

Graph-partitioning based convolutional neural network for earthquake detection using a seismic array

Keisuke Yano^{1,1}, Takahiro Shiina^{2,2}, Sumito Kurata^{3,3}, Aitaro Kato^{4,4}, Fumiyasu Komaki^{5,5}, Shin'ichi Sakai^{5,5}, and Naoshi Hirata^{5,5}

¹The institute of Statistical Mathematics

²Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology

³The University of Tokyo

⁴Earthquake Research Institute, The University of Tokyo

⁵University of Tokyo

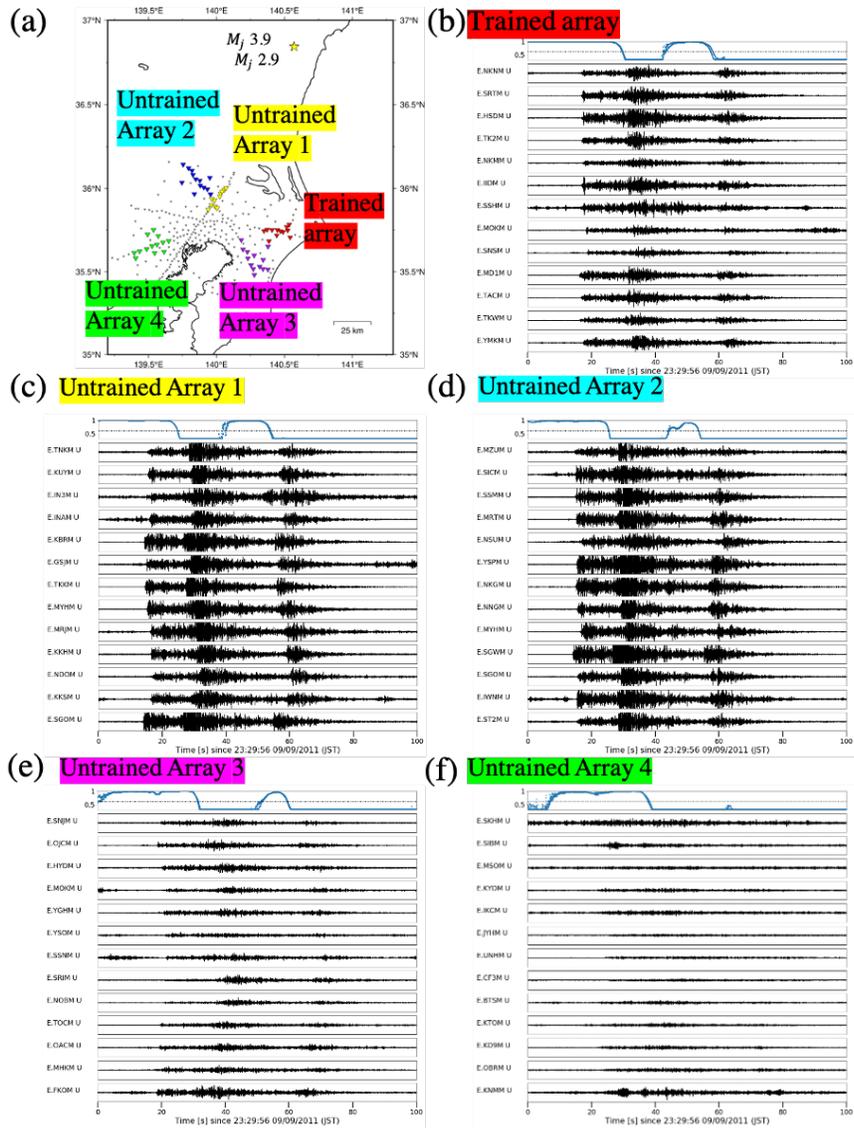
November 30, 2022

Abstract

We present a deep-learning approach for earthquake detection using waveforms from a seismic array consisting of multiple seismographs. Although automated, deep-learning earthquake detection techniques have recently been developed at the single-station level, they have potential difficulty in reducing false detections owing to the presence of local noise inherent to each station. Here, we propose a deep-learning-based approach to efficiently analyze the waveforms observed by a seismic array, whereby we employ convolutional neural networks in conjunction with graph partitioning to group the waveforms from seismic stations within the array. We then apply the proposed method to waveform data recorded by a dense, local seismic array in the regional seismograph network around the Tokyo metropolitan area, Japan. Our method detects more than 97% percent of the local seismicity catalogue, with less than 4% percent false positive rate, based on an optimal threshold value of the output earthquake probability of 0.61. A comparison with conventional deep-learning-based detectors demonstrates that our method yields fewer false detections for a given true earthquake detection rate. Furthermore, the current method exhibits the robustness to poor-quality data and/or data that are missing at several stations within the array. Synthetic tests demonstrate that the present method has the potential to detect earthquakes even when half of the normally available seismic data are missing. We apply the proposed method to analyze 1-hour-long continuous waveforms and identify new seismic events with extremely low signal-to-noise ratios that are not listed in existing catalogs. (241words)

Hosted file

multistationcnn_revisejgr_submit_essoar.docx available at <https://authorea.com/users/539552/articles/607953-graph-partitioning-based-convolutional-neural-network-for-earthquake-detection-using-a-seismic-array>



Graph-partitioning based convolutional neural network for earthquake detection using a seismic array

Keisuke Yano¹, Takahiro Shiina², Sumito Kurata³, Aitaro Kato⁴,
Fumiyasu Komaki^{3,5}, Shin'ichi Sakai^{4,6}, Naoshi Hirata^{4,7}

¹The Institute of Statistical Mathematics, 10-3 Midori cho, Tachikawa City, Tokyo 190-8562, Japan.

²Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Tsukuba City, Ibaraki 305-8567, Japan.

³Graduate School of Information Science and Technology, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.

⁴ Earthquake Research Institute, the University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo, 113-0032, Japan.

⁵RIKEN Center for Brain Science, 2-1 Hirosawa, Wako City, Saitama 351-0198, Japan.

⁶Interfaculty Initiative in Information Studies, the University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan.

⁷National Research Institute for Earth Science and Disaster Resilience, 3-1 Tennodai, Tsukuba City, Ibaraki, 305-0006, Japan.

Key Points:

- We developed a convolutional neural network (CNN) for automated earthquake detections using waveforms from a non-equispaced seismic array. (138 characters)
- Stations with similar waveforms are clustered via graph partitioning, substantially improving the CNN-based earthquake detection accuracy. (138 characters)
- The proposed CNN detected small-magnitude earthquakes with very low signal-to-noise ratios that were not listed in existing catalogs. (133 characters)

Corresponding author: Keisuke Yano, yano@ism.ac.jp

Abstract

We present a deep-learning approach for earthquake detection using waveforms from a seismic array consisting of multiple seismographs. Although automated, deep-learning earthquake detection techniques have recently been developed at the single-station level, they have potential difficulty in reducing false detections owing to the presence of local noise inherent to each station. Here, we propose a deep-learning-based approach to efficiently analyze the waveforms observed by a seismic array, whereby we employ convolutional neural networks in conjunction with graph partitioning to group the waveforms from seismic stations within the array. We then apply the proposed method to waveform data recorded by a dense, local seismic array in the regional seismograph network around the Tokyo metropolitan area, Japan. Our method detects more than 97 percent of the local seismicity catalogue, with less than 4 percent false positive rate, based on an optimal threshold value of the output earthquake probability of 0.61. A comparison with conventional deep-learning-based detectors demonstrates that our method yields fewer false detections for a given true earthquake detection rate. Furthermore, the current method exhibits the robustness to poor-quality data and/or data that are missing at several stations within the array. Synthetic tests demonstrate that the present method has the potential to detect earthquakes even when half of the normally available seismic data are missing. We apply the proposed method to analyze 1-hour-long continuous waveforms and identify new seismic events with extremely low signal-to-noise ratios that are not listed in existing catalogs. (241words)

1 Introduction

Enormous volumes of continuous seismic data have been acquired from seismograph networks over the past decade, with these datasets consisting of observations from multiple seismic stations. Dense seismograph networks, such as the Japanese Metropolitan Seismic Observation network (MeSO-net) and the Southern California Seismic Network, monitor real-time seismicity and provide continuous waveforms from their respective network stations. Efficient and thorough analyses of these datasets should be of great benefit to seismology. The main objective of the present work, which represents a novel approach to and advance in seismic data analysis, is the development of an improved earthquake detection technique for these massive seismic network datasets (Bergan & Beroza, 2018; Li et al., 2018; Peng & Zhao, 2009).

Deep neural networks have been attracting increasing interest as tools for analyzing seismic big data owing to their ability to handle such massive data volumes and improve data processing performance. Convolutional neural networks (CNNs) are a deep neural network architecture that is often used in seismic analyses. CNNs have been utilized for picking P-wave arrival times and discriminating first-motion polarities (Hara et al., 2019; Ross et al., 2018; Wang et al., 2019), detecting earthquakes and their localization (Perol et al., 2018), picking both compressional- and shear-wave (P- and S-wave) arrival times (Zhu & Beroza, 2018), and discriminating between usual earthquakes and tectonic tremors (Nakano et al., 2019). These studies have shed light on the effectiveness of employing CNNs to analyze large seismic datasets. Many other deep-learning-based studies have also been developed for various seismic analyses (Kong et al., 2019).

Here, we present a new CNN-based method for earthquake detection using waveform data recorded in a seismic array. Earthquake detection at the single-station level has a potential inability to prevent false detections, because each station produces its own local noise that originates from the near-surface environment. The use of multiple seismic stations can help reduce false detections. Waveform similarity searches, such as template matching and auto-correlation approaches, have shown that simultaneous detections at multiple stations can be used to identify very weak seismic signals and reduce false positives (Bergan

77 & Beroza, 2018; Peng & Zhao, 2009). Our goal here is to develop CNNs for earthquake
 78 detection using seismographs recorded by a local seismic array.

79 However, we need to carefully assess CNN approaches to waveforms from a local seismic
 80 array, as CNN inputs are designed to be equispaced (see Chapter 9 in Goodfellow et al.
 81 (2016)), whereas the actual station distribution in an array is usually non-equispaced. We
 82 employ a graph-partitioning method, modularity maximization, to accommodate such a
 83 non-equispaced distribution by treating the seismograph network as a graph, with each node
 84 representing a seismic station in the array. Modularity maximization is a node-clustering
 85 algorithm that automatically produces sub-graphs from a given graph. We first group the
 86 seismic stations into clusters via modularity maximization and then stack the waveforms,
 87 which are convolved with learned filters in CNNs, from each partition. This step is regarded
 88 as average pooling with respect to the seismic stations. In conjunction with the max
 89 pooling with respect to time, this step reduces effective parameters in CNNs and leads
 90 to the improvement of the performance.

91 We apply our method to waveform data recorded by MeSO-net (Hirata et al., 2009;
 92 Kano et al., 2017; Kasahara et al., 2009; Sakai & Hirata, 2009), with a focus on the
 93 earthquakes that occurred in the Kanto district, Japan, from 4 to 16 September 2011.
 94 MeSO-net is an online high-density seismic network that consists of approximately 300
 95 accelerometers installed in and around the metropolitan area in the Kanto district (Figure
 96 1). We adopt the 13 seismic stations that are located approximately in the center of the
 97 seismic activity around the Kanto district as the seismic array for our analysis. We compare
 98 our method with CNNs for a single station and a seismic array consisting of multiple stations
 99 without any pooling with respect to the stations. This comparison illustrates that our
 100 method markedly improves the detection accuracy of the seismic signals, particularly those
 101 with a low signal-to-noise ratio (SNR). We also demonstrate the robustness of our method for
 102 analyzing incomplete seismic records (e.g., due to sensor problems) and temporary increases
 103 in noise levels at several stations.

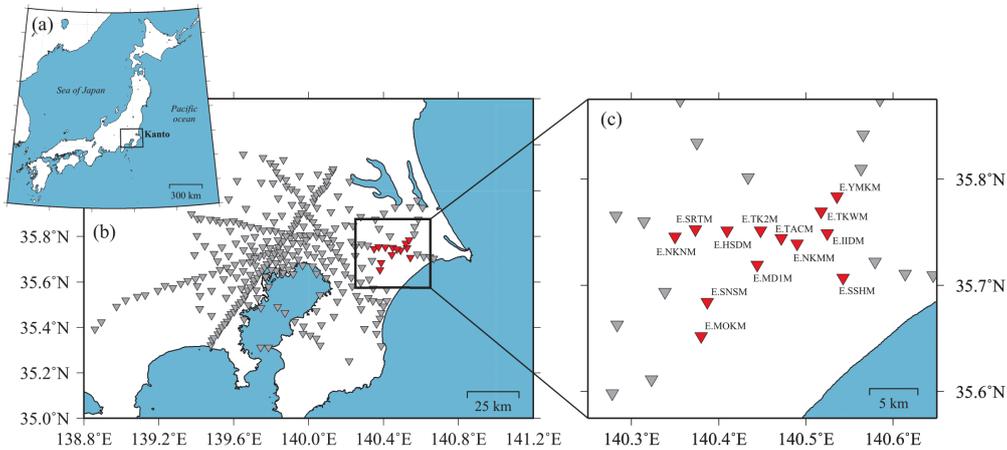


Figure 1. (a) Map of the study area, the Kanto district, relative to Japan. (b) Locations of the MeSO-net stations (gray and red inverted triangles) in the Kanto district. (c) The 13-station seismic array (red inverted triangles) used in this study.

104 We note that CNNs have already been applied to waveform records in seismic networks,
 105 although not for earthquake detection (Kriegerowski et al., 2019; McBrearty et al., 2019;
 106 Zhang et al., 2020). Our idea is not task specific and can therefore be combined with CNN
 107 architecture therein.

2 Methods

An outline of our method is presented in Figure 2. The input consists of two-dimensional (2D) arrays of the three-channel waveforms (UpDown, NorthSouth, and EastWest) recorded at the seismic stations within our local seismic array, where the horizontal axis of the 2D arrays is the elapsed time from a reference time before the P-wave onset, and the vertical axis corresponds to the station alignment. The output is the probability of the input data containing an earthquake record. Our method uses an undirected weighted graph that represents the geometry of the seismic stations within the array. The nodes of the graph represent the stations within the seismic array, and the weights of the edges represent the closeness of the stations, which are discussed below. We first introduce the two building blocks (CNN and graph partitioning) of our method, following which we present the resultant architecture that is a combination of these two blocks.

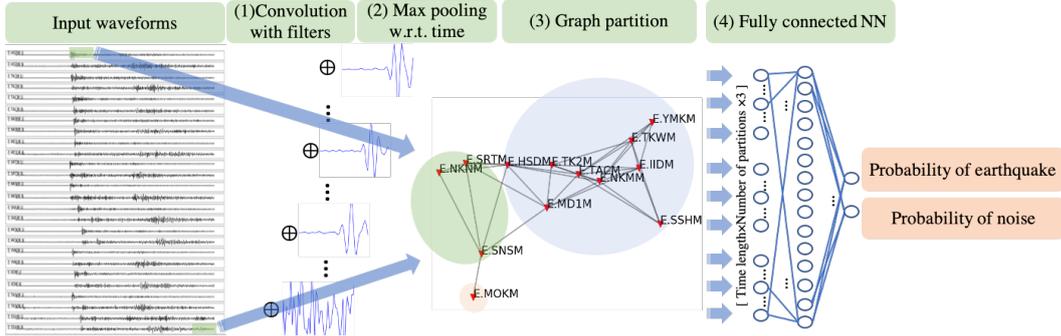


Figure 2. Schematic diagram of the resultant CNN architecture. The process is four-fold. The input consists of the three-component waveforms observed at multiple stations within the seismic array. (1) Convolution with learned filters; (2) Max pooling with respect to (w.r.t.) the time axis; (3) Average pooling with respect to the station axis via graph partitioning, with stations within the same partition lying within the same ellipse; and (4) Fully connected neural network (NN).

2.1 Convolutional neural networks

CNNs are one of the most powerful non-linear feature extractors and can learn a large number of digital filters in combination with fully connected neural networks. CNNs use these learned filters to also provide a function that maps an input (e.g., three-component 2D arrays that represent the observed waveforms from a seismic array) into an output value (e.g., the probability that the input data contains an earthquake).

CNNs build upon a series of layers that sequentially process the input data as follows: (a) convolve (or, more precisely, cross correlate) the input with a set of learnable filters; (b) downsample the convolution output; and (c) apply a non-linear transformation referred to as the “activation function” to the downsampled output. The first step (a) is termed the “convolution layer”: An input waveform $\{w_t : t = 1, \dots, T\}$ of length T and a learnable filter $\{h_t : t = 1, \dots, F\}$ of length F return a convolution layer that consists of the waveform convolved with the filter $\{\sum_{i=1}^F h_i w_{t+i} : t = 1, \dots, T\}$. Each filter $\{h_t : t = 1, \dots, F\}$ is learned during the training step (i.e., a learnable filter), and detects the presence of specific features in the original input. The second step (b) is called the pooling layer. This step drastically reduces the dimensionality, which is the number of parameters and computations in the network, to avoid overfitting caused by excess parameters. This step also introduces a degree of local translation invariance and robustness to local perturbations (see Section 9.3 in Goodfellow et al. (2016) and references therein). There are two common pooling methods:

139 max pooling and average pooling. Max pooling is a pooling operation that calculates the
 140 maximum value for each point of each waveform convolved with a learnable filter, whereas
 141 average pooling is a pooling operation that calculates the local average for each point of
 142 each waveform convolved with a learnable filter. These operations can be written in the
 143 following mathematical forms: For the waveform $\{\tilde{w}_t : t = 1, \dots, T\}$ convolved with a
 144 filter, max pooling with size P returns $\{\max\{\tilde{w}_{t+m} : m = 1, \dots, P\} : t = 1, \dots, T\}$, and
 145 average pooling with size P returns $\{\sum_{m=1}^P \tilde{w}_{t+m}/P : t = 1, \dots, T\}$. This study uses zero
 146 padding, where $\tilde{w}_{T+i} = 0$ for $i = 1, \dots, P$ to maintain the same output dimension of the
 147 pooling operation as its input dimension. These pooling operations have different roles: Max
 148 pooling highlights the sharp changes in the convolved waveform, whereas average pooling
 149 smooths these changes. The third step (c) is the activation layer, which makes the output
 150 non-linear and therefore enables the network to express complex features.

151 However, the challenge of accommodating a CNN to the spatial distribution (geometry)
 152 of the stations to boost the power of the CNN remains when applying the CNN to
 153 seismograms from a seismic array. Here, we take advantage of graph partitioning to address
 154 this challenge, as discussed below.

155 2.2 Graph partitioning

156 Graph partitioning is a way to reduce a graph to smaller sub-graphs by partitioning
 157 a set of nodes into mutually exclusive groups. Several methods for graph partitioning
 158 exist (Bichot & Siarry, 2013; Fortunato & Castellano, 2012). Here, we employ modularity
 159 maximization (Newman, 2006; Newman & Girvan, 2004), which finds a graph partition that
 160 maximizes the modularity score. The modularity score for a given graph partition is defined
 161 by:

$$162 \quad Q = \frac{1}{2m} \sum_{i,j} \left(A_{i,j} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j),$$

163 where $A = (A_{i,j})$ is the adjacency matrix of the graph, m is the number of edges of the
 164 graph, and k_i is the degree of node i , c_i is the partition to which node i belongs, and
 165 $\delta(a, b) = 1$ if $a = b$ and 0 otherwise. It is shown that $k_i k_j / 2m$ is the expected number
 166 of edges between nodes i and j under the assumption that a graph is randomly generated
 167 according to the configuration model. Therefore, the modularity score represents the number
 168 of edges within given partitions minus the expected number of randomly generated edges
 169 in an equivalent graph. Modularity maximization for graph partitioning stems from the
 170 idea that each partition within the “best” graph partition has edges that do not appear
 171 in the configuration model. This approach has been applied to a variety of research fields,
 172 including bioinformatics, brain science, and social science (Mill et al., 2008; Xia et al., 2013).

173 We use modularity maximization in the CNN pooling layer via the following process: (i)
 174 We first create a graph that represents the seismic array by defining the weighted undirected
 175 graph with nodes that correspond to the stations and an adjacency matrix:

$$176 \quad A_{i,j} := \exp(-\alpha d(i, j)) \text{ if } d(i, j) \leq L \text{ and } A_{i,j} := 0 \text{ otherwise,} \quad (1)$$

177 where $d(i, j)$ is the distance between stations i and j , α is a sensitivity parameter, and L
 178 is the effective distance parameter. We note that the determination of $d(i, j)$ is important.
 179 Here, we use the usual Euclidean distance between stations because this distance is generally
 180 robust for earthquake patterns that have various epicenters and ray-paths. (ii) We then
 181 apply modularity maximization to the weighted undirected graph to obtain the graph
 182 partition. (iii) We finally convolve the waveforms with learnable filters in the CNN at
 183 each station and stack the waveforms at the stations belonging to the same partition. We
 184 essentially conduct average pooling with respect to the station axis by using the partitions.

185 The core idea of this operation is that the waveforms convolved with a learned filter
 186 at the closer stations become similar to each other. Therefore, stacking (average pooling)
 187 along with graph partitioning can further improve the quality of the learned features. The

188 choices of d , α , and L may be important, but these can be chosen in a data-driven way, as
 189 discussed in the following subsection.

190 2.3 Resultant architecture

191 We propose the following architecture, which is a combination of the CNN and graph
 192 partitioning. We first convolve the input waveform at each station with filters, followed by
 193 waveform processing at each station via max pooling with respect to the time axis. We
 194 then conduct average pooling, along with graph partitioning (as described in the previous
 195 subsection) to the waveforms. We finally apply the fully connected neural network to the
 196 stacked waveforms and obtain the output probability.

197 In this study, we specified the hyperparameters in the CNN as follows. The number
 198 of learnable filters in the convolution layer was 30, with each being 1 second in length.
 199 The activation function was a rectified linear unit, $x \mapsto \max\{x, 0\}$. The length of max
 200 pooling was 0.25 second. The number of units in the fully connected layer was 40. The
 201 batch size during the training step was 20. The filters in the CNN were optimized during
 202 the training step that was conducted by minimizing the cross-entropy loss using stochastic
 203 gradient descent optimization with momentum, where the learning rate was 0.005 and the
 204 momentum was 0.9. The cross-entropy loss is a function that maps a pair of true label y
 205 ($y = 1$ for earthquake and $y = 0$ for noise) and output probability p for an earthquake into
 206 $-y \log p - (1 - y) \log(1 - p)$. Stochastic gradient descent optimization with momentum uses
 207 a linear combination of the gradient multiplied by the learning rate and the previous update
 208 multiplied by the momentum as the next update. The l_1 penalty with the regularization
 209 parameter (1.0) was added to the loss.

210 The hyperparameters α and L in Equation (1) were determined via hold-out validation.
 211 We set the candidates for α and L to $\{0.1, 0.5, 1.0, 5.0\}$ and $\{4, 6, 8, 10, 12\}$, respectively, with
 212 $\alpha = 5.0$ and $L = 8$ determined. We randomly split the training dataset into the training
 213 part (80% of the dataset) and the test part (the remaining 20%). The training part was
 214 used for the training step, and α and L were chosen to minimize the cross-entropy loss of
 215 the test part.

216 3 Data

217 We applied the proposed method to seismic records from a dense regional seismograph
 218 network, MeSO-net, deployed in the Tokyo metropolitan area, Japan, to demonstrate its
 219 performance. MeSO-net consists of 296 accelerometers at an average station interval of
 220 approximately 5 km. The station intervals in the central part of the station network are in
 221 the 2-3 km range. The observed waveforms contain various types of noise signal, such as
 222 from human activity, that originated from near-surface environments (Kawakita & Sakai,
 223 2009), even though the accelerometers were installed approximately 20 m below the surface.
 224 These noise signals appear only randomly in time at individual stations, whereas the seismic
 225 signals do not. Therefore, analysis of the information commonly recorded in the dense
 226 station network potentially improves the capability of earthquake detection.

227 Here, we selected 13 stations located at the eastern part of MeSO-net as the target
 228 seismic array (Fig. 1) because they are distributed approximately in the center of the
 229 seismically active area in and around the Kanto district (Fig. 3). Therefore, this seismic
 230 array provided the most earthquakes determined by visual picking that are available for
 231 supervised learning by the CNNs. Three-component continuous data were recorded at a
 232 200 Hz sampling rate at each seismic station in the array. We downsampled the records to
 233 25 Hz and then applied a 2-8 Hz bandpass filter to remove stationary background noises.
 234 We detrended an input waveform and normalized it using the median of the maximums of
 235 each waveform.

236 The Earthquake Research Institute, The University of Tokyo, identified 599 earthquakes
 237 from MeSO-net data that spanned the 4–16 September 2011 time period (Fig. 3a). A total
 238 of 6241 noise signals were also obtained as a by-product of applying the short-term average
 239 over long-term average (STA-LTA) method to the waveform data. We manually eliminated
 240 all of the signals with very small SNRs at the target array after checking all of the potential
 241 earthquakes recorded by the array. Finally, we recognized 527 earthquake signals contained
 242 in the waveforms that were recorded by the seismic array.

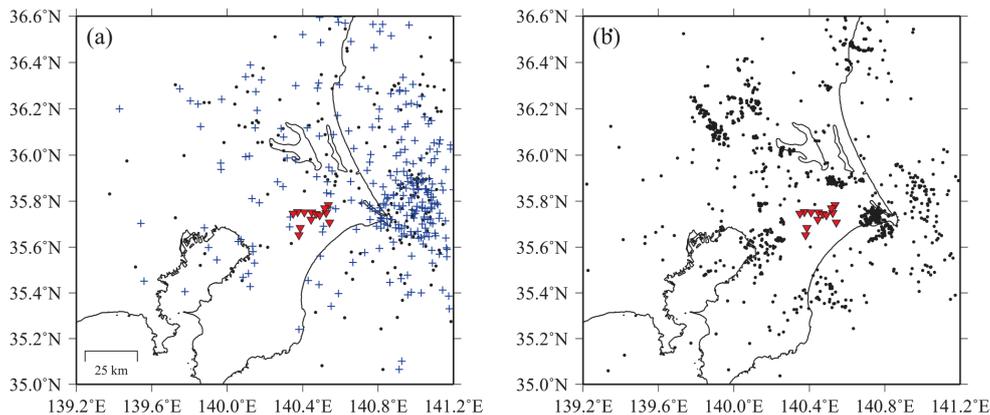


Figure 3. (a) Earthquake locations detected during 4–16 September 2011, and the seismic array used in this study, which consisted of 13 stations (red inverted triangles). Black dots denote the earthquake epicenters for the training dataset, and blue crosses denote those for the validation dataset. (b) Locations of the additional earthquakes (2011–2016 time period) that were used for the training (black dots).

243 We selected the waveforms from 209 available earthquakes that were detected during
 244 4–8 September 2011 as the training earthquake dataset. We also used the waveforms from
 245 927 $M_L > 2.8$ earthquakes that occurred during the 2011–2016 period (excluding 4–16
 246 September 2011) for the training earthquake dataset. We extracted 20-s-long time windows
 247 that contained the onsets of both P- and S-wave arrivals for all of the available events
 248 where at least three stations recorded detectable P- and S-wave onsets. Figure 3 (b) shows
 249 the distribution of epicenters of these training earthquakes. We employed the waveforms
 250 from 318 earthquakes that were detected during 9–16 September 2011 as the validation
 251 earthquake dataset. We used the noise signals identified by the STA-LTA technique, with
 252 each waveform starting within ± 5 s of the onset of each noise signal, as noise. A total of
 253 6241 noise signals with 20-s-long time windows from 4–16 September 2011 were used: those
 254 from 4–8 September for training, and those from 9–16 September for validation.

255 We employed the dataset augmentation technique owing to the relatively small size
 256 of the original training dataset. Specifically, we injected zero-mean Gaussian white noise
 257 signals with randomly determined scales into the training windows. The distribution of
 258 scales was an exponential distribution with a mean of 0.001.

259 We briefly mention the results from applying modularity maximization to the target
 260 seismic array when L and α in Equation (1) are fixed to values that are specified via hold-out
 261 validation. The resultant partition is illustrated in Figure 2. Modularity maximization
 262 using the Euclidean distance achieves a reasonably equispaced station distribution within
 263 the same partition. Furthermore, histograms of the S-P arrival-time differences are quite

264 similar to each other within the same partition, as shown in Figure S1, indicating that graph
 265 partitioning with the optimized hyperparameters works well.

266 4 Results

267 4.1 Performance of our method

268 We begin with reporting the performance of our method by varying minimum threshold
 269 values for the output earthquake probability. Hereafter, we refer to this threshold as the
 270 “minimum detection probability (MDP).” Both true positives (classifying a true earthquake
 271 as an earthquake) and false positives (classifying a true noise as an earthquake) become
 272 smaller as MDP increases.

Table 1. Resultant true positives and false positives of our method for three different minimum detection probabilities (MDPs). The true positives are events that have been classified as earthquakes from true earthquakes, and false positives are events that have been classified as earthquakes from true noise. We used 318 true earthquakes and randomly selected 500 true noise signals.

Minimum detection probability (MDP)	0.5	0.61	0.9
True Positives	312 (98.1%)	310 (97.5%)	289 (90.9%)
False Positives	38 (7.6%)	17 (3.4%)	2 (0.4%)

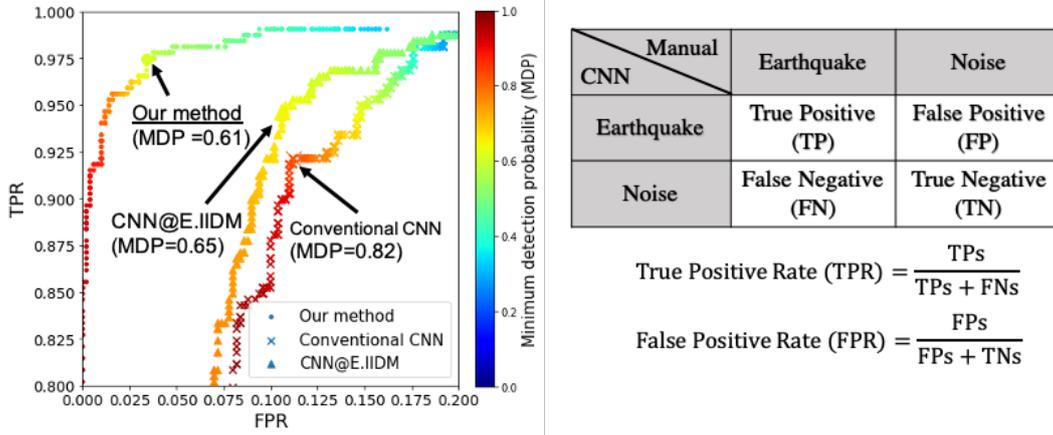


Figure 4. Receiver operating characteristics (ROC) curves for our method and two different CNNs (conventional CNN and CNN@E.IIDM). The ROC curves are plotted as their true positive rates (TPRs) over their false positive rates (FPRs) by changing the minimum detection probability (MDP). The optimal MDP that corresponds to the FPR-TPR point closest to the upper left corner is marked by an arrow for each CNN.

273 As typical cases we present results calculated for three MDPs (0.5, 0.61, 0.9) in Table
 274 1. Our method detects almost all of the earthquakes in the case where MDP is set to
 275 0.5, but it has a relatively large number of false positives. The number of false positives
 276 is significantly reduced when MDP is set to 0.9, but fewer earthquakes are detected in the
 277 process. MDP of 0.61 results in > 97% true positives among the true earthquakes and < 4%

278 false positives among the true noise signals. We chose 0.61 of MDP as an optimal MDP for
 279 our method from the receiver operating characteristics (ROC) curve in Figure 4. The ROC
 280 curve measures the classifier performance as the tradeoff between the true positive and false
 281 positive along the range of MDPs. The choice of 0.61 is the closest to the upper-left corner,
 282 which indicates that the proposed method yields nearly the ideal/best performance at this
 283 MDP. The ROC curve of our method also shows that the false positives are $< 1\%$ when
 284 MDP is large (> 0.8), $< 10\%$ when MDP is > 0.5 , and increases markedly when MDP is
 285 $< \sim 0.35$.

286 We also compared the ROC curve of our method with those of two different CNNs,
 287 namely, a CNN for a single station (CNN@station name) and the conventional CNN for
 288 multiple seismic stations (conventional CNN), as shown in Figure 4. The conventional CNN
 289 means a CNN without any pooling operation along the station axis, which is commonly
 290 employed in multi-station set-ups. We note that our method covers the conventional CNN
 291 as a specific case when L is less than the minimum of the distances between stations, and
 292 also the CNN for a single station when applied to a single station. We adopt E.IIDM as
 293 the single station used in a CNN for a single station because E.IIDM has the highest SNR
 294 among the stations in the seismic array (Fig. S2).

295 The ROC curves of the three methods highlight that our method outperforms the other
 296 methods (Fig. 4). Our method has the lowest false positive rate among the three methods
 297 when the same true positive rate is given. Furthermore, a comparison of the three methods
 298 with the optimal values of MDPs (the closest points to the upper-left corner in the ROC
 299 curves) indicates that our method attains both the most true positives and fewest false
 300 positives.

301 4.2 Temporal variation in output probability when detecting an event

302 We checked time variation of the performance of our method. We calculated the output
 303 probability by applying our method to 2-8 Hz-bandpass-filtered waveform records. We
 304 obtained a time series of output probabilities by shifting the 20-s-long window in 0.02 s
 305 increments. Figure 5 shows a time variation of the output probability, which shows that
 306 the output of our method exhibits reasonable behavior sensitive enough to detect seismic
 307 signals: The output probability suddenly increases and exceeds 0.61 when P-waves have
 308 arrived in the time window at several stations (1), 0.95 when P-waves have arrivals in the
 309 time window at almost all of the stations (2), and 0.98 when both P- and S-waves appear
 310 in the window at almost all of the stations (3), and gradually reduces to < 0.61 when the
 311 S-wave arrivals are outside of the time window at several stations (4).

312 4.3 Effects of the signal-to-noise ratio on event detection

313 We introduced an SNR-based metric, the median of the ratios of the in-event median
 314 absolute deviances (MADs) to the pre-event MADs, to quantify the impact of the SNR on
 315 the performance of the method. Specifically, we defined the median MAD ratio (MADR)
 316 as:

$$317 \text{ median MADR} = \text{median} \left\{ \frac{\text{median}\{|a_{\text{in},i}(t) - \tilde{a}_{\text{in},i}| : t = 1, \dots, T\}}{\text{median}\{|a_{\text{pre},i}(t) - \tilde{a}_{\text{pre},i}| : t = 1, \dots, T\}} : i = 1, \dots, N \right\},$$

318 where T is the time length, N is the number of stations, $\{a_{\text{pre},i}(t) : t = 1, \dots, T\}$ is the
 319 pre-event recorded amplitude at station i , $\{a_{\text{in},i}(t) : t = 1, \dots, T\}$ is the in-event recorded
 320 amplitude at station i , $\tilde{a}_{\text{pre},i}$ is the median of $a_{\text{pre},i}(t)$ with respect to time t , and $\tilde{a}_{\text{in},i}$ is the
 321 median of $a_{\text{in},i}(t)$ with respect to time t . We used the 20-s-long waveforms that started 30 s
 322 prior to P-wave arrivals as the pre-event waveforms. We defined the three-component median
 323 MADR as the median of three median MADRs that were calculated from the UpDown,
 324 NorthSouth, and EastWest components. The median MADR is a median version of the
 325 root-mean-square (RMS) deviation ratio and is more robust than the RMS deviation ratio
 326 when the waveforms are contaminated (see Chapter 5 of Huber and Ronchetti (2009)).

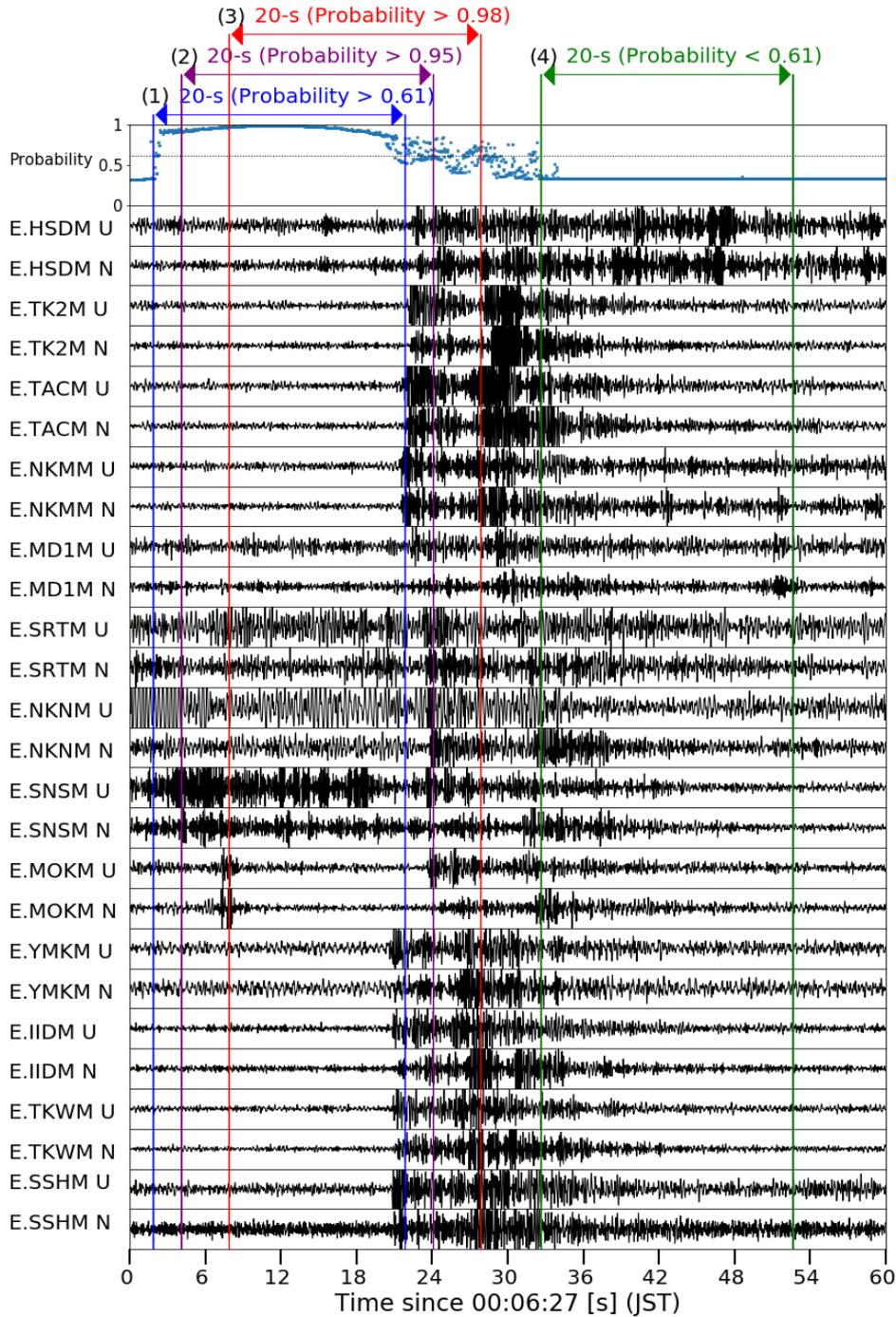


Figure 5. Performance test using the bandpassed (2-8Hz) waveforms of the manually detected earthquake that occurred at approximately 00:06 JST on 9 September 2011. The output probabilities were calculated by shifting the 20-s-long windows in 0.02 s increments. The Japan Meteorological Agency (JMA) magnitude of this event is 2.1. Blue dots at the top denote output probabilities with the dashed line denoting the 0.61 MDP line. (1,2,3,4) four 20-s-long windows.

327
328

Figure 6 (a) shows that the lower bound of the output probabilities of events with the same median MADR increases with increasing median MADR. Our method generally

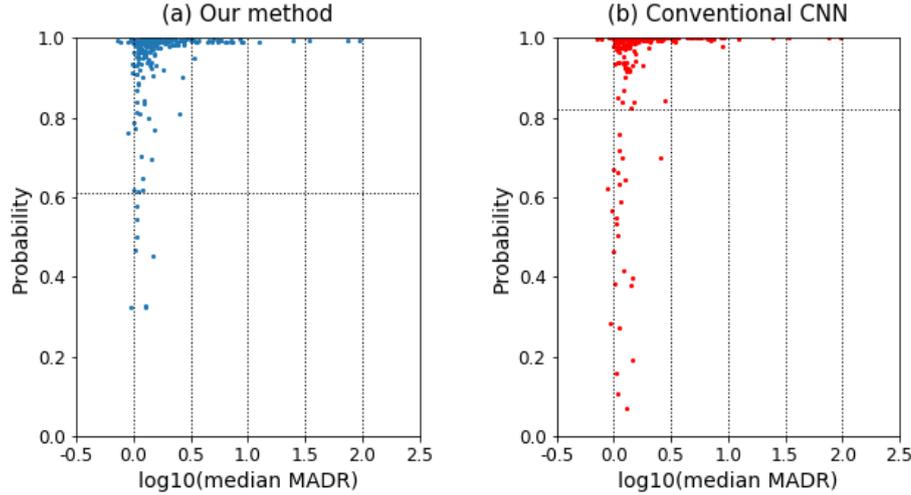


Figure 6. Output probabilities and three-component median MADRs for (a) our method and (b) the conventional CNN. The dotted horizontal lines denote the optimal MDPs of 0.61 and 0.82 in our method and the conventional CNN, respectively.

329 misses earthquakes with a small median MADR (i.e., a small SNR). Similar behaviors
 330 are observed in the conventional CNN (Fig. 6b), but our method generally yields higher
 331 output probabilities compared with the conventional CNN results. The combination of a
 332 higher output probability (Fig. 6) and lower false detection rate (Fig. 4) indicates that the
 333 appropriate clustering of stations in a seismic array greatly improves CNN-based earthquake
 334 detection.

335 It is worth noting that events with median MADR values of < 1.0 can in some cases be
 336 detected with relatively high output probabilities. This means that changes in the pre- and
 337 in-event amplitudes are not significant for such events. Further analysis of the waveform
 338 and/or amplitude features of these events may improve the detection accuracy of earthquakes
 339 with small SNRs.

340 5 Discussion

341 5.1 Robustness with respect to missing data

342 Continuous seismic records may be incomplete owing to temporary power outages,
 343 as well as failures in seismic sensors and/or data loggers. Therefore, we examined the
 344 robustness of our proposed method with respect to such incomplete data records to
 345 investigate its performance and stability. We created synthetic data that consisted of
 346 waveforms with missing data using the validation dataset (318 earthquakes). We also
 347 randomly removed seismic stations and created 10 synthetic datasets with different numbers
 348 of missing stations for the earthquakes in the validation dataset.

349 Figure 7 (a) illustrates how the number of missing stations changes output probability
 350 for 3180 trials, with a drop in performance as the number of missing stations increases.
 351 Increasing the number of missing stations from one to six results in 83%, 80%, 78%, 74%,
 352 68%, and 58% of the trials possessing output probabilities of > 0.61 , respectively. It also
 353 shows that the median output probability remains above 80%, regardless of the number of
 354 missing stations, which implies that performance deterioration is relatively low for half of the
 355 trials. These two observations suggest that data quality, in the form of missing data, impacts

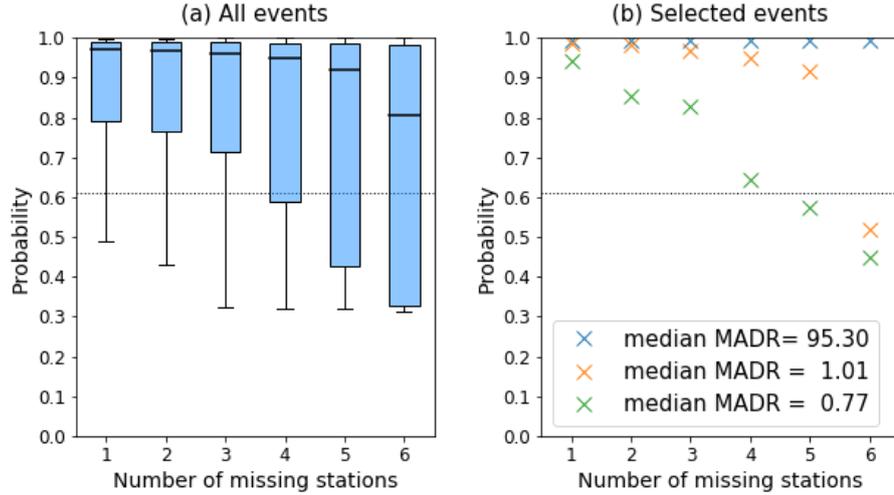


Figure 7. Changes in output probability caused by missing data relative to the optimal MDP (0.61). (a) Boxplots of output probability for different numbers of missing stations. (b) Mean output probability for three selected events that have different three-component median MADRs.

356 on performance deterioration. We selected three events with different median MADRs
 357 to illustrate this point (Fig. 7b). The event with the highest median MADR (= 95.30)
 358 possessed an output probability of approximately 1.0, even when the data loss arising from
 359 missing stations increased. The event with the median MADR nearly 1.0 possessed an
 360 output probability that decayed as a function of data loss and remained high (> 0.9) until
 361 the number of missing stations increased to six. The event with the low median MADR
 362 (= 0.77) possessed an output probability that decayed substantially, falling below 0.61 in
 363 cases where the number of missing stations increased to five. Given the high quality of the
 364 data, our CNN-based method can reliably evaluate earthquakes, even if up to half of the
 365 seismic data are missing.

366 5.2 Continuous application to the 2011 September 9 data

367 We applied our method to 1-hour-long continuous waveform data, starting at 0:00
 368 JST on 9 September 2011, to explore the feasibility of the proposed method in detecting
 369 earthquakes from continuous streaming data. We used the data because a large number
 370 of earthquakes are excited by the 2011-March-11 Tohoku-oki earthquake (M9) even in the
 371 Kanto district (e.g., Ogata et al. (2019)). We calculated the temporal change in the output
 372 probabilities by shifting the 20-s-long window in 0.02 s increments over the entire hour. We
 373 picked event candidates that possessed output probabilities of > 0.61 for more than 1 s
 374 (yellow stars and orange stars in Fig. 8b).

375 A comparison of the time series of the median MAD at the stations within the array and
 376 the output probabilities from our method (Fig. 8) reveals several sharp peaks in the output
 377 probability sequence that are synchronous with local peaks in the median MAD sequence,
 378 with four event candidates being detected with a high degree of confidence. Two of these
 379 event candidates correspond to earthquakes that had been detected by JMA and MeSO-net
 380 (orange stars in Fig. 8), thereby implying the validity of our method in detecting earthquakes
 381 using the seismic array. Although the other two candidates (yellow stars in Fig. 8b) were
 382 not detected in the routine MeSO-net detection procedure, we consider these to be actual
 383 earthquakes, as seismic signals can be identified in the nationwide seismic network in the
 384 Kanto district (e.g., high-sensitivity seismograph network in Japan (Hi-net); (Obara et al.,

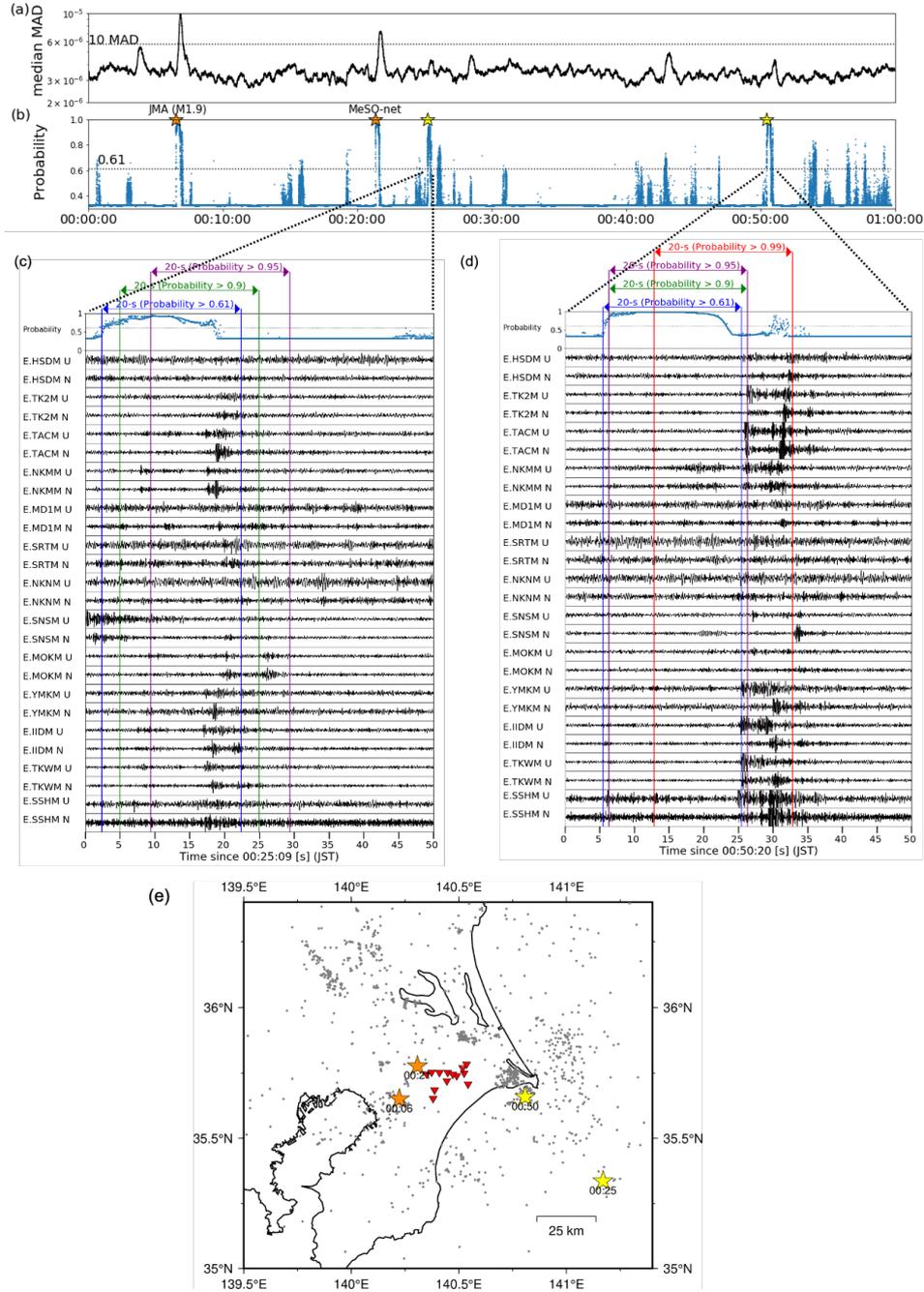


Figure 8. Application of our method to 1-hour-long waveform data starting at 0:00 JST on 9 September 2011. (a) Time series of the median MAD for the 13-station array. (b) Time series of the output probabilities, with the detected events indicated (orange stars if the catalogs exist and yellow stars if not). (c, d) Waveforms of two newly detected events. (e) Epicenters of detected events (orange and yellow stars) and training events during the time period from 4 to 11 September 2011 and the 2011–2016 time period (gray dots) relative to the 13-station array (inverted red triangles).

385 2005; Okada et al., 2004)) at the time of these two events. The relocated hypocenters of
 386 these events, which were determined from arrival times picked visually, fall within the lateral

range of the training earthquakes (Fig. 8e). The event at approximately 00:50:20 JST has a very high (> 0.99) output probability (Fig. 8d). This high output probability and the proximity of the event to a cluster of training earthquakes imply that events occurring near training earthquakes can be detected with a high output probability. MeSO-net consists of a large, dense distribution of seismic stations around the Tokyo metropolitan area compared with the nationwide seismic network. Therefore, the level of performance obtained in this study suggests that implementation of a suitable method can improve the ability to detect and monitor small earthquakes that occur in and around metropolitan areas.

5.3 Extension to large N -arrays

We shall discuss the ability to extend our method to large N -arrays (N is typically about 100 or 1000), given that we restricted our CNN-based application to 13 stations in a local array. Seismic analyses using large N -arrays have recently received increasing attention (Hayashida et al., 2020; Karplus & Schmandt, 2018; Li et al., 2018; Riahi & Gerstoft, 2017) because they allow analysis of very weak seismic signals.

The application of our method to large N -arrays may not be straightforward. A large N -array means that the CNNs will become more computationally intensive because the size of the CNN input becomes substantially larger as the number of stations within the array increases. A practical solution to this issue is the use of a divide-and-conquer approach that recursively decomposes the detection in a larger array into detection in smaller sub-arrays. Another potential issue is that the waveform patterns may become too complex to learn properly as the regional size of the seismic array increases. Incorporating hypocenter location information into the event analysis, as well as utilizing other machine-learning techniques, may assist in overcoming this issue.

6 Conclusions

We developed a CNN that incorporated graph partitioning for earthquake detection using a multi-station local seismic array. The performance of our method is validated by both its high true detection rate and low false detection rate. Our method outperforms two different conventional types of CNN-based detector. We found that the temporal variations in the output earthquake probability of our method during event detection are concordant with the P- and S-wave arrivals in the array. We used the median MADR as a measure of the SNR to confirm that the SNR has some influence on the output earthquake probability of our method; however, some low-SNR events can be detected with relatively high output probabilities. Synthetic tests demonstrated that our method exhibits robustness when some data are missing. We also identified new seismic events with extremely low SNRs that are not listed in existing earthquake catalogs by applying our method to 1-hour-long continuous waveform data. We finally note that the architecture of our method is not task specific and can be applied to other tasks using a seismic array, such as earthquake relocation and classification.

7 Acknowledgements

Seismic waveform data used in the present study is available via the website (http://www.eri.u-tokyo.ac.jp/project/iSeisBayes/MeSO-net_Narita-Array_2020). The Hi-net data (National Research Institute for Earth Science and Disaster Resilience, 2019) were used for relocating detected events. GMT software package (Wessel & Smith, 1998) was used for creating geographical figures. TensorFlow was used for training and applying CNN. NetworkX (Hagberg et al., 2008) was used for conducting graph partitioning. We sincerely thank Hiromichi Nagao, Shigeki Nakagawa, Hiroshi Tsuruoka, and Miwa Yoshida for their help in processing the MeSO-net data. This work is supported by JST CREST Grant Number JPMJCR1763, Japan.

435

References

436

Bergan, K., & Beroza, G. (2018). Detecting earthquakes over a seismic network using single-station similarity measures. *Geophysical Journal International*, *213*(3), 1984–1998. doi: <https://doi.org/10.1093/gji/ggy100>

437

438

439

Bichot, C., & Siarry, P. (2013). *Graph partitioning*. Hoboken, NJ: John Wiley & Sons.

440

441

Fortunato, S., & Castellano, C. (2012). Community structure in graphs. In R. Meyers (Ed.), *Computational Complexity* (pp. 490–512). New York, NY: Springer, New York.

442

443

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.

444

445

446

Hagberg, A., Schult, D., & Swart, P. (2008). Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)* (pp. 11–15). Pasadena, CA.

447

448

449

Hara, S., Fukahata, Y., & Iio, Y. (2019). P-wave first-motion polarity determination of waveform data in western japan using deep learning. *Earth, Planets and Space*, *71*(127). doi: <https://doi.org/10.1186/s40623-019-1111-x>

450

451

452

453

Hayashida, Y., Matsumoto, S., Iio, Y., Sakai, S., & Kato, A. (2020). Nondouble-couple microearthquakes in the focal area of the 2000 Western Tottori earthquake (M 7.3) via hyperdense seismic observations. *Geophysical Research Letters*, *47*(4). doi: <https://doi.org/10.1029/2019GL084841>

454

455

456

457

458

Hirata, N., Sakai, S., Sato, H., Satake, K., & Koketsu, K. (2009). An outline of the special project for earthquake disaster mitigation in the Tokyo metropolitan area—subproject I: Characterization of the plate structure and source faults in and around the Tokyo metropolitan area [in Japanese with English abstract]. *Bulletin of the Earthquake Research Institute, University of Tokyo*, *84*, 41–56.

459

460

461

Huber, P., & Ronchetti, E. (2009). *Robust statistics* (Second ed.). Hoboken, NJ: John Wiley & Sons. doi: <https://doi.org/10.1002/9780470434697>

462

463

464

Kano, M., Nagao, H., Nagata, K., Ito, S., Sakai, S., Nakagawa, S., ... Hirata, N. (2017). Seismic wavefield imaging of long-period ground motion in the Tokyo metropolitan area, Japan. *Journal of Geophysical Research: Solid Earth*, *122*(7), 5435–5451. doi: <https://doi.org/10.1002/2017JB014276>

465

466

467

Karplus, M., & Schmandt, B. (2018). Preface to the focus section on geophone array seismology. *Seismological Research Letters*, *89*(5), 1597–1600. doi: <https://doi.org/10.1785/0220180212>

468

469

470

471

472

Kasahara, K., Sakai, S., Morita, Y., Hirata, N., Tsuruoka, H., Nakagawa, S., ... Obara, K. (2009). Development of the Metropolitan Seismic Observation Network (MeSO-net) for detection of mega-thrust beneath Tokyo metropolitan area [in Japanese with English abstract]. *Bulletin of the Earthquake Research Institute, University of Tokyo*, *84*, 71–88.

473

474

475

Kawakita, Y., & Sakai, S. (2009). Various Types of Noise in MeSO-net [in Japanese with English abstract]. *Bulletin of the Earthquake Research Institute, University of Tokyo*, *84*, 127–139.

476

477

478

Kong, Q., Trugman, D., Ross, Z., Bianco, M., Meade, B., & Gerstoft, P. (2019). Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, *90*(1), 3–14. doi: <https://doi.org/10.1785/0220180259>

479

480

481

482

Kriegerowski, M., Petersen, G., Vasyura-Bathke, H., & Ohrnberger, M. (2019). A deep convolutional neural network for localization of clustered earthquakes based on multistation full waveforms. *Seismological Research Letters*, *90*(2A), 510–516. doi: <https://doi.org/10.1785/0220180320>

483

484

485

Li, Z., Peng, Z., Hollis, D., Zhu, L., & McClellan, J. (2018). High-resolution seismic event detection using local similarity for large- N arrays. *Scientific Reports*, *8*(1646). doi: <https://doi.org/10.1038/s41598-018-19728-w>

486

487

488

McBrearty, L., Delorey, A., & Johnson, P. (2019). Pairwise association of seismic arrivals with convolutional neural networks. *Seismological Research Letters*, *90*(2A), 503–509. doi: <https://doi.org/10.1785/0220180326>

- 489 Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., . . . Petronis, A.
490 (2008). Epigenomic profiling reveals DNA-methylation changes associated with major
491 psychosis. *American Journal of Human Genetics*, *82*(3), 696–711. doi: [https://](https://doi.org/10.1016/j.ajhg.2008.01.008)
492 doi.org/10.1016/j.ajhg.2008.01.008
- 493 Nakano, M., Sugiyama, D., Hori, T., Kuwatani, T., & Tsuboi, S. (2019). Discrimination
494 of seismic signals from earthquakes and tectonic tremor by applying a convolutional
495 neural network to running spectral images. *Seismological Research Letters*, *90*(2A),
496 530–538. doi: <https://doi.org/10.1785/0220180279>
- 497 National Research Institute for Earth Science and Disaster Resilience. (2019). *NIED Hi-net*.
498 National Research Institute for Earth Science and Disaster Resilience. doi: [https://](https://doi.org/10.17598/NIED.0003)
499 doi.org/10.17598/NIED.0003
- 500 Newman, M. (2006). Modularity and community structure in networks. *Proceedings of the*
501 *National Academy of Sciences of the United States of America*, *103*(23), 8577–8582.
502 doi: <https://doi.org/10.1073/pnas.0601602103>
- 503 Newman, M., & Girvan, M. (2004). Finding and evaluating community structure in
504 networks. *Physical Review E*, *69*(2), 026113. doi: [https://doi.org/10.1103/PhysRevE](https://doi.org/10.1103/PhysRevE.69.026113)
505 [.69.026113](https://doi.org/10.1103/PhysRevE.69.026113)
- 506 Obara, K., Kasahara, K., Hori, S., & Okada, Y. (2005). A densely distributed
507 high-sensitivity seismograph network in Japan: Hi-net by National Research Institute
508 for Earth Science and Disaster Prevention. *Review of Scientific Instruments*, *76*,
509 021301. doi: <https://doi.org/10.1063/1.1854197>
- 510 Ogata, Y., Katsura, K., Tsuruoka, H., & Hirata, N. (2019). High-resolution 3D earthquake
511 forecasting beneath the greater Tokyo area. *Earth Planets Space*, *71*(113). doi:
512 <https://doi.org/10.1186/s40623-019-1086-7>
- 513 Okada, Y., Kasahara, K., Hori, S., Obara, K., Sekiguchi, S., Fujiwara, H., & Yamamoto,
514 A. (2004). Recent progress of seismic observation networks in Japan Hi-net, F-net,
515 K-NET and KiK-net. *Earth Planets Space*, *56*, 15–28. doi: [https://doi.org/10.1186/](https://doi.org/10.1186/BF03353076)
516 [BF03353076](https://doi.org/10.1186/BF03353076)
- 517 Peng, Z., & Zhao, P. (2009). Migration of early aftershocks following the 2004 Parkfield
518 earthquake. *Nature Geoscience*, *2*, 877–881. doi: <https://doi.org/10.1038/ngeo697>
- 519 Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake
520 detection and location. *Science Advances*, *4*(2), e1700578. doi: [https://doi.org/](https://doi.org/10.1126/sciadv.1700578)
521 [10.1126/sciadv.1700578](https://doi.org/10.1126/sciadv.1700578)
- 522 Riahi, N., & Gerstoft, P. (2017). Using graph clustering to locate sources within a dense
523 sensor array. *Signal Processing*, *132*, 110–120. doi: [https://doi.org/10.1016/j.sigpro](https://doi.org/10.1016/j.sigpro.2016.10.001)
524 [.2016.10.001](https://doi.org/10.1016/j.sigpro.2016.10.001)
- 525 Ross, Z., Meier, M., & Hauksson, E. (2018). P wave arrival picking and first-motion
526 polarity determination with deep learning. *Journal of Geophysical Research: Solid*
527 *Earth*, *123*(6), 5120–5129. doi: <https://doi.org/10.1029/2017JB015251>
- 528 Sakai, S., & Hirata, N. (2009). Distribution of the Metropolitan Seismic Observation
529 network [in Japanese with English abstract]. *Bulletin of the Earthquake Research*
530 *Institute, University of Tokyo*, *84*, 57–70.
- 531 Wang, J., Xiao, Z., Liu, C., Zhao, D., & Yao, Z. (2019). Deep learning for picking seismic
532 arrival times. *Journal of Geophysical Research: Solid Earth*, *124*(7), 6612–6624. doi:
533 <https://doi.org/10.1029/2019JB017536>
- 534 Wessel, P., & Smith, W. (1998). New, improved version of generic mapping tools released.
535 *EOS Transactions, American Geophysical Union*, *79*, 579. doi: [https://doi.org/10](https://doi.org/10.1029/98EO00426)
536 [.1029/98EO00426](https://doi.org/10.1029/98EO00426)
- 537 Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: a network visualization tool for
538 human brain connectomics. *PLoS One*, *8*(7), e68910. doi: [https://doi.org/10.1371/](https://doi.org/10.1371/journal.pone.0068910)
539 [journal.pone.0068910](https://doi.org/10.1371/journal.pone.0068910)
- 540 Zhang, X., Zhang, J., Yuan, C., Liu, S., Chen, Z., & Li, W. (2020). Locating induced
541 earthquakes with a network of seismic stations in Oklahoma via a deep learning
542 method. *Scientific Reports*, *10*(1941). doi: [https://doi.org/10.1038/s41598-020-58908](https://doi.org/10.1038/s41598-020-58908-5)
543 [-5](https://doi.org/10.1038/s41598-020-58908-5)

544 Zhu, W., & Beroza, G. (2018). PhaseNet: A deep-neural-network-based seismic arrival-time
545 picking method. *Geophysical Journal International*, 216(1), 261–273. doi: [https://](https://doi.org/10.1093/gji/ggy423)
546 doi.org/10.1093/gji/ggy423