

Detecting and Locating Aftershocks for the 2020 M_w 6.5 Stanley, Idaho, Earthquake Using Convolutional Neural Networks

Bingxu Luo^{*1}, Hejun Zhu^{1,2}, Jidong Yang³, Thorne Lay⁴, Lingling Ye⁵, Zhong Lu⁶, and David Lumley^{1,2}

Abstract

Our study is to build an aftershock catalog with a low magnitude of completeness for the 2020 M_w 6.5 Stanley, Idaho, earthquake. This is challenging because of the low signal-to-noise ratios for recorded seismograms. Therefore, we apply convolutional neural networks (CNNs) and use 2D time–frequency feature maps as inputs for aftershock detection. Another trained CNN is used to automatically pick P -wave arrival times, which are then used in both nonlinear and double-difference earthquake location algorithms. Our new one-month-long catalog has 4644 events and a completeness magnitude (M_c) 1.9, which has over seven times more events and 0.9 lower M_c than the current U.S. Geological Survey National Earthquake Information Center catalog. The distribution and expansion of these aftershocks improve the resolution of two north-northwest-trending faults with different dip angles, providing further support for a central stepover region that changed the earthquake rupture trajectory and induced sustained seismicity.

Cite this article as Luo, B., H. Zhu, J. Yang, T. Lay, L. Ye, Z. Lu, and D. Lumley (2022). Detecting and Locating Aftershocks for the 2020 M_w 6.5 Stanley, Idaho, Earthquake Using Convolutional Neural Networks, *Seismol. Res. Lett.* **XX**, 1–12, doi: 10.1785/0220210341.

[Supplemental Material](#)

Introduction

The 31 March 2020 M_w 6.5 Stanley, Idaho, earthquake occurred ~115 km northeast of Boise and 30 km northwest of Stanley, Custer County. The shaking lasted ~10 s (Idaho Geological Survey). However, no injuries and only minor damage were reported because of the remote location of the earthquake. More details about the event are available in the [Data and Resources](#) (Idaho Geological Survey).

The 2020 Stanley earthquake was close to the northern extension of the late Cenozoic volcanic Basin and Range Province (BRP; Gans and Bohrsen, 1998), as well as the north side of the Snake River plain (SRP) that covers most of the major cities of Idaho (see the inset panel of Fig. 1a). In central Idaho, the SRP also intersects with the Intermountain Seismic Belt, which is an approximate north–south-trending seismic zone that extends from northwest Montana over 1300 km long and terminates in southeastern Nevada (Smith and Sbar, 1974; Richins *et al.*, 1987). Since 1960, there have been >70 earthquakes with magnitudes >4.0 within central Idaho (Fig. 1a). The largest one, the 1983 M_w 6.9 Borah Peak earthquake, occurred along the Lost River fault, which is a southwest-dipping normal fault (U.S. Geological Survey). The 2020 Stanley earthquake is the largest earthquake that has occurred in central Idaho since the 1983 M_w 6.9 Borah Peak earthquake located ~100 km to the east.

The 2020 Stanley earthquake sequence extends across the northwestern terminus of the Sawtooth fault (Fig. 1a), which is a north-northwest–south-southeast-trending northeast-dipping normal fault with a length of 55–65 km (Thackray *et al.*, 2013). To date, no strong evidence indicates that this earthquake sequence is associated with the Sawtooth or any other mapped faults within the study region. Yang *et al.* (2021) reconciled observations from backprojection (BP) imaging, teleseismic finite-fault inversion, and near-field Interferometric Synthetic Aperture Radar (InSAR) analysis and then proposed that the 2020 Stanley earthquake ruptured a pair of opposing-dip faults offset by a 10-km-wide step at the surface. They argued that this event initiated north of the Sawtooth fault and ruptured

1. Department of Geosciences, The University of Texas at Dallas, Richardson, Texas, U.S.A., <https://orcid.org/0000-0001-7075-1529> (BL); <https://orcid.org/0000-0002-7452-075X> (HZ); <https://orcid.org/0000-0003-4342-4551> (DL);
2. Department of Physics, The University of Texas at Dallas, Richardson, Texas, U.S.A.;
3. Department of Geophysics, China University of Petroleum (East China), Qingdao, China, <https://orcid.org/0000-0002-6652-5366> (JY);
4. Department of Earth and Planetary Sciences, University of California, Santa Cruz, California, U.S.A., <https://orcid.org/0000-0003-2360-4213> (TL);
5. Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen, China., <https://orcid.org/0000-0001-9689-4149> (LY);
6. Roy M. Huffington Department of Earth Sciences, Southern Methodist University, Dallas, Texas, U.S.A., <https://orcid.org/0000-0001-9181-1818> (ZL)

*Corresponding author: bingxu.luo@utdallas.edu

© Seismological Society of America

southward, then stepped over to a southern subfault semiparallel to the Sawtooth fault. They used this model to explain the 44% non-double-couple components in the W -phase moment tensor solution, which is estimated by the [U.S. Geological Survey National Earthquake Information Center \(USGS-NEIC\)](#).

In this study, we further analyze the distribution and migration of aftershocks of this event to better characterize the sequence occurrence and evolution. An aftershock catalog with a low magnitude of completeness is important for studying the ruptures of earthquakes and postseismic deformation within rupture zones ([Tajima and Kanamori, 1985](#); [Enescu et al., 2007](#); [Peng et al., 2007](#)). The USGS-NEIC ([U.S. Geological Survey, National Earthquake Information Center](#)) one-month-long aftershock catalog lists 662 events with magnitudes ranging from 1.61 to 4.8 and has a completeness magnitude of 2.8 (see Text S5 and Fig. S12, available in the supplemental material to this article). The current USGS-NEIC catalog provides limited information about this complex rupture zone. Therefore, in this study, we build a more complete aftershock catalog with accurate location information to better study the complex rupture process of the 2020 Stanley earthquake.

The procedure of locating aftershocks from continuous seismic records consists of signal or event detection, P - and S -phase arrival-time picking, association of picks, and event location. For event detection, many effective approaches have been proposed. One way is to use the short-term average/long-term average (STA/LTA) ([Allen, 1978](#); [Withers et al., 1998](#); [Saragiotis et al., 2002](#); [Nippres et al., 2010](#)), which is a fast approach for characterizing signal attributes from original waveforms, or one can use higher order statistical characteristic functions for both short and long windows, but this method fails to detect events in the presence of strong background noise. Other approaches include waveform autocorrelation ([Brown et al., 2008](#); [Gibbons and Ringdal, 2010](#)) and template matching ([Shelly et al., 2007](#); [Harris and Dodge, 2011](#); [Barrett and Beroza, 2014](#); [Benz et al., 2015](#); [Ross et al., 2019](#); [Zhou et al., 2021](#)), which use catalog events as templates to cross correlate with continuous waveform records. These cross-correlation-based methods require enough representative waveforms to thoroughly sample the overall event population and are computationally expensive. Recently, machine learning algorithms have been utilized to detect signals from seismic events ([Kortström et al., 2015](#); [Li et al., 2018](#); [Perol et al., 2018](#); [Mousavi et al., 2019](#); [Tan et al., 2021](#)). For instance, [Perol et al. \(2018\)](#) demonstrated that convolutional neural networks (CNNs) have higher detection accuracy and lower computation costs than previous methods. In this study, we use a novel CNN with 2D time–frequency feature maps as inputs, which have better performance than conventional 1D waveform inputs ([Mousavi et al., 2016](#); [Dokht et al., 2019](#); [Mousavi et al., 2019](#)), to automatically detect signals from aftershocks for the 2020 M_w 6.5 Stanley earthquake. Then we apply another trained CNN ([Ross et al., 2018](#)) to automatically pick P -wave arrival times for the detected events. We then combine a global

optimization earthquake location algorithm with a double-difference (DD) location method to improve the location of aftershocks for the 2020 Stanley earthquake.

Data and Methods

Data selection and processing

From the latest USGS-NEIC catalog, we select 1568 events that occurred within four months after the mainshock as the training dataset. Their magnitudes range from 1.1 to 4.8. Next, we choose to set the length of the detection window as 20 s, seeking no more than one event in each interval. We extract vertical-component seismograms for these catalog events for 20 s long windows from four stations (GS.ID11, IW.PLID, IE.DVCI, and US.HLID) that surround the mainshock as denoted by the blue triangles in Figure 1a. We create the event windows from the event times for station GS.ID11 and 10 s after the event times for stations IW.PLID, IE.DVCI, and US.HLID. In addition to the event dataset, we also build a noise dataset that can be distinguished from events. Here, we use the USGS Comprehensive Earthquake Catalog (ComCat) event catalog ([Benz et al., 2015](#)) and remove catalog events to build noise datasets. Finally, our training datasets include two classes (events and noise) with different labels. Thus, the CNN can convert the earthquake detection problem into a classification problem.

Raw waveform preprocessing includes detrending, demeaning, resampling data to 100 samples per second, band-pass filtering from 1 to 20 Hz, and normalization. We apply short-time Fourier transform to convert all training waveforms into 2D time–frequency feature maps to improve the training efficiency of the CNN (Fig. S1).

CNN architecture for earthquake detection

Figure 2 shows the CNN architecture used in this study. The input is a 2D time–frequency feature map with size of 2000 in the time axis and 129 in the frequency axis. Between every two layers, we use a 4D kernel w to perform convolution along both time and frequency axes and then add a bias term b :

$$y_c^l = \sigma \left(\sum_{c'=1}^{C^l} \sum_{t=1}^{32} \sum_{f=1}^3 y_{c',t,f}^{l-1} \cdot w_{c,c',t,f}^l + b_c^l \right), \quad (1)$$

in which y_c^l denotes the output feature map in the convolution layer l , for which l ranges from 0 to 8. c' and c are the indexes for input and output channels, respectively. For the 4D convolution kernel $w_{c,c',t,f}^l$, the first two indexes denote input (c') and output (c) channel numbers, and the last two indexes represent time (t) and frequency (f). C^l is the number of input channels in layer l , which is set as 32 for layer 1–8. Thus, w^1 has the size of $1 \times 32 \times 3 \times 3$ because there is only one channel for the initial input, whereas w^2, \dots, w^8 have the size of $32 \times 32 \times 3 \times 3$. b_c^l is the bias term that has the same size as the output feature map y_c^l . $\sigma(\cdot)$ is the ReLU (Rectifier Linear Unit) activation function,

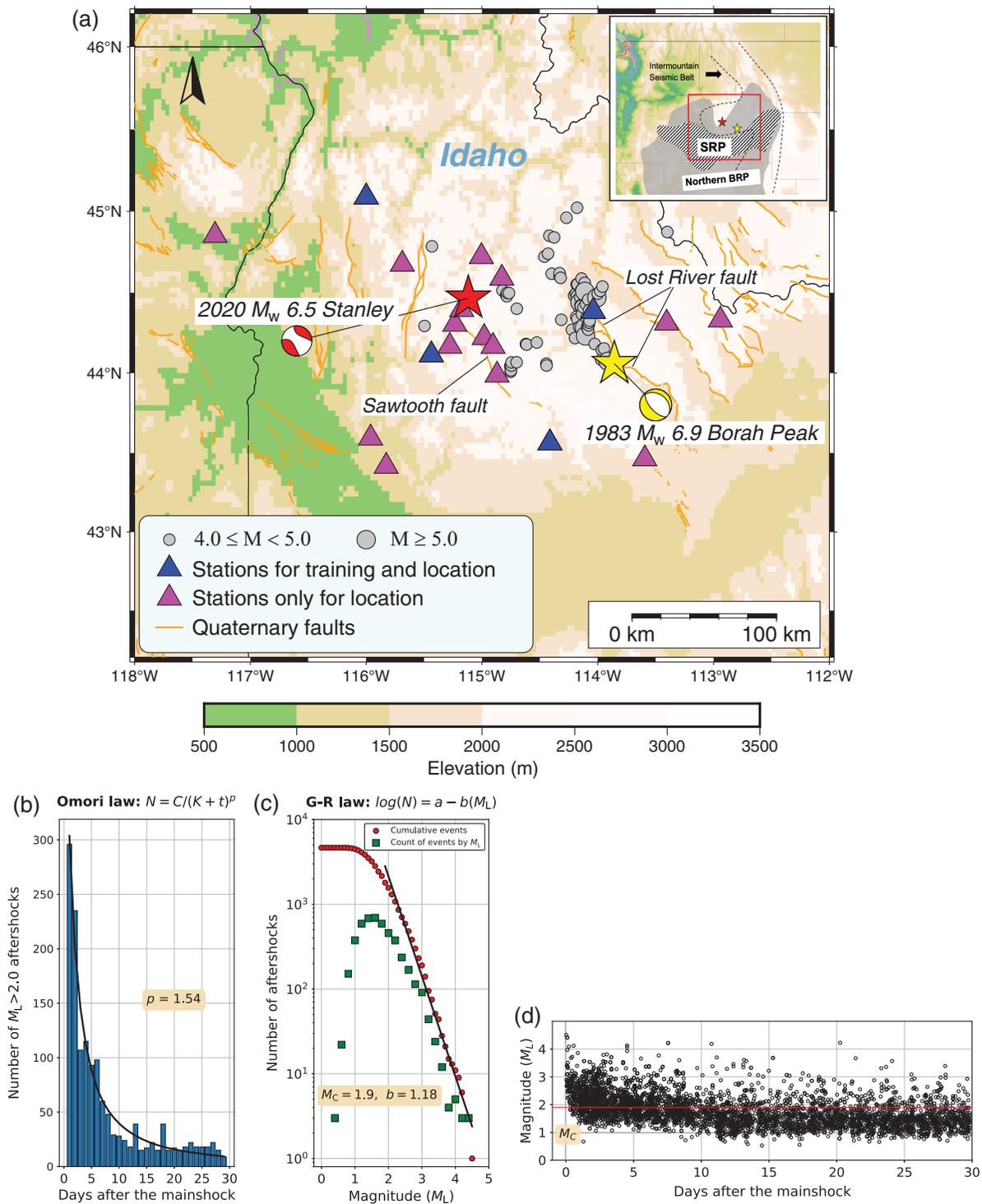
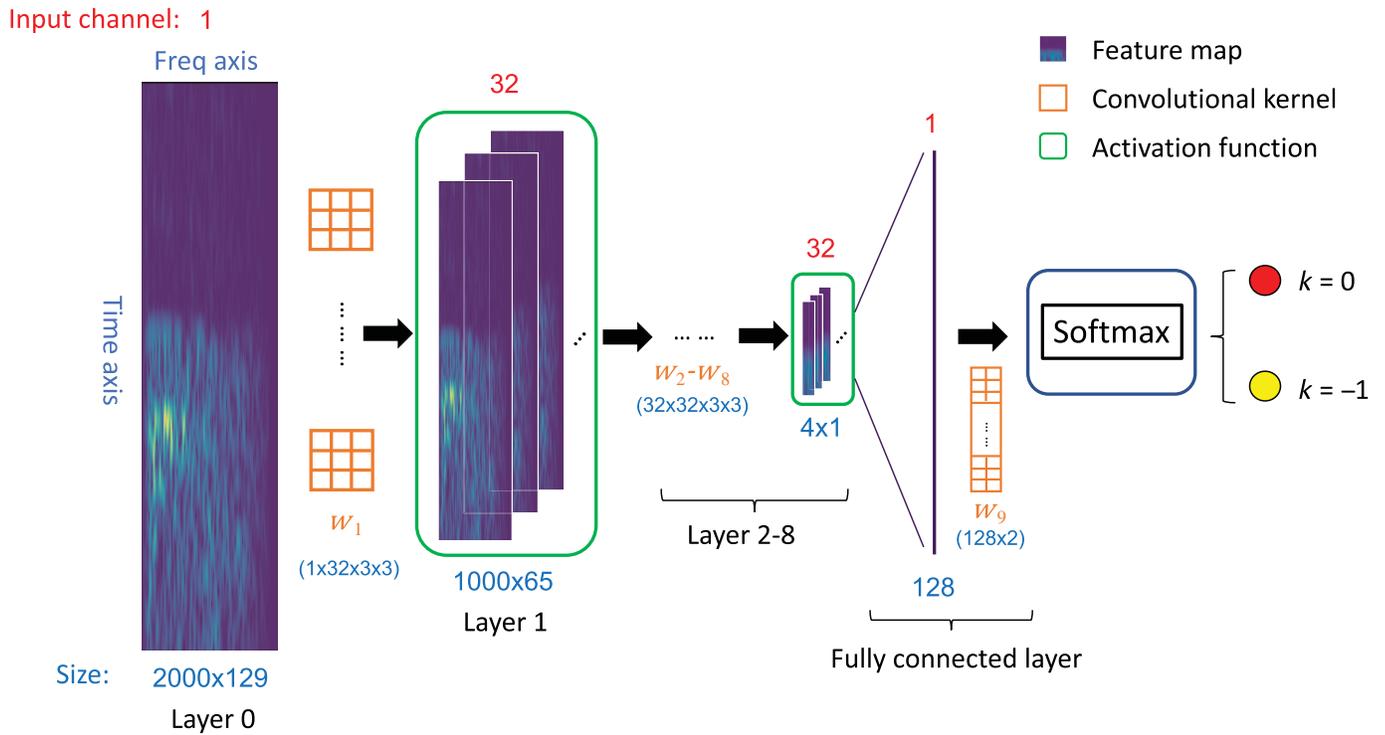


Figure 1. The seismotectonic setting of the study region and characteristics of newly detected aftershocks. (a) Seismicity in Central Idaho since 1960 is denoted by gray circles. The two largest events (the 2020 M_w 6.5 Stanley earthquake and the 1983 M_w 6.9 Borah Peak earthquake) are shown as red and yellow stars along with their U.S. Geological Survey National Earthquake Information Center (USGS-NEIC) W -phase moment tensor solutions, respectively. All triangles denote seismic stations used in this study (more details about these stations are listed in Table S1). The dominant tectonic features, Snake River plain

(SRP), Basin and Range province (BRP), and Intermountain Seismic Belt are labeled in the top right inset map. (b) The decaying curve of detected aftershocks with estimated $M_L > 2.0$, which is fit by the Omori law (Omori, 1894; Utsu, 1961). (c) The relation between the estimated local magnitudes (M_L) and aftershock frequency for our new catalog. The black line is the fit by the Gutenberg–Richter law (Gutenberg and Richter, 1944). We estimate the completeness magnitude (M_c) as 1.9. Panel (d) shows the estimated magnitudes (M_L) over time for our new catalog, the red line denotes M_c that we estimate in (c).



which is used to avoid vanishing gradients and increase the non-linearity of the networks (Hochreiter, 1991). We perform each convolution computation by a stride of 2 (with zero padding); therefore, each output channel is half the size of the input channel in both time and frequency dimensions.

After eight convolution layers, the initial input has been downsampled into 4×1 with 32 channels. We have two different classes of events ($k = 0$) and noise ($k = -1$). Then we use a fully connected layer to flatten the final output as a 1D vector \bar{y}^9 that includes 128 samples and add an additional convolution kernel w^9 with the size of 128×2 and a bias term b^9 to compute the evaluation score for the specific class k :

$$y_k = \sum_{c'=1}^{128} \sum_{k=-1}^0 \bar{y}_{c'}^9 \cdot w_{c'k}^9 + b_k^9. \quad (2)$$

This score y_k keeps information of the initial input feature map learned by the CNN model. Finally, we apply the following Softmax function to build a probability distribution $P(y_k)$ for these two classes:

$$P(y_k) = \frac{\exp(y_k)}{\sum_{k=-1}^0 \exp(y_k)}. \quad (3)$$

The logistic function can project any input data into 0 to 1. The Softmax function is a generalized probability distribution function for multiple classes, which has been widely used for classification problems in machine learning.

During the training process, we input a batch size 64 of training data into the CNN and compute the loss value ψ for each

Figure 2. The convolutional neural network (CNN) architecture for earthquake detection. There are eight convolution layers, and each layer includes trainable kernels and biases (equation 1). The final fully connected layer is used to reshape the feature map into a 1D vector (equation 2). The Softmax function produces a probability distribution function for two classes as events ($k = 0$) and noise ($k = -1$) (equation 3).

iteration. We use the following cross-entropy loss function in this study:

$$\psi = -\frac{1}{N} \sum_n \sum_{k=-1}^0 q_k^n \log(p_k^n) + \lambda \sum_{l=1}^{10} \|w_l\|^2, \quad (4)$$

in which N is the number of training samples indexed by n , q^k , and p^k are the true and predicted probabilities for class k . For all classes, $q_k^n = 1$ when k equals the true label or $q_k^n = 0$ when k is any other false label. For instance, if the n th sample's true label is -1 , only the predicted probability $p_{k=-1}^n$ with the true label ($k = -1$) can be kept because of $q_{k=-1}^n = 1$, another $p_{k=0}^n$ with the false label ($k = 0$) will be discarded. Furthermore, we add an L_2 regularization term into the loss function to mitigate the overfitting problem. λ is a constant used to balance the regularization term with the cross-entropy function, which is set as 10^{-3} in this study.

For the optimization method, we use the Adam algorithm (Kingma and Ba, 2017) to update all trainable parameters (w and b) and give a learning rate of 10^{-4} . The learning rate is an empirical parameter that determines the step size in each iteration that moves the loss function toward the minimum

level. The training dataset contains 4121 positive samples and 261,760 negative samples. It takes >28,000 iterations to finish the training process (Fig. S2). To validate the performance of the trained CNN, we use validation datasets that include 527 positive samples and 16,512 negative samples, which achieve the precision and recall of 99.6% and 91.9% for events and noise, respectively (see Table S2).

P-phase picking

To locate the detected events, we select 19 stations in total and pick *P*-wave arrival times for all available traces, which include nine permanent stations and 10 local stations from the XP network (stations are denoted as triangles in Fig. 1a; more information about stations is listed in Table S1). For each detected event, we use known aftershocks as templates to measure its peak signal-to-noise ratio (PSNR) for all available traces (stations) and remove noisy traces with average PSNR values <15. (See the method and an example in Text S6 and Fig. S4.) We only use vertical-component seismograms for *P*-wave arrival-time picking and apply the same preprocessing procedure for the training dataset. We add additional 40 s following the detected event windows to generate 60 s long picking windows for all traces.

Manually picking *P*-wave arrival times is a very time-consuming task. Here, we utilize another trained CNN model (Ross *et al.*, 2018) to automatically pick *P*-wave arrival times (see the parameters of the CNN model in Table S3). This model has been trained with 273,882 *P*-wave windows picked by experts from the Southern California Earthquake Data Center. Because the trained CNN model requires an input window that consists of 400 samples to control picking accuracy, we use the STA/LTA method to approximately determine the middle point of each input window. Then we extend the signal window as a 4 s long window from the determined middle point. To evaluate the picking quality of the trained CNN model, we consider the one-month-long USGS-NEIC aftershock catalog for the 2020 Stanley mainshock as the ground truth, then apply a 1D averaged velocity model (Shen and Ritzwoller, 2016) to calculate *P*-wave travel times for all 19 stations. To validate the trained CNN model, we use another automatic phase picker tool: PhasePapy, which picks with a dynamic threshold by processing the characteristic functions of waveforms (Chen and Holland, 2016). Then we compute travel-time residuals for these two pickers. Figure S3 shows the trained CNN model allows us to achieve smaller bias and standard deviation compared with results from PhasePapy.

Combining *P* and *S* phases would usually allow us to better locate earthquakes, especially for estimating focal depths and origin times (Gomberg *et al.*, 1990). However, the *S* phases must be correctly identified and measured; poor *S* data will lead to inaccurate solutions even with robust location algorithms (Gomberg *et al.*, 1990). In addition, *S* phases can be better identified and picked from horizontal components and require different processing. In this study, we only use *P* phases because our detection and picking algorithms mainly handle these signals.

Relocation of aftershocks

After we use two well-trained CNNs to perform aftershock detection and *P*-phase picking, we then use a global optimization location program NonLinLoc (Lomax *et al.*, 2000) to initially estimate locations for the detected events (see more details in Text S1). This approach is based on the Bayes' theorem:

$$P(m|d) = \frac{P(m)P(d|m)}{P(d)}, \quad (5)$$

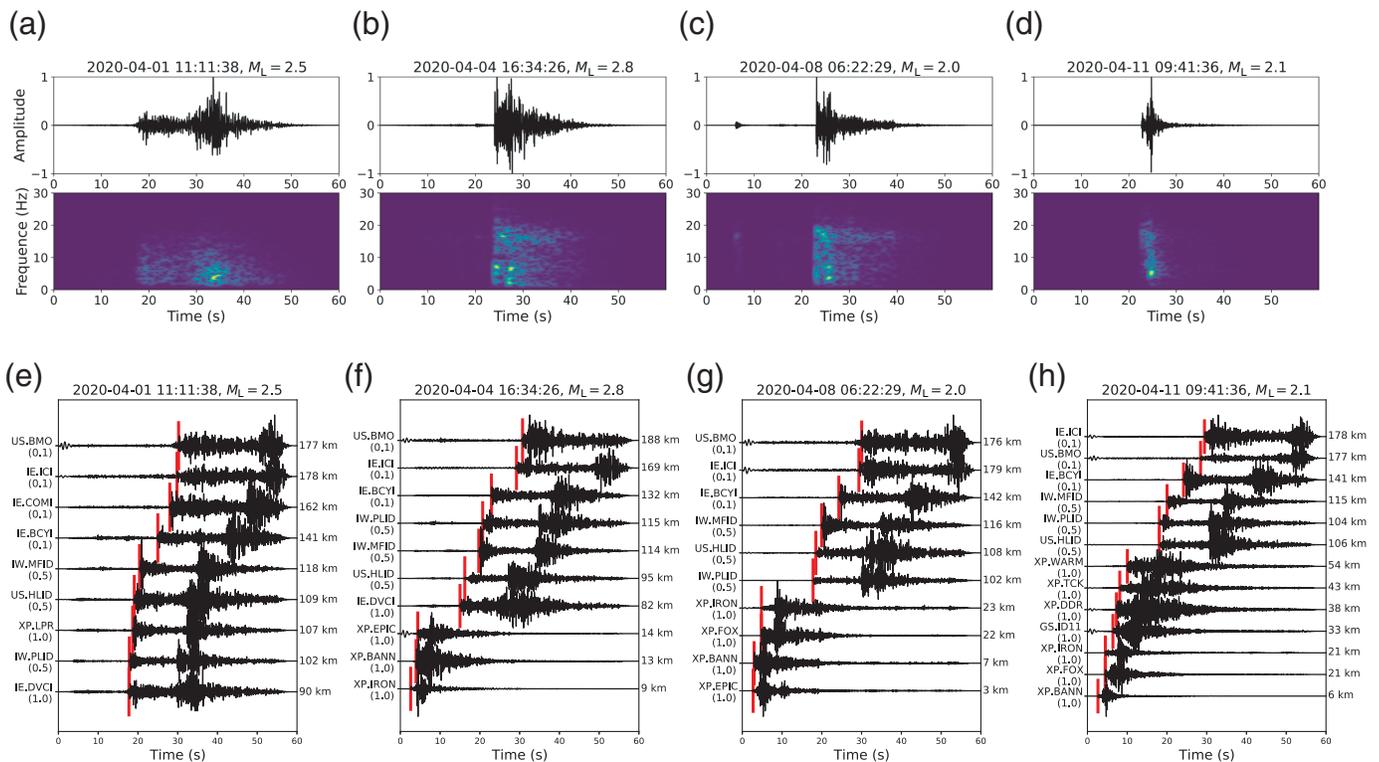
in which $m = (x, y, z, t_0)$ represents the event's hypocenter and starting time, $d = (t_1, t_2, t_3, \dots, t_n)$ are the observed travel times, and n is the total number of measurements. We estimate a 1D averaged velocity profile from a 3D crustal and uppermost mantle velocity model (Shen and Ritzwoller, 2016) to compute travel times between each station and every potential source location (see the velocity model in Table S4). Then we use the Oct-tree searching method to sample the posteriori probability density function (see Text S1). This procedure intrinsically provides association of arrivals from common events and preliminary location estimates.

We then apply a DD location technique (Waldhauser and Ellsworth, 2000) to further improve the resolution of results from the global search (see more details in Text S2). The DD technique takes advantage of two events with relatively small hypocentral separation compared with the event-station distance and the scale of velocity heterogeneity. Thus, the common travel paths for these two events can be eliminated. Then the inversion will be mainly sensitive to the hypocentral separation, which can be adjusted by fitting observed DD travel times, that is, the travel-time difference between two earthquakes for the same station. This technique is very useful for regions with a high-seismicity rate, which helps us to better study complex geological structures and seismogenic zones (Weller *et al.*, 2012).

Results

Location of detected aftershocks

We select one-month-long vertical-component continuous records after the mainshock for station XP.BANN as our testing datasets. Since station XP.BANN was installed at 22:34:38, 1 April 2020 (Coordinated Universal Time, UTC), so we use station IE.DVCI to supply records for the first day after the mainshock (23:52:30, 31 March 2020). Here, we show four examples of detected aftershocks that are not included in the current USGS-NEIC catalog (Fig. 3a–d). We also provide *P*-phase picking results from all available stations for these four examples (Fig. 3e–h). For instance, Figure 3b shows a waveform and its 2D time–frequency feature map for a detected event that occurs at 16:34:26, 4 April 2020, and Figure 3f presents all traces after PSNR filtering for this newly detected event. In addition, red bars denote the *P*-wave arrival times picked by the trained CNN model, and the a priori weight labeled at each trace works as the weighting factor in the



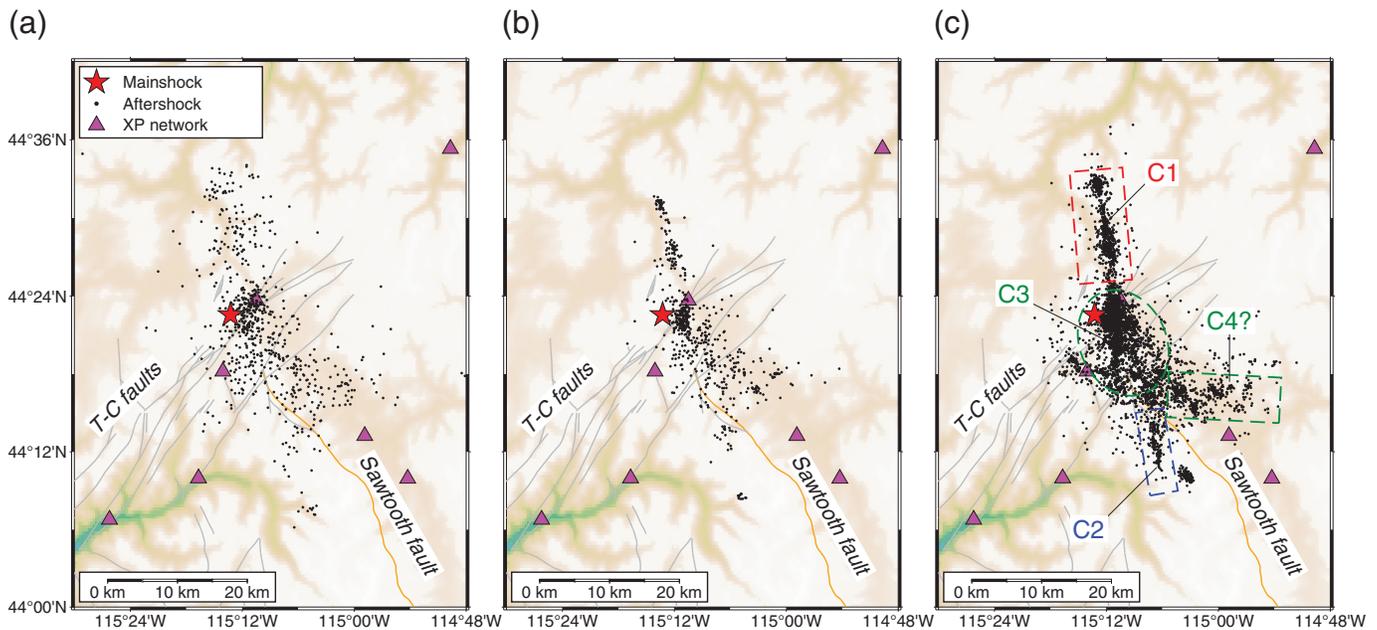
following DD relocation (see details in Text S3). We can observe the moveout trend of detected arrivals for all available stations.

The first location method based on the Bayesian theory gives us associations and initial location estimates for detected events. After the first location, we remove events with semimajor uncertainty axis lengths >10 km and those that are far away from our study region. Then the DD relocation technique is utilized to strengthen the links for pairs of events and remove weak linkages from the initial locations. (See the comparison between the initial locations of detected aftershocks and ones after the DD relocation in Figure S6 and an example of initial location uncertainty in Table S5 and Fig. S7.) In the DD relocation, we assign the a priori weights for all traces to represent the P -wave picking quality (Fig. 3e–h). The a priori weights are set according to the epicentral distances of stations because we find strong coherent background noise for stations with long distances. During the DD iterations, we use a total of 10 iterations. The first five iterations for the initial weighting and next five for the reweighting, in which the data are reweighted to significantly reduce the residual. (See details about data weighting and reweighting in Text S3 and the DD iterations in Fig. S8.)

The final relocated aftershocks are shown in Figure 4c. We successfully relocate 4644 events within one month after the mainshock. For comparison, we also present the locations of 662 aftershocks from the current USGS-NEIC catalog (Fig. 4a) and their relocation with the DD technique (Fig. 4b). The mainshock denoted by the red star in Figure 4 comes from Yang *et al.* (2021). In addition, we use the same 1D velocity

Figure 3. Four examples of event windows detected from stations IE.DVCI (event a) and XP.BANN (events b–d) that are not included in the current USGS-NEIC catalog. (a–d) Their waveform windows (top) and 2D time–frequency feature maps (bottom). (e–h) P -wave arrival-time picks for all available traces. Station codes and a priori weights for all traces are labeled at the left side, and red bars mark the P -wave arrival times picked by the trained CNNs (Ross *et al.*, 2018). We also list distances between each aftershock to stations at the right side. Estimated event times and local magnitudes are shown in the title of each panel.

model (Table S4) to calculate P -wave travel times for our new catalog and use the same P -wave picking algorithm and velocity model to process the USGS-NEIC catalog. Then we compare travel-time residuals from these two catalogs. Here, the travel-time fits work as another quality test for the new catalog, which demonstrates the self-consistency between event origin times and locations. The travel-time residuals for our new catalog satisfy a Gaussian distribution, with smaller mean (0.05 s) and standard deviation (0.51 s) (Fig. S9) than the residuals from the USGS-NEIC catalog (0.16 and 1.20 s) (Figure S3b). From the new catalog, we observe a >40 -km-long, north-northwest-trending seismogenic zone (Fig. 4c). We characterize the aftershock distribution as three main clusters: the northern linear cluster (labeled as C1 in Fig. 4c), the north-northwest-trending shorter linear cluster in the southern segment (labeled as C2 in Fig. 4c), and the central dense cluster (labeled as C3 in Fig. 4c) that includes the relocated mainshock. Owing to our new catalog, we observe another branch that extends to the east, and we



approximate it as a possible cluster C4. We also observe a north-northwest-trend linear distribution in the relocated USGS-NEIC catalog (Fig. 4b), but it is difficult to delineate the central and southern sections of the seismogenic zone because of its low quantity and sparse distribution.

Statistical analysis of detected aftershocks

To better evaluate the relocated aftershocks, we use 19 stations to estimate their local magnitudes (M_L) (see the method in Text S4). We compare and evaluate our estimates using the same strategy to recalculate M_L for the USGS-NEIC catalog (see Text S4 and Fig. S10). The Omori law and Gutenberg–Richter law are two important empirical relations to describe the characteristics of aftershock sequences following large earthquakes. The Omori law (Omori, 1894; Utsu, 1961), $N = \frac{C}{(k+t)^p}$, describes the decay of aftershock activity with time after the mainshock. We fit the Omori relation using $M_L > 2.0$ aftershocks in our new catalog and estimate the p -value as 1.54 (Fig. 1b), which is slightly higher than the normal range of 0.9–1.5 in most natural cases (Utsu et al., 1995). Furthermore, we estimate the completeness magnitudes (M_c) for both the one-month-long USGS-NEIC aftershock catalog and our new catalog (see the method in Text S5). M_c is commonly defined as the lowest magnitude for earthquakes that can be uniformly detected in a space–time distribution (Rydelek and Selwyn Sacks, 1989), so a lower M_c value indicates a catalog with more complete magnitude range and reliable magnitude estimates. Compared with the $M_c = 2.8$ for the USGS-NEIC catalog (Fig. S12), our new catalog has an M_L range from 0.5 to 4.5 and a lower completeness magnitude of $M_c = 1.9$. We also fit the Gutenberg–Richter relation $\log(N) = a - b(M)$ (Gutenberg and Richter, 1944) using aftershocks over M_c , as shown in Figure 1c. From our new catalog, we estimate a b -value of 1.18, which is less anomalous to the global average level (1.0)

Figure 4. Aftershocks occurred within one month after the mainshock. Panels (a) and (b) show 662 aftershocks in the USGS-NEIC catalog before and after double-difference (DD) relocation, respectively. Panel (c) presents 4644 detected events using the CNN. Three well-defined aftershocks clusters are labeled as C1, C2, and C3 in (c). The question mark in (c) represents an additional possible cluster C4, which is further discussed in the Results and Discussion sections. The Quaternary Sawtooth fault is denoted by an orange curve. The northeast-trending Trans-Challis fault system (T-C faults) is mapped with gray curves. The red star denotes the relocated mainshock from Yang et al. (2021).

compared with the 1.38 value inferred from the higher completeness magnitude USGS-NEIC catalog (Fig. S12).

Spatial and temporal distributions of detected aftershocks

To better investigate the expansion of aftershocks, we analyze their spatiotemporal distribution. We first introduce the finite-fault coseismic slip model from Yang et al. (2021), which includes a northern fault plane F1 and a southern fault plane F2 (Fig. 5a). The peak slip of ~ 1 m is located near the stepover of these two fault planes at a depth of 13 km (Yang et al., 2021). We plot the aftershocks within the first seven days (black circles in Fig. 5a), and the dense central cluster is consistently located around the peak slip region. To better visualize the aftershock distribution, we plot aftershocks in our new catalog with estimated $M_L \geq 2.0$ (1217 events) in Figure 5b. There are two main fault systems within our study region; one is the north-northwest–south-southeast-trending northeast-dipping Sawtooth fault (the orange curve in Fig. 5b), and the other is the Trans-Challis fault zone (T-C faults, the gray curves in Fig. 5b), which is a northeast-striking Eocene normal fault

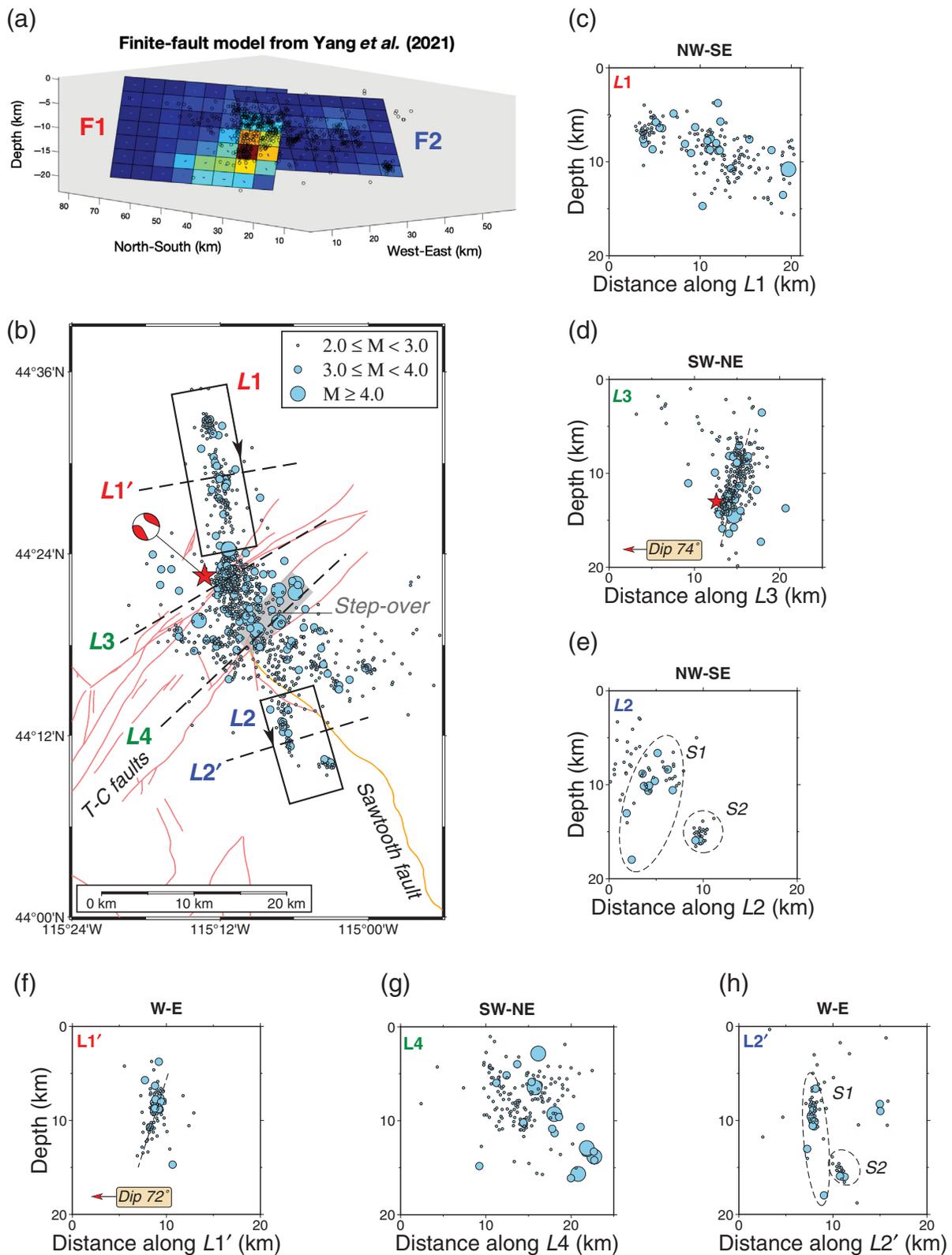


Figure 5. Spatial distributions of detected aftershocks with $M_L \geq 2.0$ in the first month (1217 events) after the mainshock (red star). (a) The finite-fault model from Yang *et al.* (2021), warm colors represent large slips, and black circles denote aftershocks. (b) The map view of the one-month-long aftershocks, events are size-coded by their magnitude ranges. The gray box represents

the stepover region that is analyzed in the Discussion section. (c–h) Six vertical cross sections along three aftershock clusters C1, C2, and C3, two small clusters S1 and S2. Aftershocks within the two dashed boxes and within 3 km from four black dashed lines are plotted on each cross section.

system (Lewis *et al.*, 2012). The northern terminus of the Sawtooth fault reaches to our central dense cluster C3, whereas the T-C faults cross the whole seismogenic zone in cluster C3 with a northeast trend.

Furthermore, we present six cross sections, which are along the possible strike and dip directions of clusters C1 ($L1, L1'$) and C2 ($L2, L2'$), and cross the central cluster C3 ($L3, L4$). Overall, the aftershocks reach to shallower depths, especially for cluster C1. Figure 5c,e shows the cross sections of clusters C1 and C2 along the strike directions. The aftershocks in cluster C1 have a much denser and wider concentration than cluster C2, which is supported by higher coseismic slip on F1 plane in Figure 5a. Laterally, we observe a west-dipping trend consistent with the fault orientation F1 inferred from finite-fault inversion by Yang *et al.* (2021) even though there are some diffuse deep events disturbing the dip angle (Fig. 5f). The overall west-dipping trend probably indicates an unmapped fault plane. We fit the west-dipping angle along the major linear trend as 72° , which is in agreement with the estimates of 70° – 74° for the west-dipping fault (Liberty *et al.*, 2020; Yang *et al.*, 2021). Besides the southern cluster C2, there is a small event cluster (labeled as S2 in Fig. 5e,h), which is separated from the major segment of cluster C2 (labeled as S1 in Fig. 5e,h) after the DD relocation. This small cluster S2 is disconnected from S1 in the relocated USGS-NEIC catalog and our initial global optimization location (see Fig. 4b; Fig. S6a). It also has a possible east-dipping trend (closer to the Sawtooth fault), which is different from the steep dipping trend of S1 (see Fig. 5h). Therefore, we consider S2 as a small separated cluster instead of one segment of cluster C2. Another two cross sections along $L3$ and $L4$ for the central cluster C3 cover two major branches of the Trans-Challis fault system (Fig. 5b). The aftershocks along $L3$ (Fig. 5d), including the relocated mainshock, have the densest concentration. The similar dipping trend with cluster C1 (Fig. 5f) indicates that the primary north-northwest-trending and west-dipping fault extends southward into the northern branch of the T-C faults. In addition, we observe that all large aftershocks with $M_L \geq 4.0$ are located along the T-C faults; three of them are near the mainshock, and six are along the southern branch of the T-C faults in cross-section $L4$ (see Fig. 5b,g). To further explore the temporal evolution of detected aftershocks, we also color code the one-month-long catalog by their occurrence times in Figure S11. During the first month after the mainshock, the sustained aftershocks (e.g., red circles) mostly occurred within cluster C1 and extended southward into cross section $L3$ (Fig. S11). Another concentration of sustained events extends from cross section $L4$ to the east, which is approximated as cluster C4 (Fig. S11). The temporal 3D evolution for the new catalog can be found in the supplemental Video S1.

Discussion

Owing to the well-trained CNN architecture and two earthquake location methods, we successfully build a more complete

one-month-long aftershock catalog that has seven times more events than the current USGS-NEIC catalog. This sequence is decaying during the first month, with decreasing number of larger events (Fig. 1d). In the first several hours of the sequence, there is a clear reduction of lower magnitude events relative to later time periods. So, any secondary events in the 20 s detection intervals are predominantly events below the completeness magnitude M_c level (Fig. 1d). We focus on the main activity in the first month as this has the clearest relationship with the mainshock faulting. By analyzing the spatiotemporal distributions of our new catalog, we observe two linear aftershock clusters C1 and C2 that enable us to delineate a multifaulting system. Cluster C1 delineates a northern north-northwest-striking and 72° west-dipping unmapped fault, which has an opposite dipping trend compared with the east-dipping Sawtooth fault. Cluster C2 probably delineates another northwest-striking and steeply dipping unmapped fault in the southern seismogenic zone and is close to the mapped Sawtooth fault. The central cluster C3 spans the northern terminus of the Sawtooth fault, indicating a complex seismicity zone in the crossover section between clusters C1 and C2. This interpretation might be consistent with the notion that recent slip and potential activity of the Sawtooth fault are migrating northward (Thackray *et al.*, 2013). It is likely that the cluster C4 delineates some additional west–east structures. The focal mechanisms in the central crossover section also suggest some normal faulting (Yang *et al.*, 2021). Thus, we conclude that cluster C4 likely delineates additional west–east faults, but it appears diffuse because of structural heterogeneity or other reasons.

Yang *et al.* (2021) combined BP images, InSAR ground deformation, teleseismic finite-fault inversion, and one-month-long USGS-NEIC aftershock relocation and proposed a multifault rupture model, which includes a northern predominantly strike-slip subfault with a strike of 162° and dipping to the southwest, a southern predominantly normal subfault with a strike of 156° and steep northeastward dip, and a stepover (10 km wide at the surface and narrowing with depth) that allows the rupture to transition to the southern fault. Pollitz *et al.* (2020) also analyzed the first two weeks aftershocks and geodetic data, then proposed a faulting model. It includes a predominant south-southeast-trending fault that accommodated most of coseismic slip and inferred afterslip on two northeast-trending faults that may represent reactivated fabrics of the Trans-Challis fault system. In comparison, our new catalog further delineates the major north-northwest-striking and west-dipping fault trend. We also observe similar dipping trends between aftershocks in cross-sections $L1'$ and $L3$ (Fig. 5d,f), which indicate the major north-northwest-striking fault may extend further southward and is probably intersected by the southern branch of the Trans-Challis fault system (cross-section $L4$). We note that the east dipping of fault F2, defined in the finite-fault model of Yang *et al.* (2021) (Fig. 5a), is different from the possible west-southwest-steep dipping direction of our cluster C2 (Fig. 5h). One possible

explanation is that the southern fault F2 has a shorter extension (<10 km), compared with the northern fault F1's ~20 km extension, and it does not generate a wide aftershock zone within the first month. Thus, we need more evidence to determine the accurate dipping angle of the southern branch. Furthermore, we do not resolve clear northeast-striking faults as inferred by Pollitz *et al.* (2020), which correlate with the Trans-Challis fault system.

According to the earthquake rupture process determined by the BP imaging technique (Yang *et al.*, 2021), this event propagated to the southeast ~20 km. When it reached the vicinity of the northwestern terminus of the Sawtooth fault at ~4 s after the origin, the rupture changed its trajectory to the southwest. After a few seconds, the rupture continued propagating to the southeast ~25 km along a northwest–southeast-striking and steeply dipping fault. This overall north–northwest rupture trajectory is reinforced by our expanded aftershock catalog. From the spatial distribution of aftershocks within the first month, we observe there are six large aftershocks with $M_L \geq 4.0$ in cross section L4, presenting a northeast–southwest-trending expansion along the southern branch of the T-C faults (Fig. 5b,g); this has a quite different trend from cluster C1. The absence of west-dipping trend in cross-section L4 also indicates that the southeast earthquake rupture trajectory was changed at cross-section L4. Therefore, based on the expanded aftershock analysis, we determine the stepover region covers the southern branch of the T-C faults and northern terminus of the Sawtooth fault (represented by the gray box in Fig. 5b), which changes the rupture of the mainshock and induces strong seismicity. From the temporal evolution of aftershocks, besides the relocated mainshock vicinity, the determined stepover region is steadily active during the first month (more red circles in Fig. S11 and supplemental Video S1). This sustained seismic activity, including the possible cluster C4, also suggests that the stepover region contributes not only coseismic slip but also possible afterslip during the first month. Our analysis from the expanded aftershocks provides additional support for the complex stepover region possibly being related to old crustal fractures from the Trans-Challis fault zone, which were dynamically reactivated by regional stress changes from the mainshock (Liberty *et al.*, 2020; Yang *et al.*, 2021). In addition, the peak slip rate (Fig. 5a) in cross sections L3 and L4 indicate that the old fault reactivation accounts for the earthquake rupture trajectory.

Conclusion

In this study, we transform event and noise datasets into 2D time–frequency feature maps, which are used to train our neural networks. The trained CNN is used to detect potential events within one-month-long continuous records after the mainshock of the 2020 M_w 6.5 Stanley, Idaho, earthquake. After detection, we apply another trained CNN to automatically pick *P*-wave arrival times for the detected events. In the study region, we use 19 permanent and local temporary stations that surround the mainshock to locate the detected

events. With respect to the previous studies, our primary contributions for this earthquake include (1) a more complete (4644 events and a completeness magnitude $M_c = 1.9$) one-month-long aftershock catalog; (2) a high-resolution geometry for the north–northwest-striking, west-dipping north fault and northwest-striking, steep-dipping south fault; and (3) further documentation of the earthquake rupture trajectory and the central sustained seismicity. This new aftershock catalog is compatible with results from regional BP and coseismic slip distribution analysis. It helps to constrain rupture details for the 2020 M_w 6.5 Stanley, Idaho, earthquake.

Data and Resources

The supplemental material for this article includes a main supporting material, our new detected and relocated one-month-long aftershock catalog, and Video S1 for the spatial and temporal evolution of the new catalog. All seismic data from the permanent networks (US, IE, IW, and US) and temporary network (XP) are archived at Incorporated Research Institutions for Seismology (IRIS). The continuous seismic records were downloaded and preprocessed using ObsPy (Beyreuther *et al.*, 2010). The USGS-NEIC earthquake catalog was available at <https://earthquake.usgs.gov/earthquakes/search> (last accessed November 2021). The original convolutional neural network (CNN) architecture for earthquake detection was built by Perol *et al.* (2018), and the library is available at <https://github.com/tperol/ConvNetQuake> (last accessed April 2022). The CNN architecture for *P*-wave arrival times picking was built by Ross *et al.* (2018). Earthquake location programs were developed by Lomax *et al.* (2000) and Waldhauser and Ellsworth (2000). The velocity model for location was derived from Shen and Ritzwoller (2016). All figures were plotted using Generic Mapping Tools (GMT) version 6.1.0 (Wessel *et al.*, 2019) and Matplotlib version 3.3.0 (Hunter, 2007). More details about the Stanley earthquake event are available at <https://www.idahogeology.org/geologic-hazards/earthquake-hazards/stanley-earthquake> (last accessed April 2022).

Declaration of Competing of Interests

The authors acknowledge that there are no conflicts of interest recorded.

Acknowledgments

The authors appreciate thoughtful reviews by three anonymous reviewers. Hejun Zhu's research is supported by U.S. National Science Foundation Grant Number EAR2042098 and The University of Texas at Dallas (UTD) faculty startup funding. Thorne Lay's research on earthquakes is supported by U.S. National Science Foundation Grant Number EAR1802364. Lingling Ye's earthquake study is supported by National Natural Science Foundation of China (Grant Number 41874056).

References

- Allen, R. V. (1978). Automatic earthquake recognition and timing from single traces, *Bull. Seismol. Soc. Am.* **68**, no. 5, 1521–1532.
- Barrett, S. A., and G. C. Beroza (2014). An empirical approach to subspace detection, *Seismol. Res. Lett.* **85**, no. 3, 594–600.

- Benz, H., N. McMahon, R. Aster, D. Mcnamara, and D. Harris (2015). Hundreds of earthquakes per day: The 2014 Guthrie, Oklahoma, earthquake sequence, *Seismol. Res. Lett.* **86**, 1318–1325.
- Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010). ObsPy: A Python toolbox for seismology, *Seismol. Res. Lett.* **81**, no. 3, 530–533, doi: [10.1785/gssrl.81.3.530](https://doi.org/10.1785/gssrl.81.3.530).
- Brown, J. R., G. C. Beroza, and D. R. Shelly (2008). An autocorrelation method to detect low frequency earthquakes within tremor, *Geophys. Res. Lett.* **35**, no. 16, 428–451.
- Chen, C., and A. A. Holland (2016). PhasePApy: A robust pure Python package for automatic identification of seismic phases, *Seismol. Res. Lett.* **87**, no. 6, 1384–1396.
- Dokht, R. M. H., H. Kao, R. Visser, and B. Smith (2019). Seismic event and phase detection using time-frequency representation and convolutional neural networks, *Seismol. Res. Lett.* **90**, no. 2A, 481–490.
- Enescu, B., J. Mori, and M. Miyazawa (2007). Quantifying early aftershock activity of the 2004 mid-Niigata Prefecture earthquake (Mw6.6), *J. Geophys. Res.* **112**, no. B4, doi: [10.1029/2006JB004629](https://doi.org/10.1029/2006JB004629).
- Gans, P. B., and W. A. Bohrsen (1998). Suppression of volcanism during rapid extension in the basin and range province, United States, *Science* **279**, no. 5347, 66–68.
- Gibbons, S. J., and F. Ringdal (2010). The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Roy. Astron. Soc.* **165**, no. 1, 149–166.
- Gomberg, J. S., K. M. Shedlock, and S. W. Roecker (1990). The effect of S-wave arrival times on the accuracy of hypocenter estimation, *Bull. Seismol. Soc. Am.* **80**, no. 6A, 1605–1628.
- Gutenberg, B., and C. F. Richter (1944). Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.* **34**, no. 4, 185–188.
- Harris, D. B., and D. A. Dodge (2011). An autonomous system for grouping events in a developing aftershock sequence, *Bull. Seismol. Soc. Am.* **101**, no. 2, 763–774.
- Hochreiter, S. (1991). Untersuchungen zu dynamischen neuronalen netzen, *Diploma, Technische Universität München*, **91**, no. 1, 1–65 (in German).
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9**, no. 3, 90–95, doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).
- Idaho Geological Survey. The Stanley earthquake, available at <https://idahogeology.org/DrawOnePage.aspx?PageID=41> (last accessed July 2022).
- Kingma, D. P., and J. Ba (2017). Adam: A method for stochastic optimization, doi: [10.48550/arXiv.1412.6980](https://doi.org/10.48550/arXiv.1412.6980).
- Kortström, J., M. Uski, and T. Tiira (2015). Automatic classification of seismic events within a regional seismograph network, *Comput. Geosci.* **87**, no. 2016, 22–30.
- Lewis, R. S., P. K. Link, L. R. Stanford, and S. P. Long (2012). Geologic map of Idaho, *Idaho Geol. Surv.* Geological Map 9, scale 1:750,000.
- Li, Z., M.-A. Meier, E. Hauksson, Z. Zhan, and J. Andrews (2018). Machine learning seismic wave discrimination: Application to earthquake early warning, *Geophys. Res. Lett.* **45**, no. 10, 4773–4779.
- Liberty, L. M., Z. M. Lifton, and T. D. Mikesell (2020). The 31 March 2020 Mw 6.5 Stanley, Idaho, earthquake: Seismotectonics and preliminary aftershock analysis, *Seismol. Res. Lett.* **92**, no. 2A, 663–678.
- Lomax, A., J. Virieux, P. Volant, and C. Berge-Thierry (2000). Probabilistic earthquake location in 3D and layered models, in *Advances in Seismic Event Location*, C. H. Thurber (Editor), Springer Netherlands, Dordrecht, The Netherlands, 101–134, doi: [10.1007/978-94-015-9536-0_5](https://doi.org/10.1007/978-94-015-9536-0_5).
- Mousavi, S. M., S. P. Horton, C. A. Langston, and B. Samei (2016). Seismic features and automatic discrimination of deep and shallow induced-microearthquakes using neural network and logistic regression, *Geophys. J. Int.* **207**, no. 1, 29–46.
- Mousavi, S. M., W. Zhu, Y. Sheng, and G. C. Beroza (2019). CRED: A deep residual network of convolutional and recurrent units for earthquake signal detection, *Sci. Rep.* **9**, no. 1, doi: [10.1038/s41598-019-45748-1](https://doi.org/10.1038/s41598-019-45748-1).
- Nippress, S. E. J., A. Rietbrock, and A. E. Heath (2010). Optimized automatic pickers: Application to the ANCORP data set, *Geophys. J. Int.* **181**, no. 2, 911–925.
- Omori, F. (1894). On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo* **7**, 111–200.
- Peng, Z., J. E. Vidale, M. Ishii, and A. Helmstetter (2007). Seismicity rate immediately before and after main shock rupture from high-frequency waveforms in Japan, *J. Geophys. Res.* **112**, no. B3, doi: [10.1029/2006JB004386](https://doi.org/10.1029/2006JB004386).
- Perol, T., M. Gharbi, and M. Denolle (2018). Convolutional neural network for earthquake detection and location, *Sci. Adv.* **4**, e1700578, doi: [10.1126/sciadv.1700578](https://doi.org/10.1126/sciadv.1700578).
- Pollitz, F. F., W. C. Hammond, and C. W. Wicks (2020). Rupture process of the M 6.5 Stanley, Idaho, earthquake inferred from seismic waveform and geodetic data, *Seismol. Res. Lett.* **92**, no. 2A, 699–709.
- Richins, W., J. Pechmann, R. Smith, C. Langer, S. Goter, J. Zollweg, and J. King (1987). The 1983 Borah Peak, Idaho, earthquake and its aftershocks, *Bull. Seismol. Soc. Am.* **77**, 694–723.
- Ross, Z. E., M.-A. Meier, and E. Hauksson (2018). P wave arrival picking and first-motion polarity determination with deep learning, *J. Geophys. Res.* **123**, no. 6, 5120–5129.
- Ross, Z. E., D. T. Trugman, E. Hauksson, and P. M. Shearer (2019). Searching for hidden earthquakes in Southern California, *Science* **364**, no. 6442, 767–771.
- Rydelek, P. A., and I. Selwyn Sacks (1989). Testing the completeness of earthquake catalogues and the hypothesis of self-similarity, *Nature* **337**, 251–253.
- Saragiotis, C. D., L. J. Hadjilentiadis, and S. M. Panas (2002). PAI-S/K: A robust automatic seismic P phase arrival identification scheme, *IEEE Trans. Geosci. Remote Sens.* **40**, no. 6, 1395–1404.
- Shelly, D. R., G. C. Beroza, and S. Ide (2007). Non-volcanic tremor and low-frequency earthquake swarms, *Nature* **446**, no. 7133, 305–307.
- Shen, W., and M. H. Ritzwoller (2016). Crustal and uppermost mantle structure beneath the United States, *J. Geophys. Res.* **121**, no. 6, 4306–4342.
- Smith, R. B., and M. L. Sbar (1974). Contemporary tectonics and seismicity of the western United States with emphasis on the intermountain seismic belt, *Geol. Soc. Am. Bull.* **85**, no. 8, 1205–1205.
- Tajima, F., and H. Kanamori (1985). Global survey of aftershock area expansion patterns, *Phys. Earth Planet. In.* **40**, no. 2, 77–134.
- Tan, Y. J., F. Waldhauser, W. L. Ellsworth, M. Zhang, W. Zhu, M. Michele, L. Chiaraluce, G. C. Beroza, and M. Segou (2021). Machine-learning-based high-resolution earthquake catalog reveals how complex fault structures were activated during the 2016–2017 Central Italy sequence, *Seismic Record* **1**, no. 1, 11–19.

- Thackray, G. D., D. W. Rodgers, and D. Streutker (2013). Holocene scarp on the Sawtooth fault, central Idaho, USA, documented through lidar topographic analysis, *Geology* **41**, no. 6, 639–642.
- U.S. Geological Survey. Quaternary fault and fold database of the United States, available at <https://earthquake.usgs.gov/cfusion/qfault> (last accessed July 2022).
- U.S. Geological Survey, National Earthquake Information Center. Mw 6.5 Stanley earthquake event page, available at <https://earthquake.usgs.gov/earthquakes/eventpage/us70008jr5/executive> (last accessed June 2021).
- Utsu, T. (1961). A statistical study on the occurrence of aftershocks, *Geophys. Mag.* **30**, 521–605.
- Utsu, T., Y. Ogata, and R. Matsuura (1995). The Centenary of the Omori Formula for a decay law of aftershock activity, *J. Phys. Earth* **43**, no. 1, 1–33.
- Waldhauser, F., and W. Ellsworth (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward Fault, California, *Bull. Seismol. Soc. Am.* **90**, 1353–1368, doi: [10.1785/0120000006](https://doi.org/10.1785/0120000006).
- Weller, O., D. Lange, F. Tilmann, D. Natawidjaja, A. Rietbrock, R. Collings, and L. Gregory (2012). The structure of the Sumatran Fault revealed by local seismicity, *Geophys. Res. Lett.* **39**, no. 1, L01306, doi: [10.1029/2011GL050440](https://doi.org/10.1029/2011GL050440).
- Wessel, P., J. F. Luis, L. Uieda, R. Scharroo, F. Wobbe, W. H. F. Smith, and D. Tian (2019). The generic mapping tools version 6, *Geochem. Geophys. Geosys.* **20**, no. 11, 5556–5564, doi: [10.1029/2019GC008515](https://doi.org/10.1029/2019GC008515).
- Withers, M., R. Aster, C. Young, J. Beiriger, and J. Trujillo (1998). A comparison of select trigger algorithms for automated global seismic phase and event detection, *Bull. Seismol. Soc. Am.* **88**, no. 1, 95–106.
- Yang, J., H. Zhu, T. Lay, Y. Niu, L. Ye, Z. Lu, B. Luo, H. Kanamori, J. Huang, and Z. Li (2021). Multifault opposing-dip strike-slip and normal-fault rupture during the 2020 Mw 6.5 Stanley, Idaho earthquake, *Geophys. Res. Lett.* **48**, no. 10, e2021GL092510, doi: [10.1029/2021GL092510](https://doi.org/10.1029/2021GL092510).
- Zhou, Y., H. Yue, L. Fang, S. Zhou, L. Zhao, and A. Ghosh (2021). An earthquake detection and location architecture for continuous seismograms: Phase picking, association, location, and matched filter (PALM), *Seismol. Res. Lett.* **93**, no. 1, 413–425, doi: [10.1785/0220210111](https://doi.org/10.1785/0220210111).

Manuscript received 30 November 2021
Published online 21 July 2022