FULL PAPER

Application of deep learning-based neural networks using theoretical seismograms as training data for locating earthquakes in the Hakone volcanic region, Japan

Daisuke Sugiyama[1], Seiji Tsuboi[1*] and Yohei Yukutake[2,3]

Abstract

In the present study, we propose a new approach for determining earthquake hypocentral parameters. This approach integrates computed theoretical seismograms and deep machine learning. The theoretical seismograms are generated through a realistic three-dimensional Earth model, and are then used to create spatial images of seismic wave propagation at the Earth's surface. These snapshots are subsequently utilized as a training data set for a convolutional neural network. Neural networks for determining hypocentral parameters such as the epicenter, depth, occurrence time, and magnitude are established using the temporal evolution of the snapshots. These networks are applied to seismograms from the seismic observation network in the Hakone volcanic region in Japan to demonstrate the suitability of the proposed approach for locating earthquakes. We demonstrate that the determination accuracy of hypocentral parameters can be improved by including theoretical seismograms for different earthquake locations and sizes, in the learning data set for the deep machine learning. Using the proposed method, the hypocentral parameters are automatically determined within seconds after detecting an event. This method can potentially serve in monitoring earthquake activity in active volcanic areas such as the Hakone region.

Keywords: Deep learning, Hypocenter determination, Synthetic seismograms

Introduction

Currently, the machine learning technique is widely exploited in disciplines of science and technology including seismology (Maggi et al. 2017; Rouet-Leduc et al. 2017; Malfante et al. 2018; Nakano et al. 2019; Seydoux et al. 2020). Specifically, combining numerical seismograms and pattern recognition for earthquake location was proposed by Käufl et al. (2014, 2015, 2016a, b). According to these studies, the technique can be applied for determining fault parameters. In this study,

we propose a different technique for earthquake location and determining its magnitude. The proposed approach is based on theoretical seismograms from a realistic Earth model and a deep learning-based convolutional neural network (CNN). The approach relies on spatial images of seismic wave propagation at the Earth's surface.

The spectral element method (SEM) is used for numerical modeling of seismic wave propagation in a realistic three-dimensional (3D) Earth model. The 3D SEM was initially employed in seismology to perform local and regional simulations (Faccioli et al. 1997; Komatitsch 1997; Komatitsch and Vilotte 1998), and later adapted for wave propagation at the Earth scale (Komatitsch and Tromp 2002a; b; Komatitsch et al. 2005; Tsuboi et al. 2003, 2016; Carrington et al. 2008; Rietmann et al. 2012). In the present study, the SEM is used to simulate

*Correspondence: tsuboi@jamstec.go.jp
[1] Research Institute for Value-Added-Information Generation, Japan Agency for Marine-Earth Science and Technology, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan
Full list of author information is available at the end of the article

Sugiyama *et al. Earth, Planets and Space* (2021) 73:135

Page 2 of 12

a regional scale seismic wave propagation problem. The surface propagation of seismic waves at each time step is calculated for different earthquake locations and sizes, and the values are differentiated to obtain the seismic wave propagation in terms of the velocity. These seismic wave propagation and temporal evolution images are then utilized as training data for a deep machine learning CNN. Considering that setting any number of hypocenters or earthquake magnitude is possible, creating a large volume of training data is easy. Therefore, deep machine learning is expected to improve the hypocenter determination accuracy.

A CNN is a deep learning model that has attracted attention in computer visualization because of the accuracy of its high image recognition tasks, such as in the ImageNet's large scale visual recognition challenge (Vaillant et al. 1994; Krizhevsky et al. 2012; Simonyan and Zisserman 2014; Szegedy et al. 2015). Artificial neural networks are commonly used for classifying seismograms (Wang and Teng 1995, 1997), and recently, CNNs have been employed for earthquake detection (Perol et al. 2018; Ross et al. 2018; Zhu and Beroza 2019). In this study, theoretical seismograms computed for different earthquake locations and magnitudes are utilized for creating images of seismic wave propagation at the Earth's surface. These snapshots then serve as training data in the training phase of the CNN to estimate the hypocenter location, earthquake occurrence time, and earthquake magnitude.

Here we apply our proposed technique to seismograms from the seismic observation network in the Hakone volcanic region in Japan to demonstrate the suitability of the proposed approach for locating earthquakes.
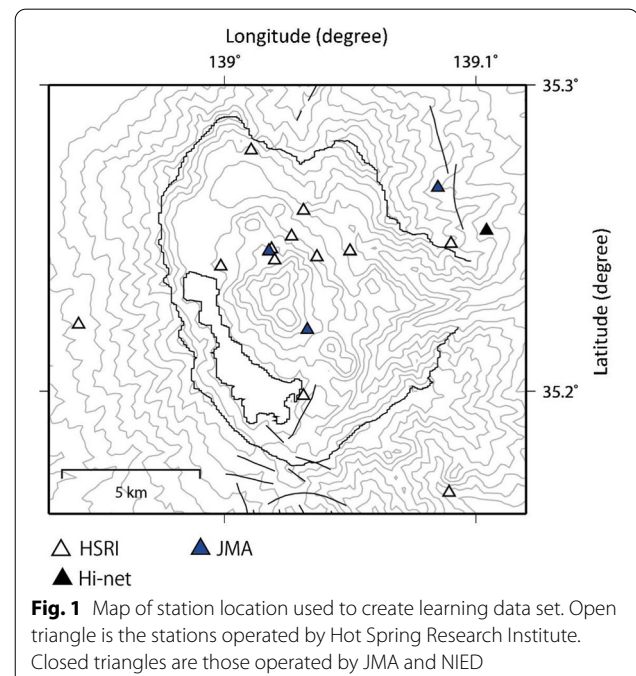
Theoretical seismograms

The SPECFEM3D program package (Komatitsch and Tromp 2002a, b; Komatitsch et al. 2005; Tsuboi et al. 2003) was used to generate theoretical seismograms for a realistic Earth model using the SEM. The theoretical seismograms were synthesized using 48×48, yielding 2304 slices. Each slice was allocated to a CPU core of a parallel supercomputer, and then subdivided into 768×768 grid points, thereby enabling generation of theoretical seismograms with an accuracy of 5.6 s and longer (e.g., Tsuboi et al. 2016). To generate a mesh for the computation, the surface topography and attenuation were included. A 3D seismic wave speed model of the Hakone region constructed by Yukutake et al. (2015) was also considered. Three-component displacement seismograms, lasting 30 s, were synthesized at seismographic stations operated by the Hot Springs Research Institute of the Kanagawa Prefecture using a sampling interval of 0.1 s. A map showing the stations in this network is shown in

Fig. 1. The theoretical seismograms were then differentiated once to create velocity seismograms. Subsequently, 32×32 images occupying the seismic network shown in Fig. 1 were generated from the seismograms synthesized for seismic stations. To create the training data set, the hypocenter was set in the Hakone region using a spatial distance of 0.02° to produce the seismic wave propagation images (Fig. 2). The hypocentral depth was varied over a vertical distance of 2 km, and the maximum depth was set at 10 km. The moment magnitude (Mw) was also randomly altered between approximately 2.2 and 4.5, to create images for the training data required for deep learning of the CNN. The moment tensor of the source can be varied while creating the training data; however, this was not performed in the present study. A typical reverse fault mechanism was assumed for the region, and was not expected to affect the estimation accuracy. An example of a wave propagation image versus time is displayed in Fig. 3. Among the 1998 earthquakes involved, 80% of the events were included in the training data set. Consequently, about 600,000 images of seismic-wave propagation at the surface were generated, with a 32×32 and time interval of 0.1 s.

Convolutional neural network

In the present study, two types of CNNs were tested. First, the test using an ordinary CNN is described. A 32×32 image served as the input, and hypocentral parameters including the latitude, longitude, depth, occurrence time, and magnitude, were estimated

Fig. 1 Map of station location used to create learning data set. Open triangle is the stations operated by Hot Spring Research Institute. Closed triangles are those operated by JMA and NIED

Sugiyama *et al. Earth, Planets and Space* (2021) 73:135
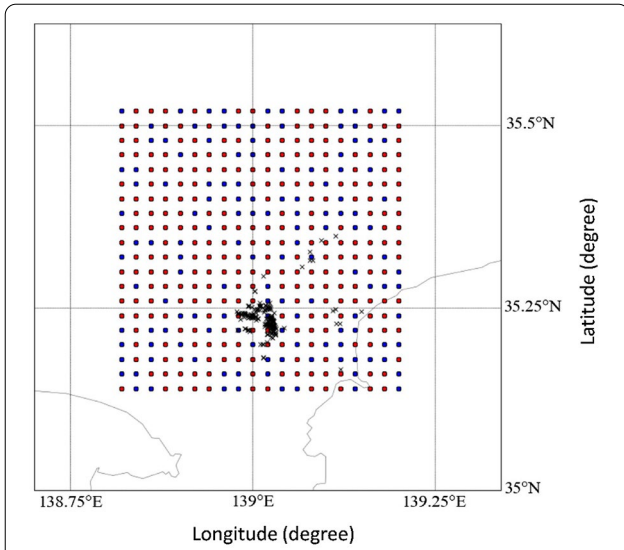
Page 3 of 12

Fig. 2 Red dots are the earthquakes used for learning and blue dots are the earthquakes used for testing. The size of the dots is proportional to the magnitude. Crosses are epicenters of observed earthquakes. The total number of earthquakes used is 2632 and we used 2236 events as the training data set. Some of the earthquakes are removed from the data set due to apparent numerical problems during the simulation
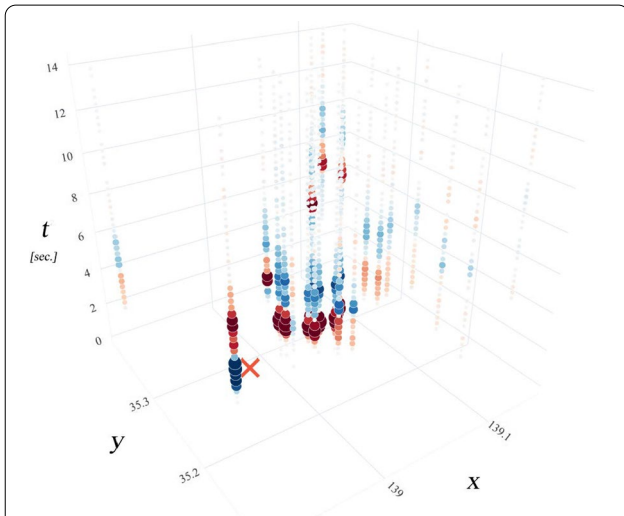

Fig. 3 Map of example propagation image used for learning process. Colors indicate vertical displacement from $-1.4e-6$ m to $1.4e-6$ m as they change from blue to red

through regression. The network architecture employed was similar to that in previous studies (LeCun et al. 1999; Onishi and Sugiyama 2017). As the convolutional layer, a set of learnable filters were applied to the input image to extract the characteristic features, while a pooling layer was used to reduce sensitivity associated with the location of the characteristic feature. The fully connected layer conducted the reasoning based on the output from the convolutional and pooling layers. These layers were sequentially connected to produce a CNN capable of associating an input image to a specific hypocenter parameter. TensorFlow was employed as the framework to establish a LeNet-based CNN (LeCun et al. 1999).

The input image data created using the seismic wave propagation at the Earth's surface comprised 32×32 pixels. The convolutional layer map and the input image were utilized to generate an output image using convolutional filters. The dimensions of the convolutional filters used were $3 \times 3 \times n$ or $7 \times 7 \times n$, and can be represented as follows, where both nW and nH $= 1$ for $3 \times 3 \times n$ and 3 for $7 \times 7 \times n$:

$$u_{i,j,k} = f\left(b_k + \sum_{i'=-nW}^{nW} \sum_{j'=-nH}^{nH} \sum_{k'=1}^{nC} Z_{i+i',j+j',k'} \cdot W_{i',j',k',k}\right), \quad (1)$$

where $u_{i,j,k}$ and $z_{i,j,k}$ are the (i, j)-th pixel of the k-th channel in the output and input images for the layer, respectively. The weights $w_{i',j',k',k}$ constitute a filter that was applied to the k'-th channel of the input image, b_k is the bias for the k-th channel, and n is the number of channels in the input layer. In this study, as only the vertical seismogram component was used, the number of channels (n) is 1. Considering $f(x)$, a ReLU function which is commonly used for regression analysis in neural networks, this can be expressed as follows:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases}. \quad (2)$$

Zero padding was not utilized on the input image, and the output image size was reduced by four pixels in two directions. The pooling layer was processed using max pooling with a 2×2 filter and a stride of 2, thereby downsampling the image size from $h \times h \times n$ to $h/2 \times h/2 \times n$ through the following expression:

$$u_{i,j,k} = \max \left\{ z_{2j+j',2i+i',k} | i', j' = (0, 1) \right\}. \quad (3)$$

In the final connected layer, to avoid overfitting, a 50% dropout was applied. Dropout is a procedure for avoiding overfitting by creating a pattern without updating the neural network weight of a designated layer. A modified stochastic gradient descent (Kingma and Ba 2015) served as the optimizer while the learning proceeded using a batch size of 50 and epoch of 250. The calculated root mean square error (RMSE) for the learning data was then employed as a loss function.

Sugiyama *et al. Earth, Planets and Space*   (2021) 73:135

Page 4 of 12

Batch normalization was applied (Ioffe and Szegedy [2015](#)) before calculating the activation function. This process renders characteristic features uncorrelated by converting the mean and standard deviation of any selected batch to 0 and 1, respectively.

The mean and standard deviation for a batch $\mathscr{B} = \{\mathbf{u}_1 \cdots \mathbf{u}_m\}$ were calculated as follows:

$$\mu_{\mathscr{B}} = \frac{1}{m} \sum_{i=1}^{m} \mathbf{u}_i,$$

$$\sigma_{\mathscr{B}}^2 = \frac{1}{m} \sum_{i=1}^{m} (\mathbf{u}_i - \mu_{\mathscr{B}})^2, \tag{4}$$

The normalized data $\mathbf{y}_i$ can be obtained using the learned parameters γ and β from the following equations:

$$\widehat{\mathbf{x}}_i = \frac{\mathbf{x}_i - \mu}{\sqrt{\sigma_{\mathscr{B}}^2 + \varepsilon}},$$
$$\mathbf{y}_i = \gamma \widehat{\mathbf{x}}_i + \beta. \tag{5}$$

where ε is a constant error term added for numerical stability.

Bayesian optimization (Shahriari et al. [2015](#)) was used to search for optimized hyperparameters for the proposed learning model to avoid manual tuning via trial and error. A grid search was unsuitable because of the high number of combinations for the search target. The Bayesian optimization is a technique for efficiently searching hyperparameters. These parameters are estimated at high resolution using approximate loss distribution functions, and assuming that the loss distribution of the hyperparameters learning model is Gaussian. Following the examination of each case, a highly accurate model was produced using few parameters. The loss function for the Bayesian optimization was the RMSE estimated for the validation data. The hyperparameter search involved the following conditions: the convolution layers varied between 1 and 5; the fully connected layers ranged between 0 and 3; the convolution filter size was between 3 and 9; the fully connected nodes varied between 50 and 1000; max pooling was applied to each convolution layer; the dropout was between 0 and 50%; and batch normalization was applied on ReLU.

Validation and testing for 2DCNN
The model was then constructed after operating 20 GPUs for approximately 2 days. Amplitude ratios of the short-time average (STA) computed for 0.5 s (for 5 images) to the long-time average (LTA) computed for 5 s (for 50 images) were determined, and hypocentral parameters estimation was started when STA/LTA exceeded 4.

The 600,000 training data images were divided into training data (80%) and validation data (20%) and the hypocentral parameters are determined. In partitioning the training data, earthquakes involved in validation were excluded from the testing (Fig. [2](#)). The network parameters were optimized for estimating the hypocentral parameters by minimizing discrepancies in the training data set. After the optimization, these were verified using the validation data set. To estimate the earthquake occurrence time, the time for the ith image was designated t_i and the occurrence time t_0 was estimated from each image. Considering that one network was used for five hypocentral parameters, some parameters such as the magnitude, demonstrated that the loss function apparently was not improved at certain stages, indicating overfitting. Therefore, the use of at least two networks; one for the longitude and latitude and another for the depth, magnitude, and occurrence time seems necessary to prevent overfitting. The network construction to suit each hypocentral parameter requires further analysis. Each seismic wave propagation image requires hypocentral parameters output, so that the evolution of the estimates can be observed as the images are sorted. The hypocentral parameter values obtained after 8 s duration were considered the estimates. The errors associated with the estimated values are comparable to those from the validation data set, and the calculated magnitudes for all earthquakes used for validation are presented in Table [1](#).

The ConvNetQuake (Perol et al. [2018](#)), a CNN used for classifying earthquakes, was also tested for estimating hypocentral parameters. The results from the ConvNetQuake presented in Table [1](#) include regression data for hypocentral parameters from the learning stage. The results reveal that although hypocentral parameters can be adequately determined, the precision values are inferior to those from the 3D CNN examined subsequently.

3D convolutional neural network
Seismic wave propagation with time was tested using 6 s images corresponding to a $32 \times 32 \times 10$ event as the input. Hypocentral parameters such as the latitude, longitude, depth, occurrence time, and magnitude were then estimated through regressions. The architecture of the 3D CNN used in the present study is displayed in Fig. [4](#), with 3D CNN reflecting an extension of 2D CNN (Tran et al. [2015](#)). This involves using a time series event (e.g., 3D structure from $x \times y \times t$ shown in Fig. [4](#)) as input for learning and estimation. 3D CNN has been used for action recognition in movies and for learning time series data. The 3D CNN learning may involve pretraining using an existing data set for human action recognition

Sugiyama *et al. Earth, Planets and Space*   (2021) 73:135

Page 5 of 12

Table 1 Root Mean Square Errors (RMSE) of hypocentral parameters estimated by various networks

Method	Training	Estimation	Root Mean Square Error (RMSE)				
			Longitude (degree)	Latitude (degree)	Depth (km)	Time (s)	Magnitude
1DCNN (ConvNetQuake[a])	Simulation data (80%)	Simulation data (20%)	0.0406	0.0365	1.679	–	0.0930
2DCNN	Simulation data (80%)	Simulation data (20%)	0.0213	0.0184	1.3267	0.5002	0.1018
3DCNN	Simulation data (80%)	Simulation data (20%)	0.0100	0.0078	0.6208	0.2095	0.0371
		Observed data	0.0301	0.0291	1.3062	1.3816	0.5830
3DCNN + CUTOUT	Simulation data (80%)	Simulation data (20%)	0.0075	0.0091	0.5298	0.2444	0.0311
		Observed data	0.0167	0.0313	1.7504	0.9427	0.6063

[a] ConvNetQuake (Perol et al. 2018) is a task to classify earthquake location into specific regions. The results show in this table for ConvNetQuake is obtained by added our task to perform regression to locate hypocentral parameters during the learning stage

Fig. 4 Network architecture used for the 3D convolutional neural network
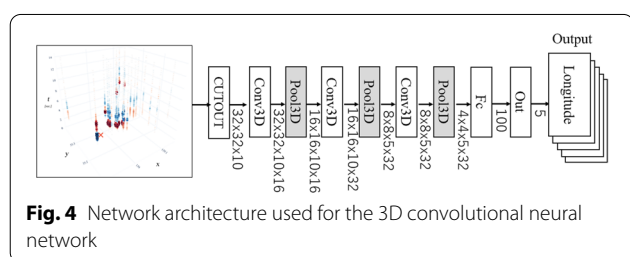
such as the Kinetics data set (Carreira and Zisserman 2017). However, such a data set for human activities and materials is unsuitable for the present study. Therefore, pretraining was not conducted, but rather, a small 3D CNN was used to construct the model experimentally. A 2DConv + LSTM model (Donahue et al. 2015) combining 2D CNN and long short-term memory (LSTM) models for time series learning is available. In addition, a two-stream model (Carreira and Zisserman 2017) that extracts motion characteristics and learns RGB simultaneously for time series learning also exists. Searching for optimal parameters using these models should be considered in future research. Hypocentral parameter estimation starts when the STA/LTA ratio is higher than 4 as in the ordinary CNN. A duration of 6 s (60 images) was employed, with estimates performed 20 times (up to 2 s after event detection) using 60 snapshots images, thus, yielding hypocentral parameters 8 s after an event detection.

A 3D CNN model generally involves a longer learning time compared to a 2D CNN. In the present study, we realized that this was linked to the I/O random access, instead of the GPU performance. To handle this problem, the complete training data set was loaded to the memory, thereby shortening the learning time to 1 epoch in 54 min which amounts to 108 GPU h for 120 epochs involving the learning of 5 parameters. Owing to the learning time for a model exceeding the resources of available computers, parallel GPUs will be required to search models in the future.

In addition, time series modeling using 3D CNN appears to overfit the learning data when input data, such as the duration, is increased. Therefore, overfitting should be avoided by introducing a generalization technique, and this can be considered for future studies.

Validation and testing for 3DCNN

The 600,000 training data images were divided into training data (80%) and validation data (20%). In partitioning the training data, earthquakes involved in validation were excluded from the testing in the same manner as those in 2DCNN (Fig. 2). The network parameters were optimized for estimating the hypocentral parameters by minimizing discrepancies in the training data set. After the optimization, these were verified using the validation data set. One network was used for the estimation of five hypocentral parameters. Each seismic wave propagation image requires hypocentral parameters output, so that the evolution of the estimates can be observed as the images are sorted. The testing of each hypocentral parameter during the 3D CNN validation process is shown in Fig. 5. The hypocentral parameter values obtained after 8 s duration were considered the estimates. The errors associated with the estimated values are comparable to those from the validation data set, and the calculated magnitudes for all earthquakes used for validation are presented in Table 1. According to the results, the estimates from the 3D CNN exhibit higher precision compared to those from the ordinary CNN. This indicates that seismic wave propagation evolution images are important for estimating hypocentral parameters using neural networks.

Generalization for real data

Subsequently, we attempted to generalize the developed neural network for real seismographic data. Seismograms recorded in stations operated by the Hot
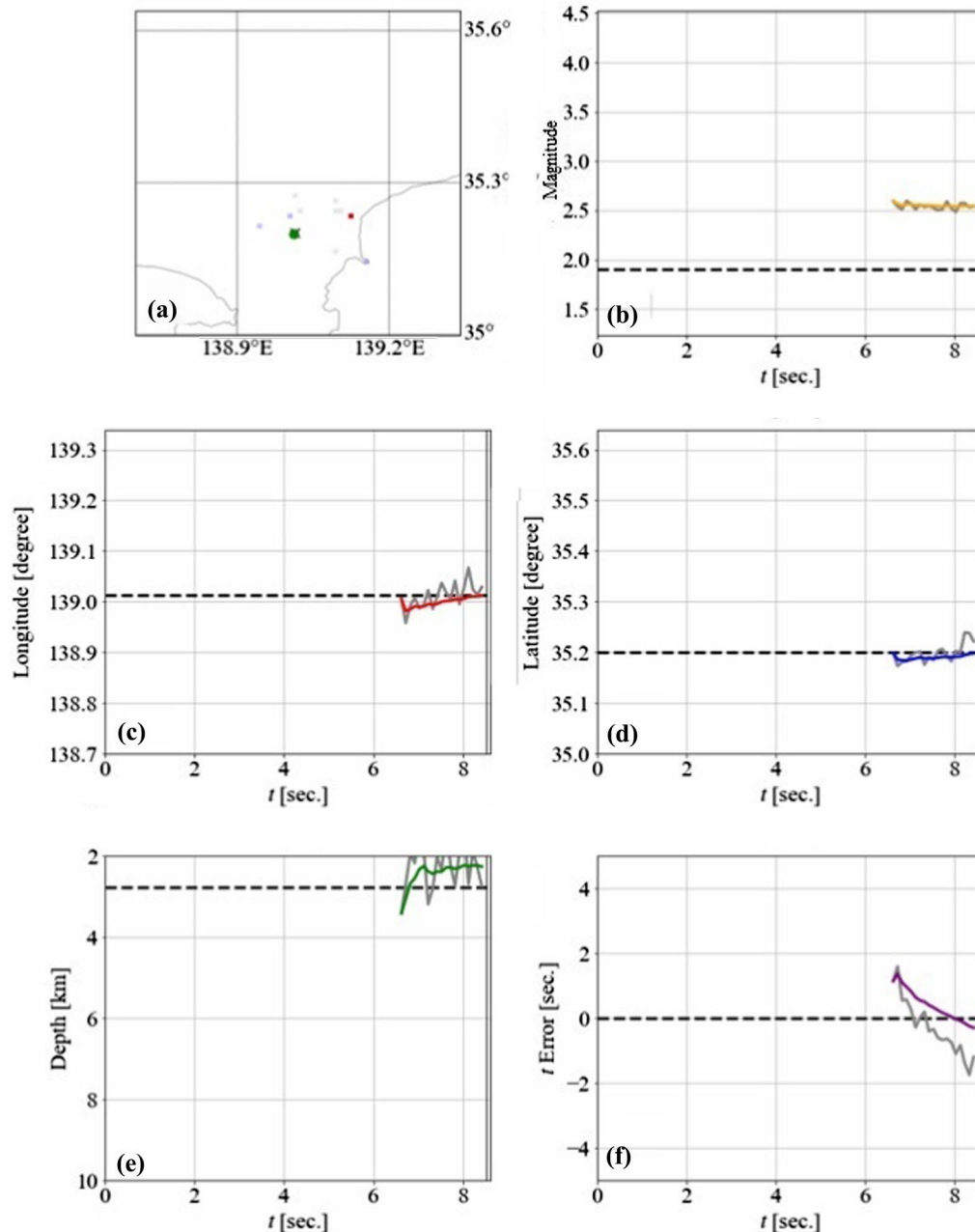
Sugiyama *et al. Earth, Planets and Space*   (2021) 73:135

Page 6 of 12

Fig. 5 Examples of hypocenter estimation in the verification phase for one earthquake (top left, **a**) Seismic wave propagation images starting at approximately 6 s after the origin time. The time evolution of the estimated hypocenter parameters is shown for magnitude (top right, **b**), longitude (middle left, **c**), latitude (middle right, **d**), depth (bottom left, **e**) and origin time (bottom right, **f**). The horizontal axis of each panel represents the time in seconds, and the estimation starts at approximately 6 s after the origin time when the absolute amplitude exceeds 1.0e−7 m/s. The dotted line in each figure represents the parameter to be estimated by each snapshot. The gray line in each figure shows the value estimated using each snapshot. The colored line shows the average of the estimated value up to each time instant. It should be noted that in the verification of the latitude, shown in **d**, the estimated value is in agreement with the actual value, and the overlap of the three lines indicates that no differences exist.

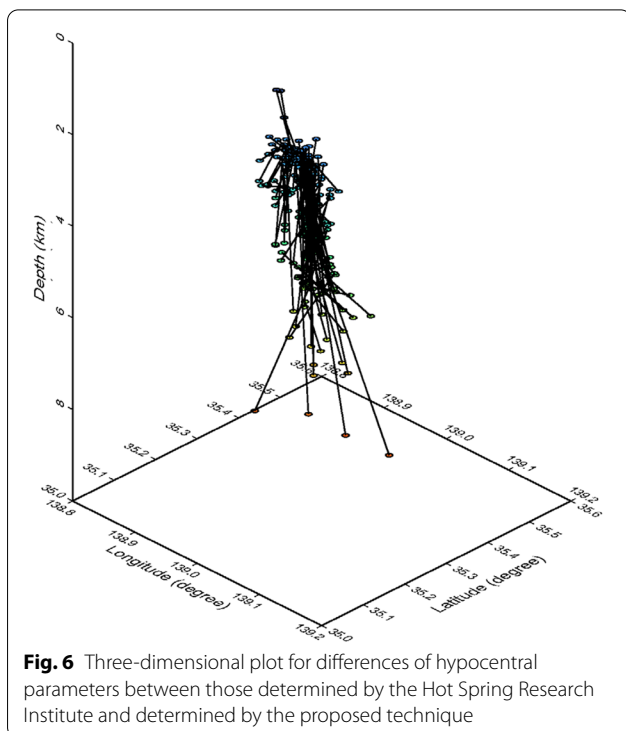Springs Research Institute of the Kanagawa Prefecture, the National Research Institute for Earth Science and Disaster Resilience, and Japan Meteorological Agency were obtained. The data were examined and 159 earthquakes with magnitude of approximately 2.0, which occurred in the Hakone region, Japan, between 2015

Sugiyama *et al. Earth, Planets and Space* (2021) 73:135

Page 7 of 12

and 2019 were selected. The data obtained from the stations were 1 min three-components velocity seismograms associated with the earthquakes. The original 100/200 Hz sampled accelerograms were converted to velocity seismograms, decimated to 10 Hz, and utilized for generating a 32×32 image for each time step. Since the accuracy of the synthetic seismograms is 5.6 s, 10 images per second represents an appropriate number of images for estimating the hypocentral parameters. The 32×32 images were employed as inputs for neural network to estimate the parameters. The RMSE between the estimated parameters and those reported by the Hot Springs Research Institute of the Kanagawa Prefecture for these earthquakes are presented in Table 1. According to comparisons of the results in Figs. 6 and 7, the proposed neural network provides good results. However, for several earthquakes, significant discrepancies emerged between the results from our method and those reported by the Hot Springs Research Institute of the Kanagawa Prefecture. These discrepancies are likely because some stations within the seismic network were impacted by noise, as discussed in the next section. The data in Table 1 show that estimates of the hypocenter depth involve higher errors compared to those of the epicenter. This is probably because the learning data set for the epicenter was larger than that for the hypocenter depth in the learning stage. Therefore, all possible parameter ranges must be included in the learning stage for the

accurate determination of hypocentral parameters, especially for shallow earthquakes as in the case studied. This can be achieved through enlargement of the training data set by computing theoretical seismograms for all possible hypocentral parameters. In addition, Fig. 7 reveals slight differences in the magnitudes. This is because the magnitude estimated using our method is the moment magnitude, which is based on the scalar moment M0 of the earthquake, whereas the magnitude reported by the Hot Spring Research Institute is the micro earthquake magnitude, which is associated with the maximum amplitude and the hypocentral distance of the seismic wave (Watanabe 1971). Therefore, the scale of the differences in magnitude between the two sets of results accounts for the shift. The results also suggest that hypocentral parameters can be estimated using images of seismic wave propagation evolution without considering the arrival times of the seismic waves at observatories. Comparison of the estimated errors between the observation and simulation data are shown in Fig. 8, including the regional dependence. Apparently, the accuracies of the hypocentral parameters exhibit no dependence.

Effect of noise

The proposed method was generalized to estimate hypocentral parameters using actual seismograms. Considering that actual seismograms involve different types of noise, examining their effects on the estimates is necessary. During the generalization of our method for actual seismograms, some stations were noted to be associated with higher noise than others. The noise linked to such stations may adversely affect the hypocentral parameters estimation and account for the significant differences between our results and those reported by the Hot Spring Research Institute (Fig. 9). To address this problem, we introduced cutout (DeVries and Taylor 2017), and created the neural network by randomly selecting at most 3 stations and setting their amplitudes as 0. This procedure is also applicable for temporarily shut stations with no results. The hypocentral parameters were then estimated using the network for the test data set and the observed data, and the results are presented in Table 1. The results show that the hypocentral parameters estimates improved slightly. In general, adding noise as a data augmentation process during the learning process may improve the network performance. Therefore, including physical processes causing noise, such as the scatter associated with the heterogeneous structure of the crust, or the ambiguity of earthquake source parameters using seismic wave propagation simulation to improve the generalization technique may be necessary, and this can be investigated in future.
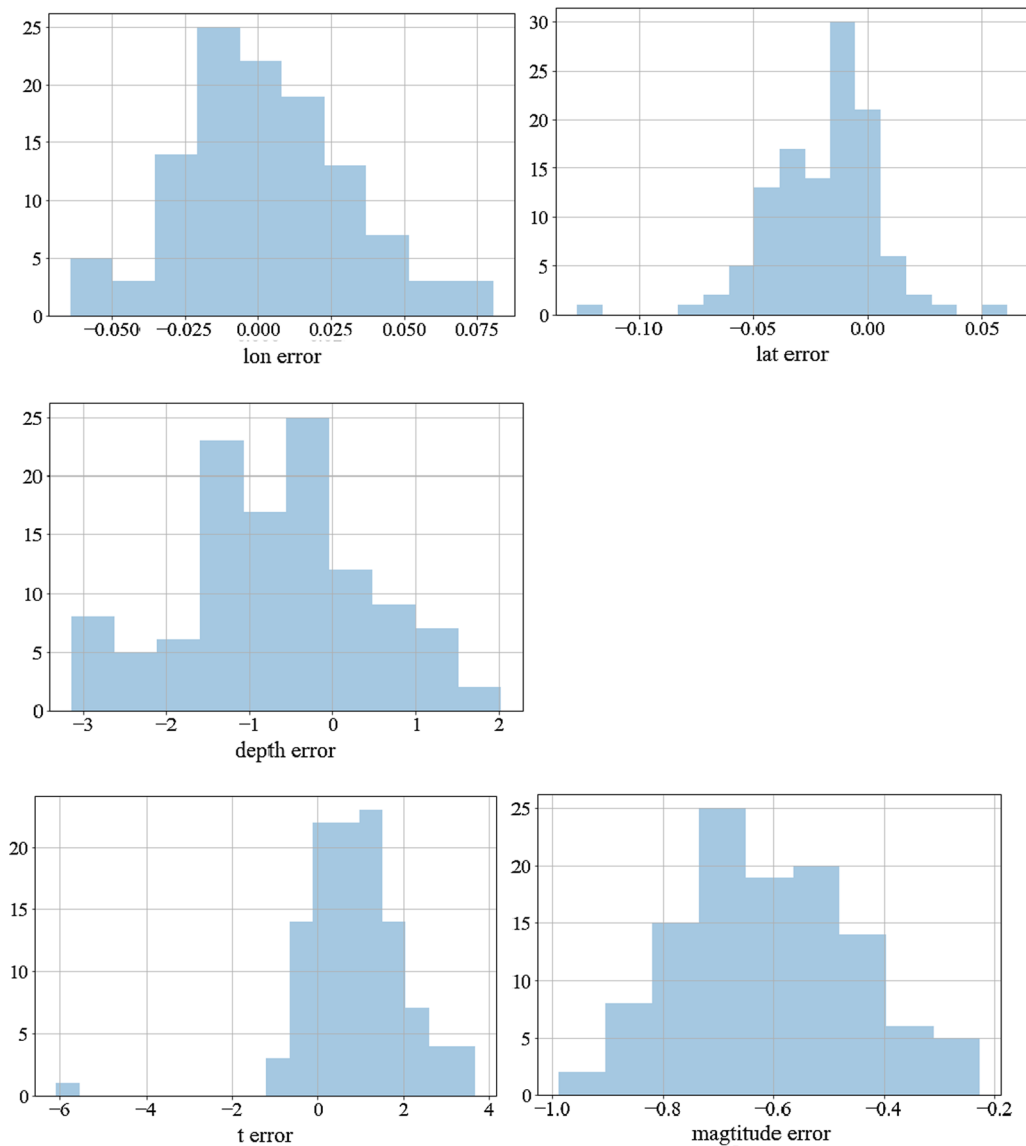

Fig. 6 Three-dimensional plot for differences of hypocentral parameters between those determined by the Hot Spring Research Institute and determined by the proposed technique

Fig. 7 Histograms of errors for hypocentral parameters determined by the proposed technique

Discussion

The advantage of using our technique in real time hypocenter determination system is its speed. The learning procedure for creating a neural network is time-intensive, especially if a large training data set is involved. However, after establishing the neural network, the hypocentral parameters are efficiently estimated within 0.1 s from each seismic wave propagation image using GPGPUs, even if multiple seismic stations are involved. Considering that the creation of images from seismic stations requires little time, efficient hypocenter determination is possible through the implementation of our technique in a local seismic network. In the present study,

the proposed technique is tested in the Hakone volcanic region in Japan, where earthquake swarms closely linked to volcanic activity are frequent. Regarding earthquake swarms, in 2015, for example, numerous micro earthquakes occurred in the region within a short duration. Monitoring such activities using conventional automated hypocenter location software is commonly challenging. In the present study, we demonstrate the application of the proposed technique for accurate and rapid micro earthquake monitoring in the Hakone volcanic region in Japan.

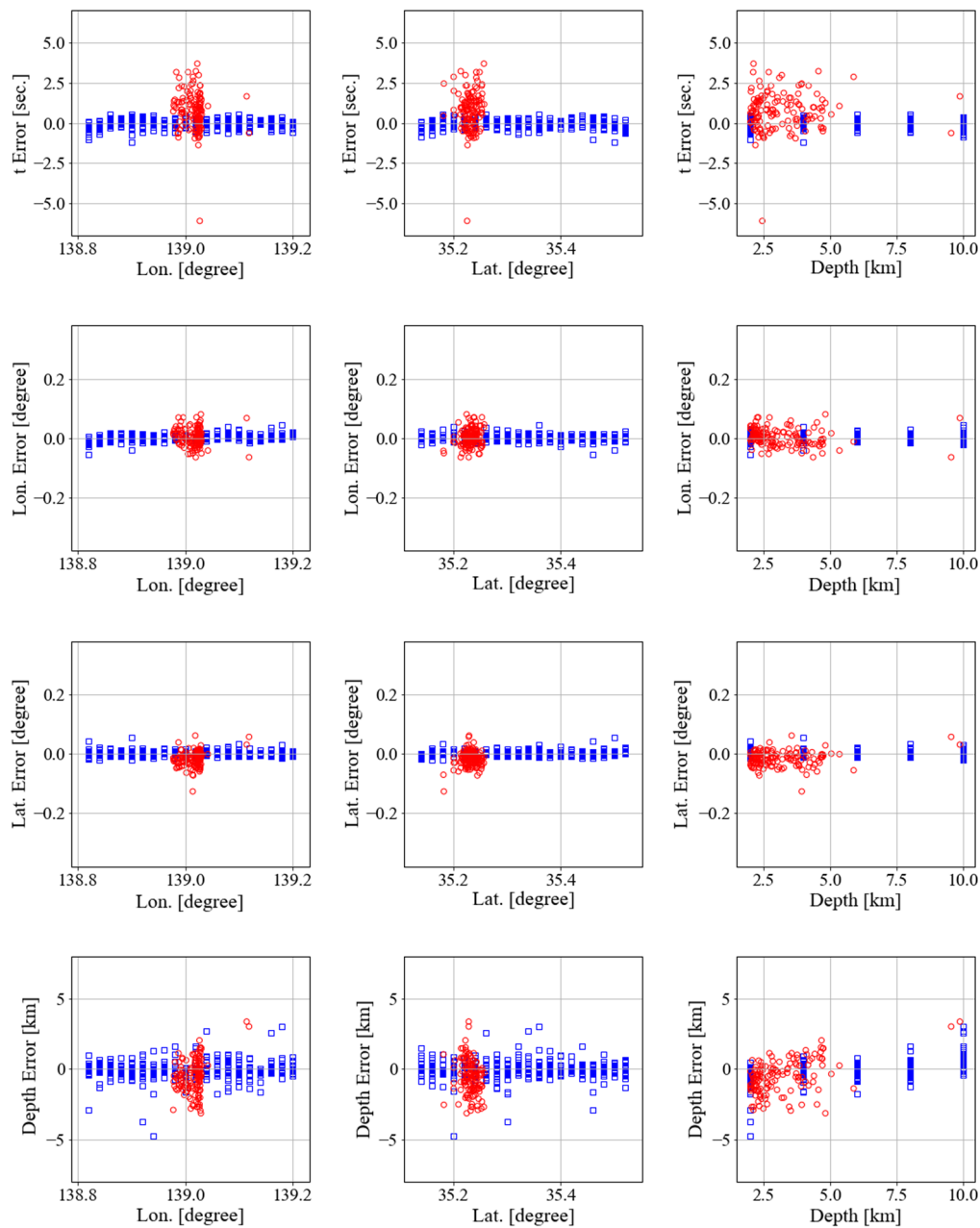Another advantage of our technique is the magnitude determination. The magnitude range can be set

Sugiyama *et al. Earth, Planets and Space*   (2021) 73:135

Page 9 of 12

Fig. 8 Distribution of errors for hypocentral parameters determined by the proposed technique. Each column shows geographical distribution of errors for hypocentral parameters. Each row shows errors of origin time, longitude, latitude and depth, from top to bottom. Red symbols show observation and blue symbols show simulation

arbitrarily if the training data are generated from theoretical seismograms. Therefore, any moment tensor solution can be employed for producing the theoretical seismograms and generating the training data set. This enables the incorporation of high magnitude events, which rarely occur. The lower limit for magnitude determination using this technique will be determined when the system

is implemented in a local seismic network and the *S*/*N* ratio is obtained by monitoring seismic activity. As previously mentioned, the images used for the training data set are generated via modeling of seismic wave propagation using a 3D Earth model and the moment tensor solutions. Therefore, the magnitude estimation is also expected to be accurate when applied to a real seismic
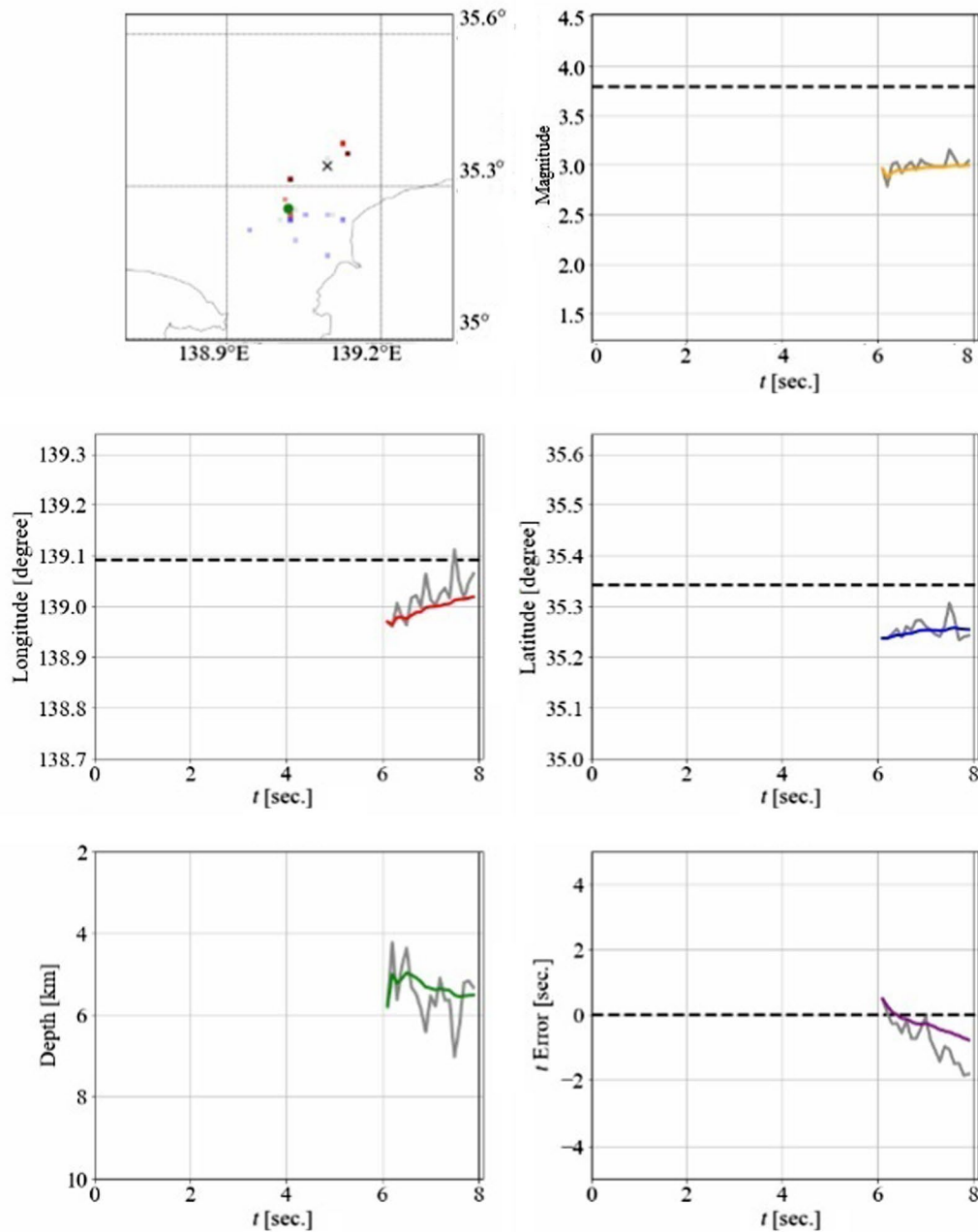
Sugiyama *et al. Earth, Planets and Space*   (2021) 73:135

Page 10 of 12

Fig. 9 Same as Fig. 5 but for noisy case. For details of figures see caption to Fig. 5

network. Although the examples are of limited scope, the approach exhibits potential for real time monitoring of seismic activity.

The performance of our proposed technique depends on the spatial grid size of assumed hypocenters for computing synthetic seismograms to be used as training data. We have tried to locate earthquake hypocenter using larger grid size to check how the grid size of the learning

data set affects the accuracy of the hypocenteral parameters and have shown the results in Additional file 1: Table S1. We have doubled the latitude grid size of the hypocenter in the learning data and found that the RMS errors for hypocentral parameters become worse for both simulation and observation data. This result demonstrates that the accuracy of the estimated hypocentral parameters depends on the grid size of the learning

Sugiyama *et al. Earth, Planets and Space*   (2021) 73:135

Page 11 of 12

data set. We may get much higher performance if we use much finer grid size. We should point out, however, that there still exists a limitation in the accuracy of Earth model to be used for the computation of synthetic seismograms. Even though we may use finer grid size for the selection of hypocenter to be used as training data set, it may be difficult to get enough accuracy in the Earth structure model. We may not be able to get accurate three-dimensional Earth structure model with the resolution of kilometer even for local seismic network scale, as we need to use seismic waves to get three-dimensional model. Therefore, the performance of the proposed approach is controlled by the accuracy of the three-dimensional Earth model. This is the same situation as that of conventional approach to locate earthquake hypocenter based on the arrival times of P and S waves. Our approach may be better than the conventional approach, because our approach handles seismic wave propagation rigorously even for three-dimensional Earth structure.

Summary

In the present study, we established neural networks for determining hypocentral parameters such as the epicenter, depth, occurrence time, and magnitude using synthetic seismograms as a training data set. The theoretical seismograms were synthesized based on a realistic 3D Earth model, and the seismograms were exploited to produce spatial images for seismic wave propagation at the Earth's surface. The generalized neural network results for obtaining actual seismic data indicated potential for estimating earthquake hypocentral parameters using CNN trained via theoretical seismograms. Although more events are required to validate generalizing the proposed neural network, this framework is suitable for determining hypocentral parameters in a local seismic network, and has potential for seismic and volcanic activity monitoring.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40623-021-01461-w.

Additional file 1: Table S1. Comparisons of Root Mean Square Errors (RMSE) of hypocentral parameters estimated by 3DCNN and 3DCNN with the latitude grid size of the learning data set to be doubled.

Authors' contributions
All authors contributed equally to this research. DS: conceptualization, methodology, data curation and visualization, writing; ST: methodology, investigation, software, writing and editing; YY: methodology, data curation and visualization, writing. All authors read and approved the final manuscript.

Availability of data and materials
The seismic waveforms utilized were provided by the Hot Spring Research Institute of the Kanagawa Prefecture, the National Research Institute for Earth Science and Disaster Resilience, and Japan Meteorological Agency. The open-source program package SPECFEM3D from the Computational Infrastructure for Geodynamics (CIG; geodynamics.org) was employed for computation, and the GMT software (Wessel and Smith 1998) was used to generate many figures presented.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
There is no competing interests.

Author details
[1]Research Institute for Value-Added-Information Generation, Japan Agency for Marine-Earth Science and Technology, Kanazawa-ku, Yokohama, Kanagawa 236-0001, Japan. [2]Hot Springs Research Institute of Kanagawa Prefecture, Odawara, Kanagawa, Japan. [3]Present Address: Earthquake Research Institute, University of Tokyo, Tokyo, Japan.

References
Carreira J, Zisserman A (2017) Quo Vadis, action recognition? A new model and the kinetics dataset. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). pp 4724–4733

Carrington L, Komatitsch D, Laurenzano M, Tikir M M, Michéab D, Le Goff N, Snavely A, Tromp J (2008) High-frequency simulations of global seismic wave propagation using SPECFEM3D GLOBE on 62K processors. In: SC'08: proceedings of the 2008 ACM/IEEE conference on supercomputing. https://doi.org/10.1109/SC.2008.5215501

DeVries T, Taylor GW (2017) Improved regularization of convolutional neural networks with cutout. arxiv:1708.04552v2

Donahue J, Hendricks LA, Guadarrama S, Rohrbach M, Venugopalan S, Saenko K, Darrell T (2015) Long-term recurrent convolutional networks for visual recognition and description. In: The IEEE conference on computer vision and pattern recognition (CVPR). pp 2625–2634

Faccioli E, Maggio F, Paolucci R, Quarteroni A (1997) 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method. J Seismol 1:237–251

Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 32nd international conference on machine learning, Lille Grand Palais, Lille, 6–11 July 2015

Käufl P, Valentine AP, O'Toole TB, Trampert J (2014) A framework for fast probabilistic centroid-moment-tensor determination–inversion of regional static displacement measurements. Geophys J Int 196:1676–1693

Käufl P, Valentine AP, de Wit R, Trampert J (2015) Robust and fast probabilistic source parameter estimation from near-field displacement waveforms using pattern recognition. Bull Seismol Soc Am 105:2299–2312

Käufl P, Valentine AP, de Wit R, Trampert J (2016a) Solving probabilistic inverse problems rapidly with prior samples. Geophys J Int 205:1710–1728

Käufl P, Valentine AP, Trampert J (2016b) Probabilistic point source inversion of strong-motion data in 3-D media using pattern recognition: a case study for the 2008 Mw 5.4 Chino Hills earthquake. Geophys Res Lett 43:8492–8498. https://doi.org/10.1002/2016GL069887

Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In: Proceedings of the 3rd international conference on learning representations (ICLR2015), Hilton San Diego Resort & Spa, Sandiego, 7–9 May 2015

Komatitsch D (1997) Spectral and spectral-element methods for the 2D and 3D elastodynamics equations in heterogeneous media. Dissertation, Institut de Physique du Globe, Paris, France

Komatitsch D, Tromp J (2002a) Spectral-element simulations of global seismic wave propagation-I. Validation. Geophys J Int 149:390–412

Komatitsch D, Tromp J (2002b) Spectral-element simulations of global seismic wave propagation-II. Three-dimensional models, oceans, rotation, and self-gravitation. Geophys J Int 150:303–318

Komatitsch D, Vilotte JP (1998) The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. Bull Seismol Soc Am 88:368–392

Komatitsch D, Tsuboi S, Tromp J (2005) The spectral-element in seismology. Seismic earth: array analysis of broadband seismograms, vol 157. Geophysical monograph. American Geophysical Union, Washington, D.C, pp 205–227

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv NIPS 25:1097–1105

LeCun Y, Haffner P, Bottou L, Bengio Y (1999) Object recognition with gradient-based learning. Shape, contour and grouping in computer vision, vol 1681. Lecture notes in computer science. Springer, Berlin

Maggi A, Ferrazzini V, Hibert C, Beauducel F, Boissier P, Amemoutou A (2017) Implementation of a multistation approach for automated event classification at piton de la fournaise volcano. Seismol Res Lett 88:878–891

Malfante M, Mura MD, Métaxian J-P, Mars J, Macedo O, Inza A (2018) Machine learning for volcano-seismic signals: challenges and perspectives. IEEE Signal Process Mag 35:20–30

Nakano M, Sugiyama D, Hori T, Kuwatani T, Tsuboi S (2019) Discrimination of seismic signals from earthquakes and tectonic tremor by applying a convolutional neural network to running spectral images. Seismol Res Lett 90:530–538

Onishi R, Sugiyama D (2017) Deep convolutional neural network for cloud coverage estimation from snapshot camera images. SOLA 13:235–239. https://doi.org/10.2151/sola.2017-043

Perol T, Gharbi M, Denolle M (2018) Convolutional neural network for earthquake detection and location. Sci Adv 4:e1700578

Rietmann M, Messmer P, Nissen-Meyer T, Peter D, Basini P, Komatitsch D, Schenk O, Tromp J, Boschi L, Giardini D (2012) Forward and adjoint simulations of seismic wave propagation on emerging large-scale GPU architectures. In: Proceedings of the international conference on high performance computing, networking, storage and analysis (SC12). IEEE Computer Society Press, Los Alamitos, CA, USA, Article 38

Ross ZE, Meier M-A, Hauksson E, Heaton TH (2018) Generalized seismic phase detection with deep learning. Bull Seismol Soc Am. https://doi.org/10.1785/0120180080

Rouet-Leduc B, Hulbert C, Lubbers N, Barros K, Humphreys CJ, Johnson PA (2017) Machine learning predicts laboratory earthquakes. Geophys Res Lett. https://doi.org/10.1002/2017GL074677

Seydoux L, Balestriero R, Poli P, de Hoop M, Campillo M, Baraniuk R (2020) Clustering earthquake signals and background noises in continuous seismic data with unsupervised deep learning. Nat Commun 11:3972. https://doi.org/10.1038/s41467-020-17841-x

Shahriari B, Swersky K, Wang Z, Adams RP, de Freitas N (2015) Taking the human out of the loop: a review of Bayesian optimization. Proc IEEE 104:148–175. https://doi.org/10.1109/JPROC.2015.2494218

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint. arXiv:1409.1556

Szegedy C, Liu W, Jia Y, Sermanet P, Reed SE, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proc. IEEE conference on computer vision and pattern recognition (CVPR). pp 1–9

Tran D, Bourdev L, Fergus R, Torresani L, Paluri M (2015) Learning spatiotemporal features with 3D convolutional networks. In: The IEEE international conference on computer vision (ICCV). pp 4489–4497

Tsuboi S, Komatitsch D, Ji C, Tromp J (2003) Broadband modeling of the 2002 Denali fault earthquake on the Earth simulator. Phys Earth Planet Inter 139:305–312

Tsuboi S, Ando K, Miyoshi T, Peter D, Komatitsch D, Tromp J (2016) A 1.8 trillion degrees of freedom, 1.24 petaflops global seismic wave simulation on the K computer. Int J High Perform Comput Appl 30:411–422. https://doi.org/10.1177/1094342016632596

Vaillant R, Monrocq C, Cun YL (1994) Original approach for the localization of objects in images. IEE Proc vis Image Signal Process 141:245–250

Wang J, Teng T-L (1995) Artificial neural network-based seismic detector. Bull Seismol Soc Am 85:308–319

Wang J, Teng TL (1997) Identification and picking of S phase using an artificial neural network. Bull Seismol Soc Am 87:1140–1149

Watanabe H (1971) Determination of earthquake magnitude at regional distance in and near Japan. Zisin 24:189–200. https://doi.org/10.4294/zisin1948.24.3_189

Wessel P, Smith WH (1998) New, improved version of the generic mapping tools released. EOS Trans AGU 79:579–579. https://doi.org/10.1029/90EO00426

Yukutake Y, Honda R, Harada M, Arai R, Matsubara M (2015) A magma-hydrothermal system beneath Hakone volcano, central Japan, revealed by highly resolved velocity structures. J Geophys Res Solid Earth 120:3293–3308. https://doi.org/10.1002/2014JB011856

Zhu W, Beroza GC (2019) PhaseNet: a deep-neural-network-based seismic arrival-time picking method. Geophys J Int 216:261–273

Publisher's Note