), 385–432. <https://doi.org/10.1038/nature16961>

- Silver, P. G., & Chan, W. W. (1991). Shear wave splitting and subcontinental mantle deformation. *Journal of Geophysical Research*, 96(B10), 16429–16454. <https://doi.org/10.1029/91JB00899>
- Silver, P. G., & Savage, M. K. (1994). The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers. *Geophysical Journal International*, 119(3), 949–963. <https://doi.org/10.1111/j.1365-246X.1994.tb04027.x>
- Teanby, N. A., Kendall, J. M., & van der Baan, M. (2004). Automation of shear-wave splitting measurements using cluster analysis. *Bulletin of the Seismological Society of America*, 94(2), 453–463. <https://doi.org/10.1785/0120030123>
- Titos, M., Bueno, A., Garcia, L., & Benitez, C. (2018). A deep neural networks approach to automatic recognition systems for volcano-seismic events. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1533–1544. <https://doi.org/10.1109/JSTARS.2018.2803198>
- Vecsey, L., Plomerová, J., & Babuška, V. (2008). Shear-wave splitting measurements—Problems and solutions. *Tectonophysics*, 462(1–4), 178–196. <https://doi.org/10.1016/j.tecto.2008.01.021>
- Woollam, J., Rietbrock, A., Bueno, A., & De Angelis, S. (2019). Convolutional neural network for seismic phase classification, performance demonstration over a local seismic network. *Seismological Research Letters*, 90(2A), 491–502. <https://doi.org/10.1785/0220180312>
- Yang, B. B., Liu, K. H., Dahm, H. H., & Gao, S. S. (2016). A uniform database of teleseismic shear-wave splitting measurements for the western and central United States: December 2014 update. *Seismological Research Letters*, 87(2A), 295–300. <https://doi.org/10.1785/0220150213>
- Yang, B. B., Liu, Y., Dahm, H., Liu, K. H., & Gao, S. S. (2017). Seismic azimuthal anisotropy beneath the eastern United States and its geodynamic implications. *Geophysical Research Letters*, 44(6), 2670–2678. <https://doi.org/10.1002/2016GL071227>
- Yang, Y., Gao, S. S., Liu, K. H., Kong, F., & Fu, X. (2021). Mantle flow in the vicinity of the eastern edge of the Pacific-Yakutat slab: Constraints from shear wave splitting analyses. *Journal of Geophysical Research: Solid Earth*, 126, e2021JB022354. <https://doi.org/10.1029/2021JB022354>
- Zhang, S., & Karato, S. I. (1995). Lattice preferred orientation of olivine in simple shear deformation and the flow geometry of the upper mantle of the Earth. *Nature*, 375, 774–777. <https://doi.org/10.1038/375774a0>

- Zhu, W., & Beroza, G. C. (2018). 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>
- Zhu, W., Mousavi, S. M., & Beroza, G. C. (2019). Seismic signal denoising and decomposition using deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(11), 9476–9488.
<https://doi.org/10.1109/TGRS.2019.2926772>

II. CONVOLUTIONAL NEURAL NETWORK CLASSIFICATION OF NATURAL EARTHQUAKES, MINE COLLAPSES AND EXPLOSIONS

Yanwei Zhang¹, Junhao Qu², Shaohui Zhou², Xiaoyi Fan³, and Stephen S. Gao¹

1. Department of Geoscience and Geological and Petroleum Engineering, Missouri University of Science and Technology, Rolla, MO, 65409
2. Shandong Seismic Agency, Jinan, Shandong, China
3. Tai'an Fiducial Seismic Station of Shandong Earthquake Agency, Tai'an, Shandong, China

ABSTRACT

Convolutional neural network (CNN) has been widely applied in geophysics and shows outstanding performance in various sub-areas of geophysical research. Usually, a huge amount of high-quality data is necessary to train the CNN in order to obtain reliable results. Mis-labeled and ambiguous measurements in the dataset would negatively influence the training process and reduce the performance of the CNN. In this study, we established a CNN to classify natural earthquakes, mine collapses, and explosions using seismic waveforms recorded by 287 stations in Shandong Province, China. The data set contains 1035 earthquakes, 159 mine collapses, and 586 explosions. In order to reduce the influence of unreliable measurements in the dataset, cross-validation is employed to scan the whole dataset. The measurements with different labels between human and the CNN are manually assessed and kept, corrected, or abandoned in dataset. Testing with the new dataset, the classification accuracies of the three types of events obviously increase and all above 95%. The CNN shows over-human behavior in this task and the performance is highly influenced by the quality and distribution of the dataset.

1. INTRODUCTION

With the rapidly developing demand of mineral products in the modern society, the number of non-tectonic earthquake events, like mine collapses and explosions, quickly increase as well. Therefore, reliable classification of natural earthquakes, collapses, and explosions becomes a common and tough challenge in applied seismological research (Astiz et al., 2014).

Because the events often occurred in deserted and hard-to-access areas, the features from seismic waveform are significant criteria for analysts to make event labels of tectonic and non-tectonic earthquakes. Several different automated and semiautomated methods have been developed to classify the source based on seismic waveforms. Fäh and Koch (2002) compared the ratio between P and S phases with various time windows and frequency bands to discriminate earthquakes and chemical explosions. The Lg phases or Rg phases are also analyzed to classify different types of sources (Douglas et al., 1990; Rourke and Baker, 2017). These studies reveal that seismic waveforms contain significant information of the event source, but the performance of the waveform-based techniques is highly dependent on the depth and distance of the event and background environments. Traditionally, therefore, classification of earthquakes, mine collapses, and explosions are operated by human analysts.

Recently, high-performance Machine Learning (ML) techniques have attracted the attention of geophysical researchers. Relied on the outstanding ability of feature extraction, convolutional neural network (CNN), one of the most common ML techniques, shows considerable potential in classification problems based on seismic waveforms (Perol et al., 2018; Zhu & Beroza, 2018; Linville et al., 2019; Zhang and Gao,

2022). Usually, a huge amount of high-quality data is necessary to train the CNN in order to obtain reliable results. Some previous studies, however, pointed out that mis-labeled and ambiguous measurements commonly existing in human labeled dataset can degenerate the results (Zhu & Beroza, 2018; Garcia et al., 2021; Zhang and Gao, 2022). Linville et al. (2019) reported over 50% CNN misclassified events were mislabeled by human analysts. It is the common sense in ML world that the dataset is more important than the model itself. These low-quality measurements would negatively influence the training process and reduce the performance of the CNN.

In this study, we established a 11-layers CNN on the task of classification for tectonic earthquakes, mine collapses and explosions. The human labeled dataset is from Shangdong Seismic Network Center (SSNC). To minimize the influence of unreliable measurements in the dataset, a ten-fold cross validation is employed to scan the whole dataset. The measurements with different labels between human and CNN are manually assessed and kept, corrected, or abandoned in the dataset. Comparing with the original dataset, the classification accuracies of three types of events increase from 97.3% to 99.2% for earthquake, from 84.9% to 95.8% for mine collapses, and from 93.6% to 98.1% for explosion. Our results reveal that the unreliable measurements have negative effects in ML studies and indicate that cross-validation with CNN can evaluate, correct, and enhance the dataset.

2. DATA AND METHOD

SSNC provides a seismic dataset recorded during the period from Aug. 2017 to Jan. 2022. Totally 31754 three-component seismic waveforms from 4410 events were recorded by 287 seismic stations (Figure 1). The P-wave arrival time of each waveform is manually picked and the corresponding event is categorized into earthquake, collapse, or explosion by human experts. Based on the determination method of magnitude of completeness (M_c) from Cao and Gao (2002), the earthquake frequency-magnitude distribution reveals that the M_c of the dataset from the SSNC is 1.5. Thus we excluded all

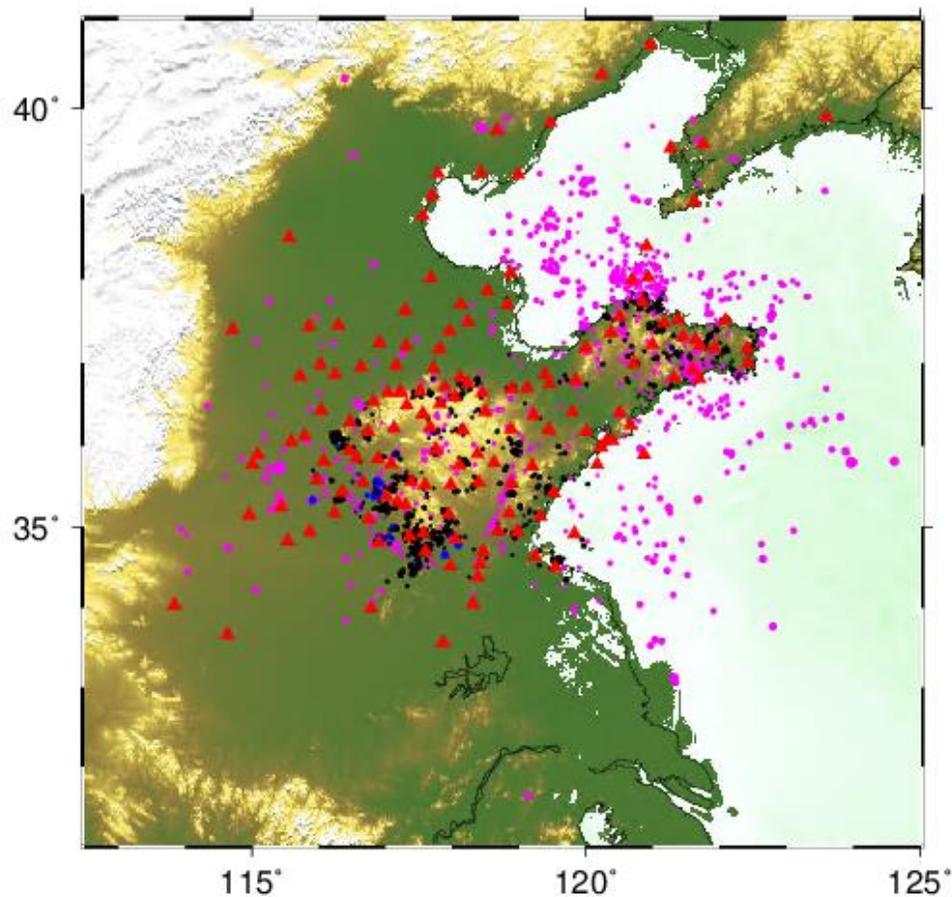


Figure 1. Stations and events distribution of dataset in the study area. The dataset is from SSNC. The red triangles are stations, the purple dots are earthquakes, the blue dots are collapses, and the black dots are explosions.

events whose magnitudes are less than 1.5. Finally, 17,557 seismic waveforms are kept in our dataset, including 11901 waveforms from 1035 earthquakes, 1511 waveforms from 159 collapses, and 4145 waveforms from 586 explosions. After detrending, seismic waveforms in the time window 10 s before and 40 s following the P-wave arrival time are used in this study, and the sampling interval is 0.01 s. Zeros are appended to the seismic waveforms that do not have enough length.

In our dataset, the number of waveforms from each event are significantly unbalanced. This situation would cause the overfitting of CNN (Goodfellow et al., 2016). To balance our dataset, different class weights are set for different types of the events (1 for the earthquakes, 8 for the collapses, and 2.5 for the explosions), which is a common approach to treat unbalanced situations (Japkowicz & Stephen, 2002; Zhang & Gao, 2022). The earthquakes are labeled as array [1 0 0], the collapses are labeled as array [0 1 0], and the explosions are labeled as array [0 0 1].

Based on practice from previous studies, we established a CNN with 10 convolutional layers and a full connected layer using Keras (Gulli & Pal, 2017) (Figure 2). The inputs are 5000 nodes long three-component seismic waveforms and after

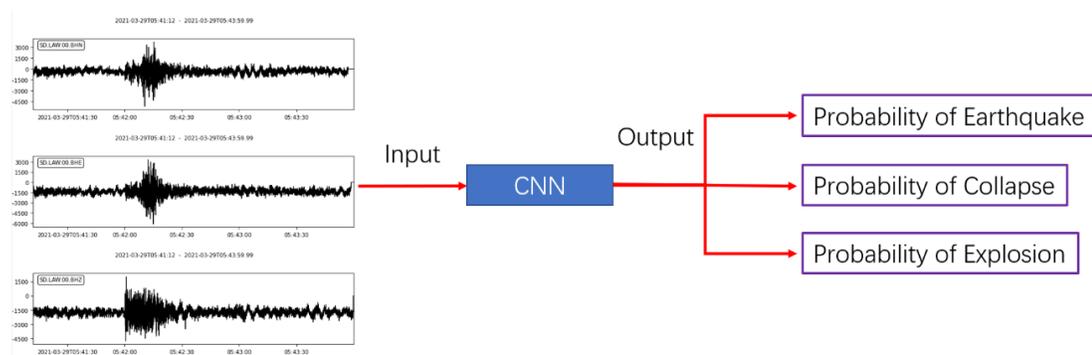


Figure 2. The workflow of CNN in this study. The input of CNN is 3 component (N, E, and Z) seismic waveforms. The output of CNN is 3 probabilities of natural earthquakes, collapses, and explosions.

convolution, the outputs are the probability of each of the three-types of events. The activation function between each layer is LeakyReLU with a 0.05 negative slope (Maas et al., 2013). During the convolutional process, 3 by 1 filters are applied to extract the features from seismic waveforms. The stride is set to 2, so the length of each layer is down sampling to half of the previous one. The Softmax is used as the activation function to classify each seismic waveform in the output layer (Goodfellow et al., 2016):

$$p(x)_i = \frac{e^{x_i}}{\sum_{j=1}^3 e^{x_j}} \quad (1)$$

where $j = 1, 2, 3$ and $i = 1, 2, 3$ represent the 3 nodes of the last layer, and $p(x)_1$, $p(x)_2$, and $p(x)_3$ represent the probability of earthquakes, collapses, and explosions. The result of CNN follows the highest value of $p(x)$.

The cross-entropy between human-labeled and CNN-predicted results is defined as the loss function of our CNN (Goodfellow et al., 2016):

$$L = - \sum_{i=1}^n p_i \log(q_i) \quad (n = 3) \quad (2)$$

where n is 3 representing the three types of events (earthquakes, collapses, and explosions), p is the probability of the CNN-predicted result given by Softmax, and q is for human-labeled result. The weights of the filters and fully connected layer are automatically optimized by minimizing the loss function during the training process. The optimizer is Adam with a learning rate of 0.001 (Kingma & Ba, 2014), and the epoch is 64.

3. RESULTS

Firstly, a ten-fold cross validation is employed to scan all the dataset. The dataset is evenly separated into 10 groups and at each iteration, 9 groups serve as the training dataset and 1 group serves as the testing dataset. Because each event can be received by multiple stations, the majority vote algorithm is applied to determine each event in the network level. The results will be the sum of output of all the stations that recorded the event. Finally, the CNN successfully recognized 97.3% (1007/1035) earthquakes, 93.7% (549/586) explosions, and 84.9% (135/159) collapses (Figure 3a, b, and c).

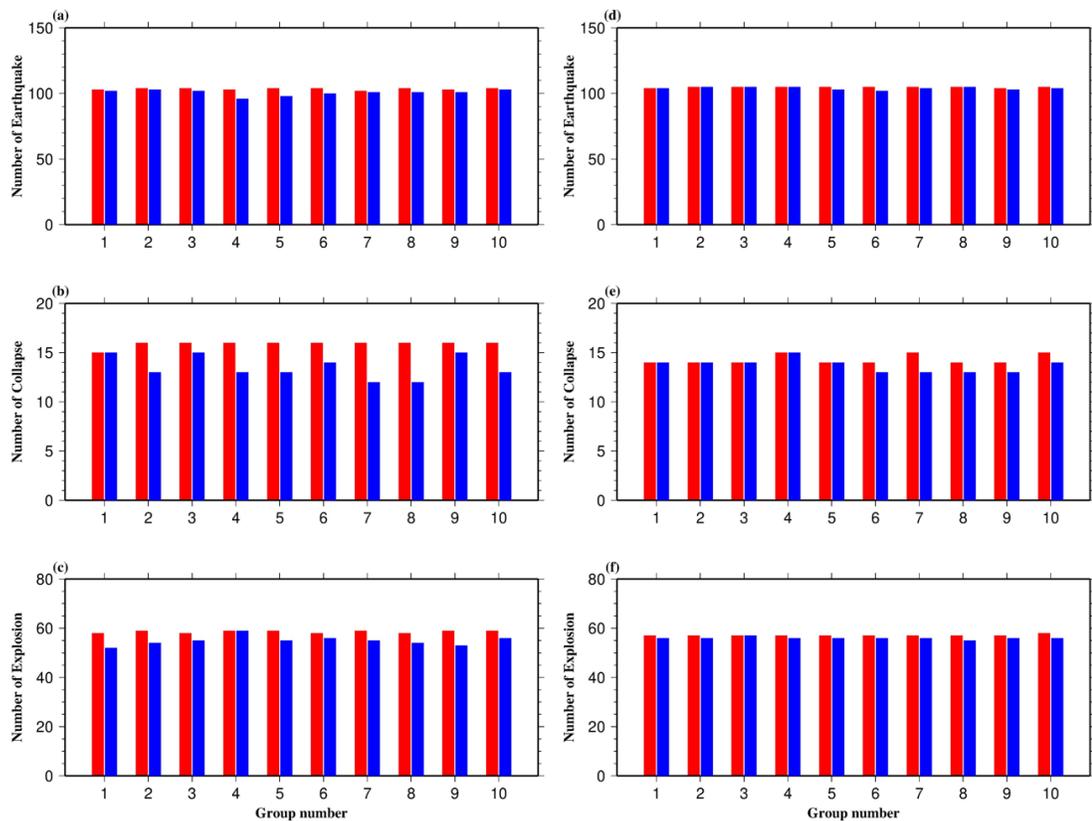


Figure 3. The performance of the CNN. The red bars are number of each type in each group. The blue bars are number of events with same label between CNN and human. (a), (b), and (c) are results with original dataset. (d), (e), and (f) are results with corrected dataset.

All the measurements which have different labels between the CNN and human are manually assessed. 10 out of 28 earthquakes, 18 out of 24 collapses, and 18 out of 37 explosions are considered as mislabeled by human; 5 earthquakes, 1 collapse, and 12 explosions are hard to classify due to unclear features on the seismograms. Even assume human is correct with all the ambiguous events, there are still over 50% measurements mislabeled by human in different-labels events. In the other word, the CNN shows over-human behavior with original dataset.

Themis-labeled and ambiguous measurements commonly exist in various datasets and would decrease the accuracy of ML models. To minimize the influence of unreliable measurements, we corrected all the mis-labeled measurements and removed ambiguous ones with unclear features. The same ten-fold cross validation is employed again with the corrected dataset. After the same process as that used for the first scanning, the accuracies of CNN increase from 97.3% to 99.2% (1040/1048) for natural earthquakes, from 84.9% to 95.9% (137/143) for collapses, and from 93.6% to 98.1% (560/571) for explosions (Figure 3d, e, and f). The overall accuracy of each type is over 95% with the corrected dataset. The CNN mislabeled 25 events, and 10 of them are the same as those in the first training with the original data, and 15 of them are unique in the second training with the corrected dataset. Additionally, the results of all the corrected events are same with the new labels.. Therefore, the CNN agrees all the manually corrected events and gives consistent label with human for 15 uncorrected mislabeled events in the first training. We also manually assessed the 15 unique events, and found that 1 earthquake and 1 collapse were considered as mislabeled by human, and features of the 2 explosions and 1 earthquake are unclear on the seismograms.

4. DISCUSSION

After correcting human mislabeled events and removing ambiguous events, the accuracy of CNN has obviously increased. For the true mislabeled events, most of them were given vague results by the CNN for which the maximum value is close to the other two values; Moreover, even after balance, the accuracy of CNN is higher with a larger number of events. The large dataset would include various situations and the CNN has a higher probability to avoid overfitting.

The distribution of true mislabeled events in the CNN results is evaluated. We plot the accuracy of the CNN versus magnitude and the number of measurements of each event. As Figure 4 shows, the accuracy of the CNN increases with the magnitude and the number of measurements, and when the magnitude larger is than 2.5 and the number of measurements is larger than 20, the CNN can successfully classify all the measurements.

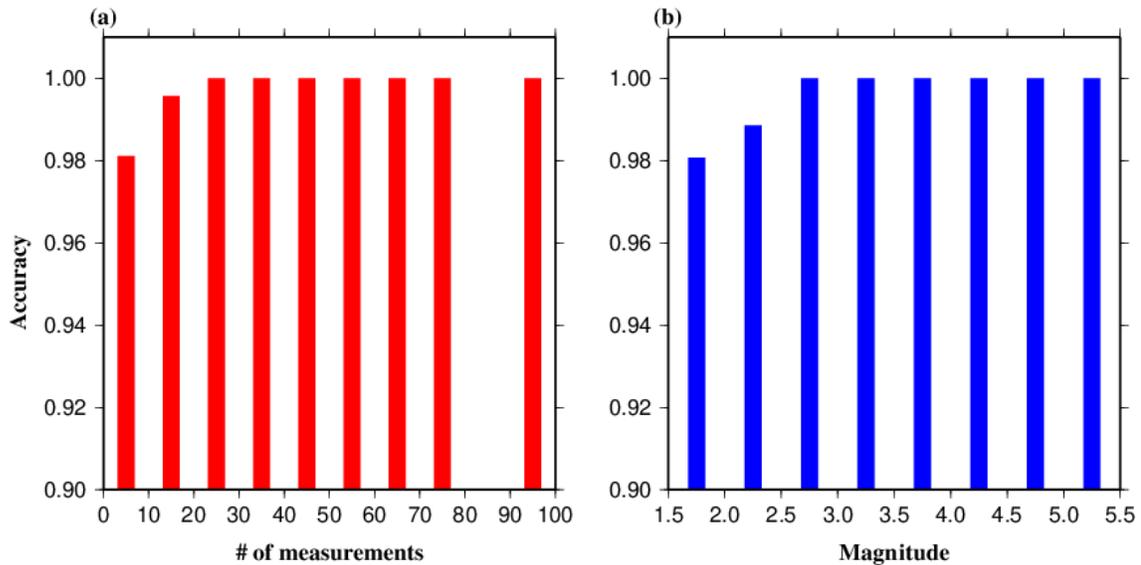


Figure 4. (a) the accuracy of CNN with corrected dataset various with number of measurements of each event. (b) the accuracy of CNN with corrected dataset various with magnitude.

It is also consistent with the common sense that larger magnitude would generate more clear features on the seismograms and a large number of measurements would reduce the influence of bias.

In the ML research, the CNN is known as possessing a poor interpretability and is hard to calibrate (Guo et al., 2017). The dataset, however, is easier to be evaluated and corrected. Data-centric AI is proposed by Ng (2021) and suggests that instead of calibrating models, AI researchers should pay more attention to make better datasets. The results of CNN in this study also reveal that the performance of CNN is highly correlated with the quality and distribution of the dataset.

5. CONCLUSIONS

With rapidly developing of ML, the CNN, with the out-standing ability of feature extraction, has shown high potential in classification problem based on seismic waveforms. In this study, we built an 11-layers CNN to classify natural earthquakes, mine collapses and explosions in Shandong Province, China. A ten-fold cross validation is applied to scan all the dataset, and the results reveal that the CNN has higher performance than human. Moreover, all the measurements with different labels between the CNN and human are assessed and kept, corrected, or abandoned in the dataset. The accuracy of the CNN has obviously improved with the corrected dataset, and all of them are above 95%. The performance of CNN is also influenced by the magnitudes, and the number of measurements of each event received by stations. This study suggests that the

performance of CNN is highly correlated with the quality and distribution of the dataset used to train and validate the neural network.

ACKNOWLEDGEMENTS

The study was partially supported by the U.S. National Science Foundation under awards 1830644 and 1919789 to S.G. The dataset is from the Shangdong Seismic Network Center, China.

REFERENCES

- Andrew Ng, Dillon Laird, and Lynn He. 2021. Datacentric ai competition. <https://https-deeplearning-ai.github.io/data-centric-comp/>.
- Astiz, L., Eakins, J. A., Martynov, V. G., Cox, T. A., Tytell, J., Reyes, J. C., ... & Vernon, F. L. (2014). The Array Network Facility seismic bulletin: Products and an unbiased view of United States seismicity. *Seismological Research Letters*, 85(3), 576-593. <https://doi.org/10.1785/0220130141>
- Baumgardt, D. R., & Young, G. B. (1990). Regional seismic waveform discriminants and case-based event identification using regional arrays. *Bulletin of the Seismological Society of America*, 80(6B), 1874-1892. <https://doi.org/10.1785/BSSA08006B1874>
- Cao, A., & Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophysical research letters*, 29(9), 48-1. <https://doi.org/10.1029/2001GL013775>
- Fäh, D., & Koch, K. (2002). Discrimination between earthquakes and chemical explosions by multivariate statistical analysis: A case study for Switzerland. *Bulletin of the Seismological Society of America*, 92(5), 1795-1805. <https://doi.org/10.1785/0120010166>

- Garcia, J. A., Waszek, L., Tauzin, B., & Schmerr, N. (2021). Automatic identification of mantle seismic phases using a Convolutional Neural Network. *Geophysical Research Letters*, 48(18), e2020GL091658. <https://doi.org/10.1029/2020GL091658>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Machine learning basics. *Deep learning*, 1(7), 98–164.
- Gulli, A., & Pal, S. (2017). *Deep learning with Keras*. Packt Publishing Ltd.
- Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017, July). On calibration of modern neural networks. In *International conference on machine learning* (pp. 1321-1330). PMLR.
- Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis*, 6(5), 429–449. <https://doi.org/10.3233/IDA-2002-6504>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. <https://doi.org/10.48550/arXiv.1412.6980>
- Linville, L., Pankow, K., & Draelos, T. (2019). Deep learning models augment analyst decisions for event discrimination. *Geophysical Research Letters*, 46(7), 3643–3651. <https://doi.org/10.1029/2018GL081119>
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml* (Vol. 30, No. 1, p. 3).
- O'Rourke, C. T., & Baker, G. E. (2017). A spectrogram-based method of Rg detection for explosion monitoring. *Bulletin of the Seismological Society of America*, 107(1), 34-42. <https://doi.org/10.1785/0120160184>
- Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2), e1700578. <https://doi.org/10.1126/sciadv.1700578>
- Zhang, Y., & Gao, S. S. Classification of Teleseismic Shear Wave Splitting Measurements: A Convolutional Neural Network Approach. *Geophysical Research Letters*, e2021GL097101. <https://doi.org/10.1029/2021GL097101>

Zhu, W., & Beroza, G. C. (2018). 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>

SECTION

2. CONCLUSION

In this study, we designed two CNNs for ranking teleseismic SWS measurements and classifying natural earthquakes, mine collapses, and explosions.

For ranking teleseismic SWS measurements, the CNN is trained by published human-labeled datasets and tested using synthetic SWS measurements to evaluate its performance against different levels of noise and its dependence on the difference between the fast orientation and the back-azimuth of the events. When the SNR is greater than 6.5, more than 97% of the non-null synthetic measurements can be correctly accepted by the CNN. Application of the CNN to data from south central Alaska shows that it can accept almost all human-accepted measurements (98.1%) as acceptable when a threshold probability of 0.5 is used. The study suggests a high potential for CNN-based methods to significantly improve the efficiency of measuring shear wave splitting parameters.

For classifying natural earthquakes, mine collapses and explosions, a ten-fold cross validation is applied to scan the entire dataset, and all the measurements with different labels between the CNN and human are assessed and kept, corrected, or abandoned in dataset. The accuracy of CNN has obviously improved with the corrected dataset, and all of them are above 95%. The performance of CNN is also influenced by the magnitudes, and the number of measurements of each event received by stations. This study suggests that the performance of CNN is highly correlated with the quality and distribution of the dataset used to train and validate the neural network.

BIBLIOGRAPHY

- Alsina, D., & Snieder, R. (1995). Small-scale sublithospheric continental mantle deformation: Constraints from SKS splitting observations. *Geophysical Journal International*, 123(2), 431–448. <https://doi.org/10.1111/j.1365-246X.1995.tb06864.x>
- Ando, M., Ishikawa, Y., & Yamazaki, F. (1983). Shear wave polarization anisotropy in the upper mantle beneath Honshu, Japan. *Journal of Geophysical Research*, 88(B7), 5850–5864. <https://doi.org/10.1029/JB088iB07p05850>
- Andrew Ng, Dillon Laird, and Lynn He. 2021. Datacentric ai competition. <https://https-deeplearning-ai.github.io/data-centric-comp/>.
- Astiz, L., Eakins, J. A., Martynov, V. G., Cox, T. A., Tyttell, J., Reyes, J. C., ... & Vernon, F. L. (2014). The Array Network Facility seismic bulletin: Products and an unbiased view of United States seismicity. *Seismological Research Letters*, 85(3), 576-593. <https://doi.org/10.1785/0220130141>
- Baumgardt, D. R., & Young, G. B. (1990). Regional seismic waveform discriminants and case-based event identification using regional arrays. *Bulletin of the Seismological Society of America*, 80(6B), 1874-1892. <https://doi.org/10.1785/BSSA08006B1874>
- Bianco, M. J., Gerstoft, P., Olsen, K. B., & Lin, F. C. (2019). High-resolution seismic tomography of Long Beach, CA using machine learning. *Scientific Reports*, 9(1), 1–11. <https://doi.org/10.1038/s41598-019-50381-z>
- Cao, A., & Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophysical research letters*, 29(9), 48-1. <https://doi.org/10.1029/2001GL013775>
- Dokht, R. M., Kao, H., Visser, R., & Smith, B. (2019). Seismic event and phase detection using time–frequency representation and convolutional neural networks. *Seismological Research Letters*, 90(2A), 481–490. <https://doi.org/10.1785/0220180308>
- Fäh, D., & Koch, K. (2002). Discrimination between earthquakes and chemical explosions by multivariate statistical analysis: A case study for Switzerland. *Bulletin of the Seismological Society of America*, 92(5), 1795-1805. <https://doi.org/10.1785/0120010166>

- Garcia, J. A., Waszek, L., Tauzin, B., & Schmerr, N. (2021). Automatic identification of mantle seismic phases using a Convolutional Neural Network. *Geophysical Research Letters*, 48(18), e2020GL091658. <https://doi.org/10.1029/2020GL091658>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Machine learning basics. *Deep learning*, 1(7), 98–164.
- Gulli, A., & Pal, S. (2017). *Deep learning with Keras*. Packt Publishing Ltd.
- Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017, July). On calibration of modern neural networks. In *International conference on machine learning* (pp. 1321–1330). PMLR.
- Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis*, 6(5), 429–449. <https://doi.org/10.3233/IDA-2002-6504>
- Jia, Y., Liu, K. H., Kong, F., Liu, L., & Gao, S. S. (2021). A systematic investigation of piercing-point-dependent seismic azimuthal anisotropy. *Geophysical Journal International*, 227(3), 1496–1511. <https://doi.org/10.1093/gji/ggab285>
- Katayama, I., & Karato, S. I. (2006). Effect of temperature on the B-to C-type olivine fabric transition and implication for flow pattern in subduction zones. *Physics of the Earth and Planetary Interiors*, 157(1–2), 33–45. <https://doi.org/10.1016/j.pepi.2006.03.005>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>
- Kong, F., Gao, S. S., & Liu, K. H. (2015). A systematic comparison of the transverse energy minimization and splitting intensity techniques for measuring shear-wave splitting parameters. *Bulletin of the Seismological Society of America*, 105(1), 230–239. <https://doi.org/10.1785/0120140108>
- Li, Z., Meier, M. A., Hauksson, E., Zhan, Z., & Andrews, J. (2018). Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophysical Research Letters*, 45(10), 4773–4779. <https://doi.org/10.1029/2018GL077870>
- Link, F., Reiss, M. C., & Rumpker, G. (2022). An automatized XKS-splitting procedure for large data sets: Extension package for SplitRacer and application to the USArray. *Computers & Geosciences*, 158, 104961. <https://doi.org/10.1016/j.cageo.2021.104961>

- Linville, L., Pankow, K., & Draelos, T. (2019). Deep learning models augment analyst decisions for event discrimination. *Geophysical Research Letters*, 46(7), 3643–3651. <https://doi.org/10.1029/2018GL081119>
- Liu, K. H., Elsheikh, A., Lemnifi, A., Purevsuren, U., Ray, M., Refayee, H., et al. (2014). A uniform database of teleseismic shear wave splitting measurements for the western and central United States. *Geochemistry, Geophysics, Geosystems*, 15(5), 2075–2085. <https://doi.org/10.1002/2014GC005267>
- Liu, K. H., & Gao, S. S. (2013). Making reliable shear-wave splitting measurements. *Bulletin of the Seismological Society of America*, 103(5), 2680–2693. <https://doi.org/10.1785/0120120355>
- Liu, K. H., Gao, S. S., Gao, Y., & Wu, J. (2008). Shear wave splitting and mantle flow associated with the deflected Pacific slab beneath northeast Asia. *Journal of Geophysical Research*, 113, B01305. <https://doi.org/10.1029/2007JB005178>
- Lomax, A., Michelini, A., & Jozinović, D. (2019). An investigation of rapid earthquake characterization using single-station waveforms and a convolutional neural network. *Seismological Research Letters*, 90(2A), 517–529. <https://doi.org/10.1785/0220180311>
- Long, M. D., & Silver, P. G. (2009). Shear wave splitting and mantle anisotropy: Measurements, interpretations, and new directions. *Surveys in Geophysics*, 30(4), 407–461. <https://doi.org/10.1007/s10712-009-9075-1>
- Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml* (Vol. 30, No. 1, p. 3).
- McBrearty, I. W., Delorey, A. A., & Johnson, P. A. (2019). Pairwise association of seismic arrivals with convolutional neural networks. *Seismological Research Letters*, 90(2A), 503–509. <https://doi.org/10.1785/0220180326>
- Mignan, A., & Broccardo, M. (2019). One neuron versus deep learning in aftershock prediction. *Nature*, 574(7776), E1–E3. <https://doi.org/10.1038/s41586-018-0438-y>
- Mousavi, S. M., & Beroza, G. C. (2020). A machine-learning approach for earthquake magnitude estimation. *Geophysical Research Letters*, 47(1), e2019GL085976. <https://doi.org/10.1029/2019GL085976>
- Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th international conference on international conference on machine learning*, 807–814.

- O'Rourke, C. T., & Baker, G. E. (2017). A spectrogram-based method of Rg detection for explosion monitoring. *Bulletin of the Seismological Society of America*, 107(1), 34-42. <https://doi.org/10.1785/0120160184>
- Peng, Z., & Ben-Zion, Y. (2004). Systematic analysis of crustal anisotropy along the Karadere–Duzce branch of the North Anatolian fault. *Geophysical Journal International*, 159, 253–274. <https://doi.org/10.1111/j.1365-246X.2004.02379.x>
- Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances*, 4(2), e1700578. <https://doi.org/10.1126/sciadv.1700578>
- Ross, Z. E., Meier, M. A., & Hauksson, E. (2018). P wave arrival picking and first-motion polarity determination with deep learning. *Journal of Geophysical Research: Solid Earth*, 123, 5120–5129. <https://doi.org/10.1029/2017JB015251>
- Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes. *Geophysical Research Letters*, 44(18), 9276–9282. <https://doi.org/10.1002/2017GL074677>
- Rümpker, G., & Silver, P. G. (1998). Apparent shear-wave splitting parameters in the presence of vertically varying anisotropy. *Geophysical Journal International*, 135(3), 790–800. <https://doi.org/10.1046/j.1365-246X.1998.00660.x>
- Savage, M. K. (1999). Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting? *Reviews of Geophysics*, 37(1), 65–106. <https://doi.org/10.1029/98RG02075>
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484. <https://doi.org/10.1038/nature16961>
- Silver, P. G. (1996). Seismic anisotropy beneath the continents: Probing the depths of geology. *Annual Review of Earth and Planetary Sciences*, 24(1), 385–432. <https://doi.org/10.1038/nature16961>
- Silver, P. G., & Chan, W. W. (1991). Shear wave splitting and subcontinental mantle deformation. *Journal of Geophysical Research*, 96(B10), 16429–16454. <https://doi.org/10.1029/91JB00899>
- Silver, P. G., & Savage, M. K. (1994). The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers. *Geophysical Journal International*, 119(3), 949–963. <https://doi.org/10.1111/j.1365-246X.1994.tb04027.x>

- Teanby, N. A., Kendall, J. M., & van der Baan, M. (2004). Automation of shear-wave splitting measurements using cluster analysis. *Bulletin of the Seismological Society of America*, 94(2), 453–463. <https://doi.org/10.1785/0120030123>
- Titos, M., Bueno, A., Garcia, L., & Benitez, C. (2018). A deep neural networks approach to automatic recognition systems for volcano-seismic events. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5), 1533–1544. <https://doi.org/10.1109/JSTARS.2018.2803198>
- Vecsey, L., Plomerová, J., & Babuška, V. (2008). Shear-wave splitting measurements—Problems and solutions. *Tectonophysics*, 462(1–4), 178–196. <https://doi.org/10.1016/j.tecto.2008.01.021>
- Woollam, J., Rietbrock, A., Bueno, A., & De Angelis, S. (2019). Convolutional neural network for seismic phase classification, performance demonstration over a local seismic network. *Seismological Research Letters*, 90(2A), 491–502. <https://doi.org/10.1785/0220180312>
- Yang, B. B., Liu, K. H., Dahm, H. H., & Gao, S. S. (2016). A uniform database of teleseismic shear-wave splitting measurements for the western and central United States: December 2014 update. *Seismological Research Letters*, 87(2A), 295–300. <https://doi.org/10.1785/0220150213>
- Yang, B. B., Liu, Y., Dahm, H., Liu, K. H., & Gao, S. S. (2017). Seismic azimuthal anisotropy beneath the eastern United States and its geodynamic implications. *Geophysical Research Letters*, 44(6), 2670–2678. <https://doi.org/10.1002/2016GL071227>
- Yang, Y., Gao, S. S., Liu, K. H., Kong, F., & Fu, X. (2021). Mantle flow in the vicinity of the eastern edge of the Pacific-Yakutat slab: Constraints from shear wave splitting analyses. *Journal of Geophysical Research: Solid Earth*, 126, e2021JB022354. <https://doi.org/10.1029/2021JB022354>
- Zhang, S., & Karato, S. I. (1995). Lattice preferred orientation of olivine in simple shear deformation and the flow geometry of the upper mantle of the Earth. *Nature*, 375, 774–777. <https://doi.org/10.1038/375774a0>
- Zhang, Y., & Gao, S. S. Classification of Teleseismic Shear Wave Splitting Measurements: A Convolutional Neural Network Approach. *Geophysical Research Letters*, e2021GL097101. <https://doi.org/10.1029/2021GL097101>
- Zhu, W., & Beroza, G. C. (2018). 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>

Zhu, W., Mousavi, S. M., & Beroza, G. C. (2019). Seismic signal denoising and decomposition using deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(11), 9476–9488.
<https://doi.org/10.1109/TGRS.2019.2926772>

VITA

Yanwei Zhang entered Northeast Petroleum University, Daqing, China in 2013 and studied Exploration Technology and Engineering. After two college years, he transferred to Missouri University of Science and Technology and focused on Geology and Geophysics. He received his bachelor's degree in December 2017 from both universities. He continued studying Geology and Geophysics in Missouri University of Science and Technology and joined Dr. Stephen S. Gao's research group for Ph.D. degree in January 2018.

He focused on application of Machine Learning in Geophysics during his Ph.D. career and published one paper on Geophysical Research Letter as the first author. He also presented his works at American Geophysical Union in 2019, 2020, 2021, and 2022. In December 2022, he received his Doctor of Philosophy in Geology and Geophysics from Missouri University of Science and Technology.