Discrimination of Seismic Signals from Earthquakes and Tectonic Tremor by Applying a Convolutional Neural Network to Running Spectral Images

by Masaru Nakano, D. Sugiyama, T. Hori, T. Kuwatani, and S. Tsuboi

ABSTRACT

Monitoring of seismic signals generated by slow deformation at convergent and transform plate boundaries worldwide, known as tectonic tremor, might provide insights into deformation processes in the source regions of megathrust earthquakes. Tremor signals occur dominantly in the 2–8 Hz frequency band and can last for tens of seconds to several minutes, in contrast to typical earthquakes that produce seismic signals at frequencies up to several tens of hertz and last less than a minute. Because tremor is caused by stochastic processes, the resultant waveforms are represented by a stochastic function and construction of deterministic measures to discriminate tremor signals from earthquakes is very difficult. In this study, we used a convolutional neural network (CNN) to discriminate the signals of tectonic tremor from those of local earthquakes in running spectral images of these signals. We developed a method (seismic running spectra-CNN [SRSpec-CNN]) that is sensitive to the absolute frequency of signal appearance, which reflects the physical properties of the signal source, but is insensitive to the time of signal onset. SRSpec-CNN has 130,211 parameters that were trained by 17,213 images of 64 × 64 pixels. Based on simultaneous analyses of the frequency contents and durations of the signals, we achieved 99.5% accuracy for our identifications of signals from tectonic tremor, local earthquakes, and noise. Because running spectra clearly differentiate the characteristic features of these signals, we were able to achieve this high accuracy by using a CNN of simple architecture.

Electronic Supplement: Figures showing classification results for images not included in the training/validation data.

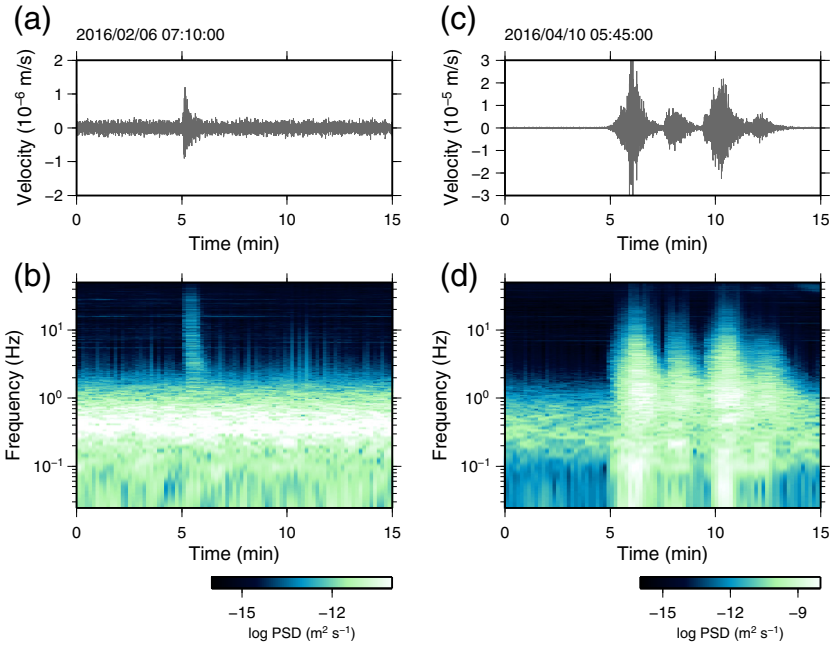
INTRODUCTION

The recent development of dense high-sensitivity geophysical observation networks has led to many observations of seismic and geodetic signals known as tectonic tremor, very-low-frequency earthquakes, and slow-slip events in subduction zones (e.g., Dragert *et al.*, 2001; Obara, 2002; Rogers and Dragert, 2003; Obara and Ito, 2005; Obara and Kato, 2016) and at transform plate boundaries such as the San Andreas fault (e.g., Nadeau and Dolenc, 2005). Because of the high sensitivity of these phenomena to crustal stress changes, monitoring such slow earthquake activity might provide information about the likelihood of impending megathrust earthquakes (Obara and Kato, 2016).

Tectonic tremor signals are often accompanied by very-low-frequency earthquakes and slow-slip events that share the same source location, which implies a common source (Rogers and Dragert, 2003; Ito *et al.*, 2007; Kaneko *et al.*, 2018; Nakano *et al.*, 2018). Of these, tectonic tremor has been the most frequently observed, and numerous studies clarified the spatial and temporal variations of tremor signal characteristics (e.g., Obara, 2010; Obara *et al.*, 2010; Yabe and Ide, 2014; Annoura *et al.*, 2016). Recent observations revealed that tremor also occurs in the shallow parts of the subduction zones as the Nankai trough and Japan trench (Sugioka *et al.*, 2012; Ito *et al.*, 2015; Yamashita *et al.*, 2015).

Because of the emergent onset of tremor signals, traditional methods of event detection based on the ratio of short-term and long-term average amplitudes often fail. Consequently, the envelope correlation method (Obara, 2002; Ide, 2010a) is commonly used to detect and locate tectonic tremor. This method searches for coherent seismic signals on the basis of the similarity of waveform envelopes among stations. Because this method does not distinguish between earthquakes and tremors, the resultant catalog might include local earthquakes. Therefore, efficient monitoring of tremor requires a method that can discriminate signals generated by earthquakes from those generated by tremor. However, because the source of tremor can be described by stochastic processes (Ide, 2008, 2010b), the observed waveforms are not always similar to each other, even

doi: 10.1785/0220180279



▲ **Figure 1.** Comparison of waveforms and spectral features of events recorded at Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET) station KMD13 (location in Fig. 2): (a,b) a local earthquake (M_L 1.9) about 17 km below the station, (c,d) a slow earthquake about 7 km below the station that shows both tremor and very-low-frequency signals. PSD, power spectral density. The color version of this figure is available only in the electronic edition.

for events that occur close together (Toh *et al.*, 2018). Accordingly, the development of classification methods based on deterministic measures is difficult.

Machine learning (ML) approaches have been applied to detect and classify seismic signals for various source types and tectonic settings (e.g., Dowla *et al.*, 1990; Dysart and Pulli, 1990; Wang and Teng, 1995; Del Pezzo *et al.*, 2003). Recent improvements of ML methodologies and computer performance (e.g., LeCun *et al.*, 2015) greatly improved the accuracy of the detection and classification of various seismic signals (e.g., Yoon *et al.*, 2015; Kong *et al.*, 2016; Rouet-Leduc *et al.*, 2017; Holtzman *et al.*, 2018; Perol *et al.*, 2018).

Tremor signals are characterized by frequencies dominantly in the 2–8 Hz band that continue for tens of seconds to several minutes, whereas earthquakes produce signals at frequencies up to tens of hertz, and have durations shorter than several tens of seconds (Fig. 1). These differences arise from the different source processes generating these signals. Simultaneous analyses of these different characteristics using running spectra, rather than raw waveforms, should better distinguish tremor signals from those of regular earthquakes. Among the various ML methodologies, the convolutional neural network (CNN) can recognize characteristic features embedded in images of a given class and is thus suitable for image classification (LeCun *et al.*, 1999, 2015). The few studies that have applied the CNN to seismic data have used time-series seismic waveforms as input (Li *et al.*, 2018; Perol *et al.*, 2018).

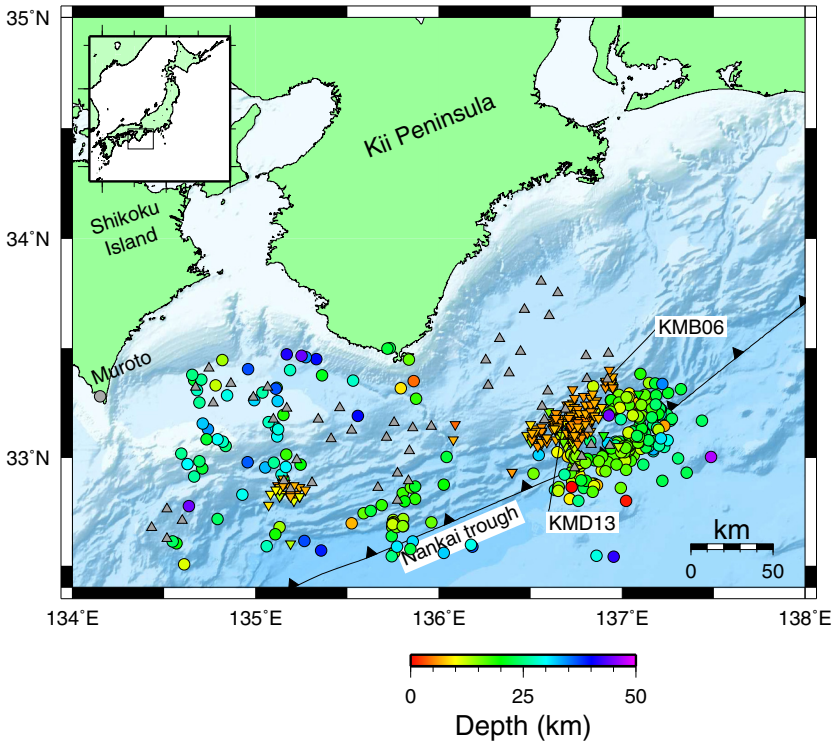
Here, we developed a method to distinguish seismic signals based on frequency–time series. We applied this method to seismometer recordings obtained from the Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET; Fig. 2) installed along the Nankai trough (Kaneda *et al.*, 2015; Kawaguchi *et al.*, 2015), where shallow tremor occurs at intervals of several years (Sugioka *et al.*, 2012; To *et al.*, 2015; Annoura *et al.*, 2017). Our objective was to distinguish between signals from regular earthquakes, tectonic tremor, and noise by applying CNN-based image recognition to running spectra of these signals. Because the frequency content of a seismic signal reflects the physical properties of its source, we developed a CNN that senses the frequencies of such signals.

The CNN is a supervised learning procedure wherein the network is first trained to recognize the characteristic features of each known class by using a training dataset. We created our training dataset from existing seismic event catalogs. After training, we evaluated performance of the method by applying it to images that had not been included in the training dataset. Our CNN successfully distinguished the seismic signals of local earthquakes, tremors, and noise with 99.5% accuracy. Use of the method presented herein will improve our ability to monitor tectonic tremor activities.

METHOD

The CNN focuses on spatial correlations of pixel data in an image to extract the characteristic features of a given object by using learnable filters (LeCun *et al.*, 1999, 2015; Ross *et al.*, 2018). This approach provides more effective recognition of objects in images than the traditional fully connected (FC) feedforward neural networks. CNNs are generally composed of convolutional, pooling, and FC layers: convolutional layers apply sets of filters to an input image (or the output of a previous layer) to extract characteristic features, pooling layers reduce the sensitivity to the location and scale where the characteristic feature appears, and FC layers perform reasoning based on the output of the convolutional and pooling layers. Sequential connection of these layers constitutes a CNN that places input images in given categories.

CNNs are normally designed to be insensitive to the location of signal appearance in an image. For application of a CNN to running spectra of seismic signals, it is the frequency component that reflects the physical properties of the signal. Hence, we needed to develop a CNN that is sensitive to frequency but insensitive to the onset time of the signal. To achieve this end, we developed a new CNN-based approach to categorize seismic waveforms of local earthquakes, tectonic tremor, and noise from running spectra, which we named seismic running



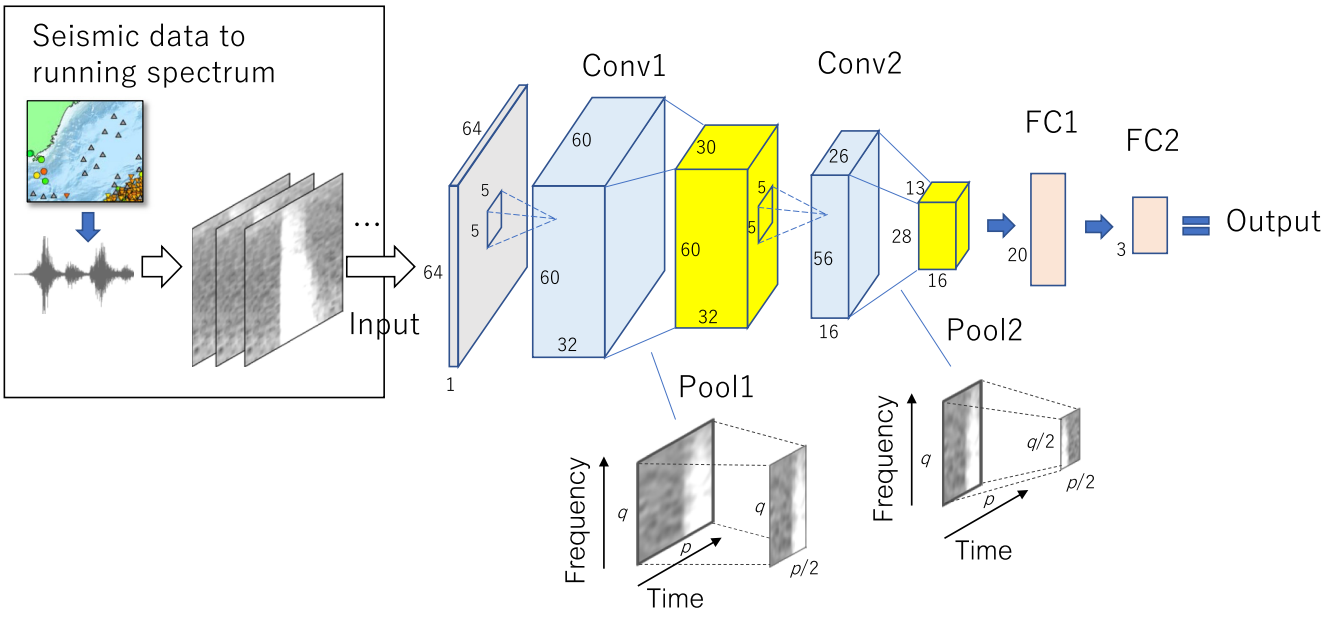
▲ **Figure 2.** Map showing the distribution of DONET stations (gray triangles), hypocenters of regular earthquakes (circles), and slow earthquakes (inverted triangles) used in this study. (Inset) Location of the plotted area. The color version of this figure is available only in the electronic edition.

spectra-CNN (SRSpec-CNN), using TensorFlow (Abadi *et al.*, 2016), an open-source library provided by Google LLC, and referring to learning and object identification as described by LeCun *et al.* (1998, 1999). The architecture of SRSpec-CNN (Fig. 3) includes a special form in its pooling layers to downsample the images but retain frequency sensitivity. The input image has 64×64 pixels, each corresponding to a particular time and frequency pair, with intensity proportional to the power of the seismic signal (see the Data section).

The convolutional layers map an input image to an output by using convolutional filters. We used convolutional filters of dimensions $5 \times 5 \times n$ represented by:

$$u_{i,j,k} = f\left(b_k + \sum_{i'=-2}^{2}\sum_{j'=-2}^{2}\sum_{k'=1}^{n} z_{i+i',j+j',k'} w_{i',j',k',k}\right), \quad (1)$$

in which $u_{i,j,k}$ and $z_{i,j,k}$ are the (i,j)-th pixel of the kth channel in the output and input images for the layer, respectively. The weights $w_{i',j',k',k}$ constitute a filter that is applied to the k'th channel of the input image, b_k is the bias for the kth channel, and n is the number of channels in the input layer. $f(x)$ is an activation function of the form



▲ **Figure 3.** Data flow and architecture of the seismic running spectra-convolutional neural network (SRSpec-CNN) used in this study. The dimensions of the matrices, width and height of input images, and channels, namely the number of parallel images in each layer, are annotated along the sides of the rectangles representing each layer. Conv, Pool, and FC represent convolutional, pooling, and fully connected (FC) layers, respectively. The color version of this figure is available only in the electronic edition.

$$f(x) = \tanh(x). \quad (2)$$

We did not use zero padding on the perimeter of the input image and the size of the output image was reduced by four pixels in both directions.

For the pooling layer, we used a maximum pooling of variable stride $[s, t]$ implemented by an $s \times t$ filter, which downsampled the image size from $p \times q \times n$ pixels to $p/s \times q/t \times n$ pixels as follows:

$$u_{i,j,k} = \max\{z_{si+i',tj+j',k} | i' \in [0, s-1], j' \in [0, t-1]\}. \quad (3)$$

To make the SRSpec-CNN sensitive to the frequency of signal appearance, that is, the vertical direction of the spectral image, we suppressed the pooling in this direction in pooling layer 1 by setting the stride to [2, 1]. The stride for pooling layer 2 was [2, 2].

The FC layers in SRSpec-CNN are of the same form as those of an ordinary neural network, in which all the input and output parameter pairs are connected as

$$u_{l,k} = f\left(b_k + \sum_{l'} \sum_{k'} z_{l',k'} w_{l',k',k}\right), \quad (4)$$

in which the index l' includes all the combinations of (i, j) pairs if the layer is connected to the last pooling layer. The second FC layer outputs the probability of each signal category. The category with the maximum probability is the estimate to which the input image belongs.

DATA

We used running spectral images of daily data recorded at DONET seafloor seismometer stations. The running spectra were computed in 20.48 s overlapping windows with 5.12 s lags for frequencies of 2–10 Hz. The original color image was converted to gray scale with tones proportional to the log of the Fourier amplitude and resized for CNN processing.

For the supervised learning procedure of SRSpec-CNN, we trained the network using training images. Training images were created by clipping running spectral images from continuous seismic records according to the origin time listed in earthquake or tremor catalogs. For local earthquakes, we used a catalog of automatic hypocenter determinations held by Japan Agency for Marine-Earth Science and Technology (JAMSTEC), including only events that occurred below the DONET at depths of 0–50 km between December 2015 and February 2016 (Fig. 2) and using only stations with *P*- and/or *S*-wave first motion readings. For tremor, we used events from a catalog created by Nakano *et al.* (2016, 2018), which includes tremor activity off Muroto during August–September 2015 and off the Kii peninsula in October 2015 and April 2016. Stations used by Nakano *et al.* (2016, 2018) for source inversion analyses were used in this study. For noise, we created training images from randomly selected time windows between December 2015 and March 2016, excluding periods that contain cataloged earthquakes. Examples for each signal class are given in Figure 4.

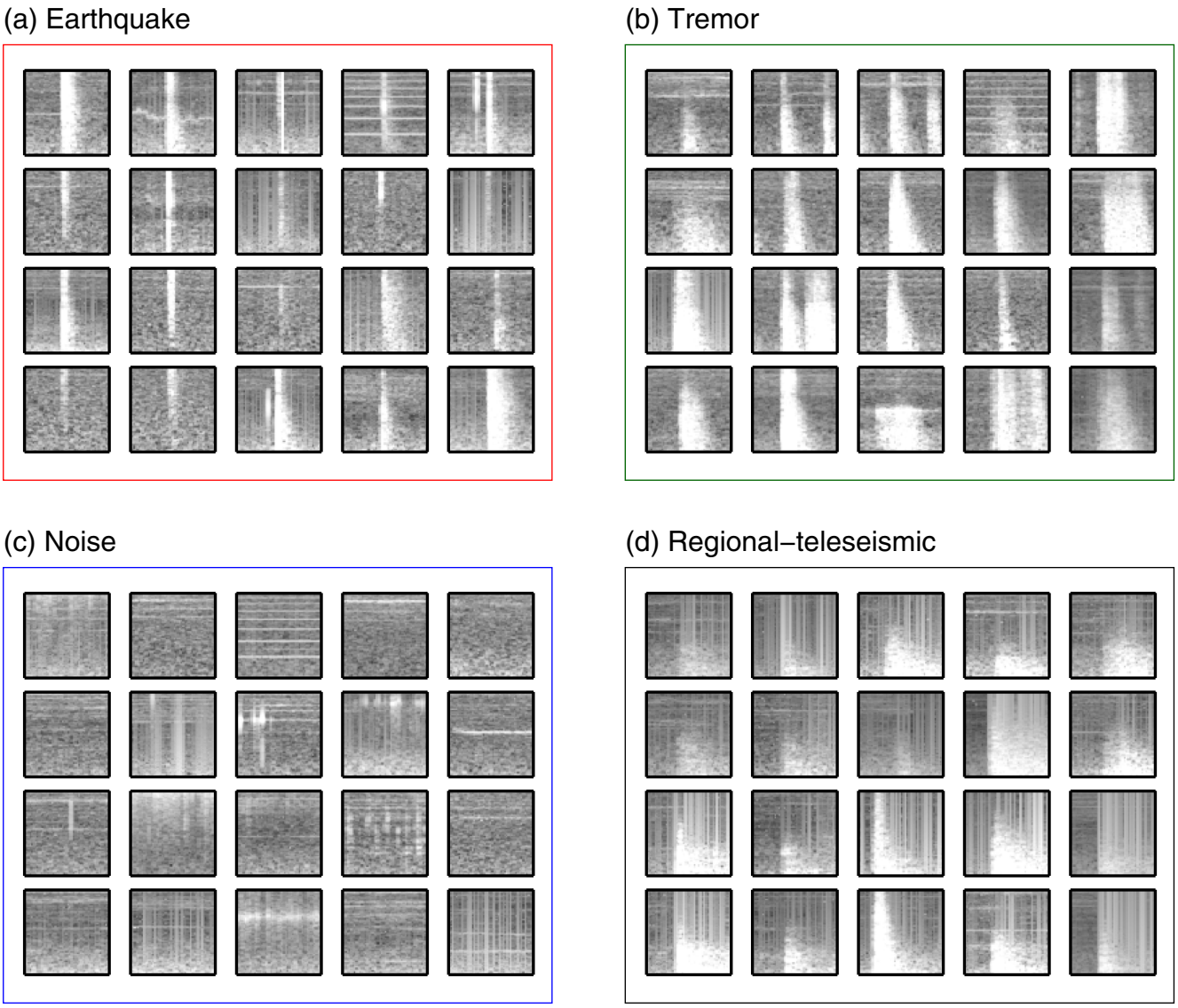
Each training image comprised 64 × 64 pixels, encompassing 225 s and 2–10 Hz in the horizontal and vertical directions, respectively. The time windows we used are comparable to those used for the envelope correlation method (several minutes) (Obara, 2002; Ide, 2010a). The availability of three-component waveforms from broadband and strong-motion seismometers at each DONET station allowed us to obtain up to six training images per station for each time window.

Although the training images were created from stations at which earthquake or tremor signals had been detected, the signals on some images were strongly contaminated with noise. This reflects the ability of the automatic hypocenter determination system to detect weak earthquake signals that were not clearly evident in the running spectral images. We visually inspected the training data and removed these images. We also removed images that included multiple events or continuous tremor signals. For the noise training data, we also removed images that included noncataloged events.

From the resultant dataset, we first removed data from stations KMB06 and KMD13 to evaluate the performance for the classifications of signals from stations not included in the training dataset. We obtained 6227, 5969, and 8148 training images for local earthquake, tremor, and noise, respectively. From these images, we randomly selected earthquake and tremor events that include about 15% images from each category to be used as validation data. For noise, randomly selected time windows were used. These images were excluded from the CNN-training data.

RESULTS AND DISCUSSION

The parameters of SRSpec-CNN were trained using a cross-entropy loss function with the Adam optimization algorithm (Kingma and Ba, 2015) in minibatches of 64 images with a GeForce GTX 1080 Ti graphical processing unit provided by NVIDIA Corporation. The 130,211 parameters of SRSpec-CNN were trained by 17,213 images of 64 × 64 pixels. After training the SRSpec-CNN, we evaluated its accuracy by applying it to the validation data. The accuracy was 97.4%. If we used the stride of [2, 2] in the pooling layer 1, the corresponding value was 95.9%. We further improved the accuracy of SRSpec-CNN by reducing the learning rate to 0.0001 from previous 0.001, and achieved a high level of accuracy of 99.5%. The confusion matrix (Table 1) shows that recall was 99.4%, 99.1%, and 99.8% for local earthquakes, tremor, and noise, respectively. The corresponding values for precision were 99.1%, 99.7%, and 99.6%. Despite the relative simplicity of the structure of SRSpec-CNN, we achieved higher accuracy than 94.8% achieved by Perol *et al.* (2018) using seven classes and eight convolutional layers, although the seven classes they used would make it more difficult to achieve a high level of accuracy than was the case in our study. We attribute the high accuracy we achieved to the clear presence in the input data of the characteristic features as the frequency contents and durations that



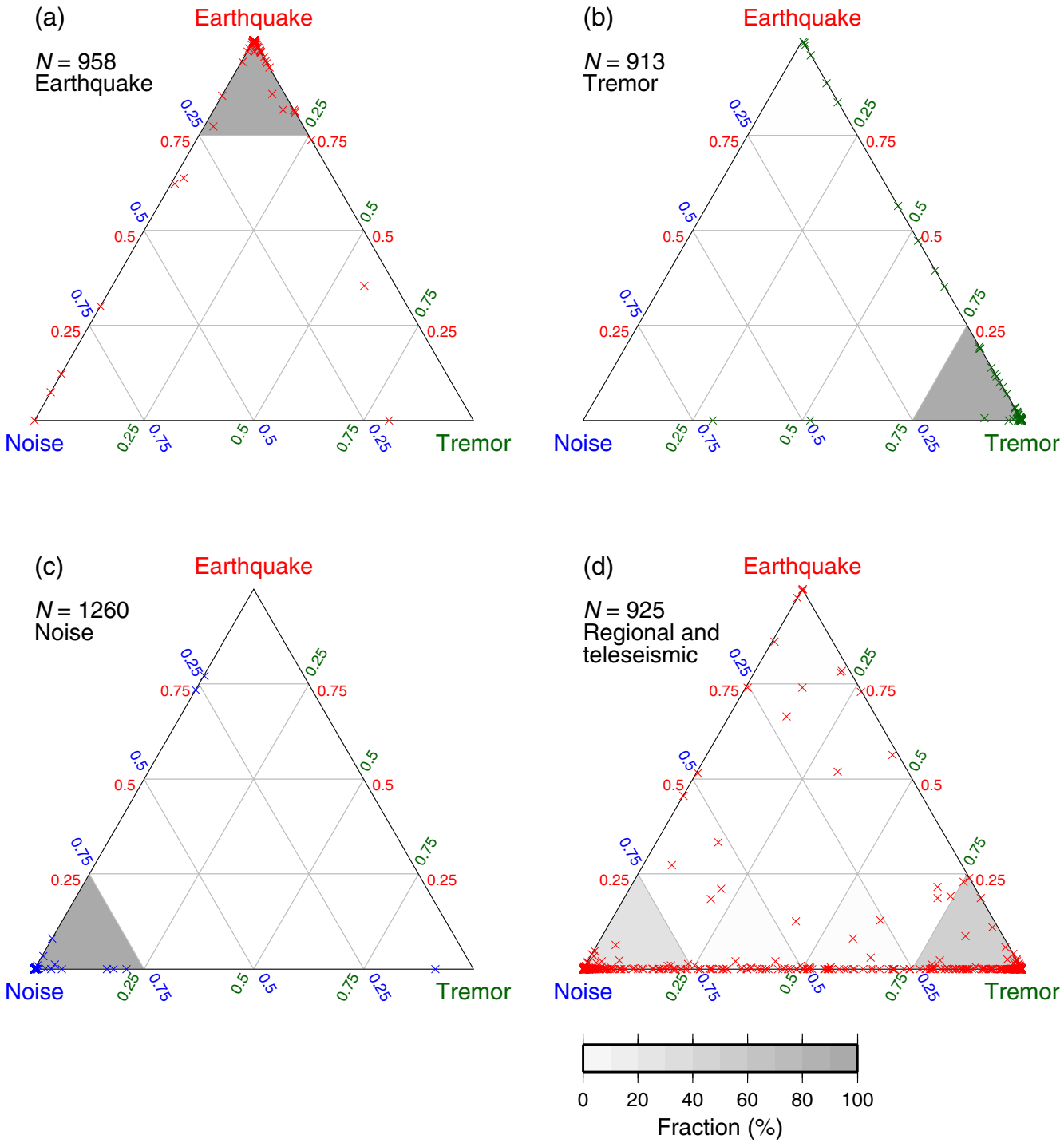
▲ **Figure 4.** Examples of running spectral images. (a) Local earthquakes, (b) tectonic tremor, (c) noise, and (d) regional and teleseismic earthquakes. The horizontal and vertical axes of each image correspond to time and frequency, respectively. The color version of this figure is available only in the electronic edition.

Table 1
Confusion Matrix for Validation Data

	Predicted Event Type			
Classification	**Earthquake (n = 961)**	**Tremor (n = 908)**	**Noise (n = 1262)**	**Recall**
In: Earthquake (n = 958)	952	2	4	0.994
In: Tremor (n = 913)	7	905	1	0.991
In: Noise (n = 1260)	2	1	1257	0.998
Precision	0.991	0.997	0.996	

distinguish our three classes of signals. The probability distributions for each input class of the validation data (Fig. 5) show that more than 98% of the input data were correctly classified with a probability of more than 90%.

In typical CNNs, pooling is applied in both the horizontal and vertical directions of the input image. In SRSpec-CNN, we suppressed pooling in the vertical direction (the frequency of signal appearance). This modification improved the accuracy of our signal classifications by about 1%. Because the frequency of signal appearance reflects the physical properties of the signal source, differences in frequency content are



▲ **Figure 5.** Ternary probability diagrams showing SRSpec-CNN results for (a–c) validation data and (d) regional and teleseismic earthquakes. (a) Local earthquakes, (b) tectonic tremor, and (c) noise. (d) Results of application of SRSpec-CNN to DONET data for regional earthquakes. The gray-scale shading of the triangular segments represents the density of data points within them. *N* is the total number of each event type analyzed. The color version of this figure is available only in the electronic edition.

crucial for seismic signal classifications (e.g., Yoon *et al.*, 2015; Holtzman *et al.*, 2018). The use of architecturally simple neural networks for direct recognition of different types of seismic signals on the basis of frequency, duration, or possibly other properties may help improve determinations of source types, magnitudes, and source locations.

We next tried using SRSpec-CNN (as trained above) to classify running spectral images of signals from stations KMB06

and KMD13 for events selected for the validation data. The accuracy was 99.2% (Ⓔ Fig. S1, available in the electronic supplement to this article). Because we excluded these stations from the training dataset, the high accuracy of our classifications indicates that SRSpec-CNN learned general, rather than site-specific, characteristics of the seismic signals in our study area. Therefore, if new stations are added to the network, no additional training should be necessary. Kong *et al.* (2016) used data from a strong-motion seismometer network to train a neural network to identify earthquake motions recorded on smartphones. Perol *et al.* (2018) showed that their CNN recognized earthquake signals not included in the training data, even recognizing events for which both the polarity of the first motion and the pulse duration were different to those of the training data. These applications indicate that neural networks have the potential to correctly classify unseen signals, or signals recorded by different devices, if properly trained.

In this study, we discriminated tremor from local earthquakes on the basis of the longer duration and characteristic depletion of high frequencies of tremor signals. Large earthquakes at regional distances can have waveform signatures that are similar to those of tremor because of scattering in the crust and attenuation of higher frequencies during wave propagation. We tried using SRSpec-CNN (as trained above) to classify running spectral images of signals from regional and teleseismic earthquakes of magnitude 7 or larger between 2015 and 2016 at epicentral distances between 3° and 83° (Fig. 4d). These signals look very similar to the tremor signals. These data were not included in the training dataset. As expected, most of these signals were classified as tremor (Fig. 5d). Although in detail, the waveforms of regional earthquakes differ from those of tremor signals, SRSpec-CNN as trained in this study could not identify them, probably because of the limited range and resolution of both time and frequency of the running spectral images we used. However, our objective was to automate the discrimination of tremor signals from local earthquakes detected by the envelope correlation method. Because the sources of regional earthquakes are far from DONET, they will not be included in the tremor catalog. Besides, by using the long-period components of the signals, the regional earthquakes are easily recognized because of clearly identifiable dispersion of the Rayleigh waves, of which dispersion is not clearly recognized for tremor signals that occurred close to the observation network.

We also tried classifications of tremor signals strongly contaminated with noise, which have been removed from the training dataset. The recall was 67.9% (Ⓔ Fig. S2), which implies that our SRSpec-CNN has the ability to identify signals contaminated with noise to some degree but misses some signals for severe conditions. Ⓔ Figure S3 shows examples of the noisy images located around the boundary of each class, in which conditions when the performance starts to drop are hard to observe and further studies are necessary to improve the accuracy. Including these noisy images to the training data, or the data augmentation, in which artificial noise is added to the training images, would improve the accuracy for noisy data.

In this study, signal classifications were based on running spectral images of limited resolution in time and frequency. The use of higher resolution data as Fourier transforms computed for successive time windows would improve the classification capability and accuracy, especially for discrimination between tremors and regional earthquakes. However, the additional computational cost arises and determining the most suitable CNN structure may be laborious for higher resolution data. Further improvements of efficient and precise signal classification methods based on SRSpec-CNN is necessary to improve our ability to monitor seismic activities from massive datasets that include observations from a number of networks.

CONCLUDING REMARKS

We developed SRSpec-CNN to be sensitive to the frequency domain of running spectral images to discriminate among tectonic tremor, local earthquakes, and noise. Because running spectra simultaneously represent the frequency contents and durations of seismic signals, we achieved 99.5% accuracy in our use of SRSpec-CNN to discriminate between these signals. Suppression of pooling in the frequency domain improved the accuracy of our results. Our input data for this study were 64×64 pixel images, though using data of higher resolution in both frequency and time may further improve the accuracy of SRSpec-CNN and allow its application to larger, combined datasets for classification, and monitoring of various types of seismic activity.

DATA AND RESOURCES

Seismograms recorded by Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET) were collected by the National Research Institute for Earth Science and Disaster Resilience (NIED) and Japan Agency for Marine-Earth Science and Technology (JAMSTEC). These data can be obtained from the NIED website at http://www.hinet.bosai.go.jp/?LANG=en (last accessed December 2018). ∎

ACKNOWLEDGMENTS

The authors appreciate the contributions of the people involved in the development and maintenance of Dense Oceanfloor Network System for Earthquakes and Tsunamis (DONET). The authors also appreciate fruitful discussions with members of the Cooperative Research Program of the Earthquake Research Institute, University of Tokyo (2018-B01). Comments from K. Bergen, M.-A. Meier, and an anonymous reviewer greatly improved the manuscript.

REFERENCES

Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isardm, *et al.* (2016). TensorFlow: A system for large-scale machine learning, *Proc. 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, Vol. 16, 265–283.

Annoura, S., T. Hashimoto, N. Kamaya, and A. Katsumata (2017). Shallow episodic tremor near the Nankai Trough axis off southeast Mie prefecture, Japan, *Geophys. Res. Lett.* **44,** 3564–3571, doi: 10.1002/2017GL073006.

Annoura, S., K. Obara, and T. Maeda (2016). Total energy of deep low-frequency tremor in the Nankai subduction zone, southwest Japan, *Geophys. Res. Lett.* **43,** 2562–2567, doi: 10.1002/2016GL067780.

Del Pezzo, E., A. Esposito, F. Giudicepietro, M. Marinaro, M. Martini, and S. Scarpetta (2003). Discrimination of earthquakes and underwater explosions using neural networks, *Bull. Seismol. Soc. Am.* **93,** 215–223.

Dowla, F. U., S. R. Taylor, and R. W. Anderson (1990). Seismic discrimination with artificial neural networks: Preliminary results with regional spectral data, *Bull. Seismol. Soc. Am.* **80,** 1346–1373.

Dragert, H., K. Wang, and T. S. James (2001). A silent slip event on the deeper Cascadia subduction interface, *Science* **292,** 1525–1528.

Dysart, P. S., and J. J. Pulli (1990). Regional seismic event classification at the NORESS array: Seismological measurements and the use of trained neural networks, *Bull. Seismol. Soc. Am.* **80,** 1910–1933.

Holtzman, B. K., A. Paté, J. Paisley, F. Waldhauser, and D. Repetto (2018). Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field, *Sci. Adv.* **4,** eaao2929, doi: 10.1126/sciadv.aao2929.

Ide, S. (2008). A Brownian walk model for slow earthquakes, *Geophys. Res. Lett.* **35,** L17301, doi: 10.1029/2008GL034821.

Ide, S. (2010a). Striations, duration, migration and tidal response in deep tremor, *Nature* **466,** 356–359, doi: 10.1038/nature09251.

Ide, S. (2010b). Quantifying the time function of nonvolcanic tremor based on a stochastic model, *J. Geophys. Res.* **115,** doi: 10.1029/2009JB000829.

Ito, Y., R. Hino, S. Suzuki, and Y. Kaneda (2015). Episodic tremor and slip near the Japan Trench prior to the 2011 Tohoku-Oki earthquake, *Geophys. Res. Lett.* **42,** 1725–1731, doi: 10.1002/2014GL062986.

Ito, Y., K. Obara, K. Shiomi, S. Sekine, and H. Hirose (2007). Slow earthquakes coincident with episodic tremors and slow slip events, *Science* **315,** 503–506, doi: 10.1126/science.1134454.

Kaneda, Y., K. Kawaguchi, E. Araki, H. Matsumoto, T. Nakamura, S. Kamiya, K. Ariyoshi, T. Hori, T. Baba, and N. Takahashi (2015). Development and application of an advanced ocean floor network system for megathrust earthquakes and tsunamis, in *Seafloor Observatories*, F. Paolo, B. Laura, and D. S. Angelo (Editors), Springer Praxis Books, Berlin, Germany, 643–662, doi: 10.1007/978-3-642-11374-1_25.

Kaneko, L., S. Ide, and M. Nakano (2018). Slow earthquakes in the microseism frequency band (0.1–1.0 Hz) off Kii Peninsula, Japan, *Geophys. Res. Lett.* **45,** 2618–2624, doi: 10.1002/2017GL076773.

Kawaguchi, K., S. Kaneko, T. Nishida, and T. Komine (2015). Construction of the DONET real-time seafloor observatory for earthquakes and tsunami monitoring, in *Seafloor Observatories*, F. Paolo, B. Laura, and D. S. Angelo (Editors), Springer Praxis Books, Berlin, Germany, 211–228, doi: 10.1007/978-3-642-11374-1_10.

Kingma, D. P., and J. Ba (2015). Adam: A method for stochastic optimization, *Proc. of the 3rd International Conference on Learning Representations (ICLR2015)*, Hilton San Diego Resort-Spa, Sandiego, 7–9 May 2015, available at http://arxiv.org/abs/1412.6980 (last accessed October 2018).

Kong, Q., R. M. Allen, L. Schreier, and Y.-W. Kwon (2016). MyShake: A smartphone seismic network for earthquake early warning and beyond, *Sci. Adv.* **2,** e1501055, doi: 10.1126/sciadv.1501055.

LeCun, Y., Y. Bengio, and G. Hinton (2015). Deep learning, *Nature* **521,** 436–444, doi: 10.1038/nature14539.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner (1998). Gradient-based learning applied to document recognition, *Proc. IEEE* **86,** 2278–2324.

LeCun, Y., P. Haffner, L. Bottou, and Y. Bengio (1999). Object recognition with gradient-based learning, in *Shape, Contour and Grouping in Computer Vision*, D. A. Forsyth, J. L. Mundy, V. di Gesú, and R. Cipolla (Editors), Lecture Notes in Computer Science, Vol. 1681, Springer, Berlin, Germany, doi: 10.1007/3-540-46805-6_19.

Li, Z., M.-A. Meier, E. Hauksson, Z. Zhan, and J. Andrews (2018). Machine learning seismic wave discrimination: Application to earthquake early warning, *Geophys. Res. Lett.* **45,** 4773–4779, doi: 10.1029/2018GL077870.

Nadeau, R. M., and D. Dolenc (2005). Nonvolcanic tremors deep beneath the San Andreas fault, *Science* **307,** no. 5708, 389, doi: 10.1126/science.1107142.

Nakano, M., T. Hori, E. Araki, S. Kodaira, and S. Ide (2018). Shallow very-low-frequency earthquakes accompany slow slip events in the Nankai subduction zone, *Nat. Comm.* **9,** 984, doi: 10.1038/s41467-018-03431-5.

Nakano, M., T. Hori, E. Araki, N. Takahashi, and S. Kodaira (2016). Ocean floor networks capture low-frequency earthquake event, *Eos* **97,** doi: 10.1029/2016EO052877.

Obara, K. (2002). Nonvolcanic deep tremor associated with subduction in Southwest Japan, *Science* **296,** 1679–1681, doi: 10.1126/science.1070378.

Obara, K. (2010). Phenomenology of deep slow earthquake family in southwest Japan: Spatiotemporal characteristics and segmentation, *J. Geophys. Res.* **115,** B00A25, doi: 10.1029/2008JB006048.

Obara, K., and Y. Ito (2005). Very low frequency earthquakes excited by the 2004 off the Kii peninsula earthquakes: A dynamic deformation process in the large accretionary prism, *Earth Planets Space* **57,** 321–326.

Obara, K., and A. Kato (2016). Connecting slow earthquakes to huge earthquakes, *Science* **353,** 253–257, doi: 10.1126/science.aaf1512.

Obara, K., S. Tanaka, T. Maeda, and T. Matsuzawa (2010). Depth-dependent activity of non-volcanic tremor in southwest Japan, *Geophys. Res. Lett.* **37,** L13306, doi: 10.1029/2010GL043679.

Perol, T., M. Gharbi, and M. Denolle (2018). Convolutional neural network for earthquake detection and location, *Sci. Adv.* **4,** e1700578, doi: 10.1126/sciadv.1700578.

Rogers, G., and H. Dragert (2003). Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip, *Science* **300,** 1942–1943, doi: 10.1126/science.1084783.

Ross, Z. E., M.-A. Meier, E. Hauksson, and T. H. Heaton (2018). Generalized seismic phase detection with deep learning, *Bull. Seismol. Soc. Am.* **108,** 2894–2901, doi: 10.1785/0120180080.

Rouet-Leduc, B., C. Hulbert, N. Lubbers, K. Barros, C. J. Humphreys, and P. A. Johnson (2017). Machine learning predicts laboratory earthquakes, *Geophys. Res. Lett.* **44,** doi: 10.1002/2017GL074677.

Sugioka, H., T. Okamoto, T. Nakamura, Y. Ishihara, A. Ito, K. Obana, M. Kinoshita, K. Nakahigashi, M. Shinohara, and Y. Fukao (2012). Tsunamigenic potential of the shallow subduction plate boundary inferred from slow seismic slip, *Nature Geosci.* **5,** 414–418, doi: 10.1038/NGEO1466.

To, A., K. Obana, H. Sugioka, E. Araki, N. Takahashi, and Y. Fukao (2015). Small size very low frequency earthquakes in the Nankai accretionary prism, following the 2011 Tohoku-Oki earthquake, *Phys. Earth Planet. In.* **245,** 40–51, doi: 10.1016/j.pepi.2015.04.007.

Toh, A., K. Obana, and E. Araki (2018). Distribution of very low frequency earthquakes in the Nankai accretionary prism influenced by a subducting-ridge, *Earth Planet. Sci. Lett.* **482,** 342–356, doi: 10.1016/j.epsl.2017.10.062.

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

Yabe, S., and S. Ide (2014). Spatial distribution of seismic energy rate of tectonic tremors in subduction zones, *J. Geophys. Res.* **119,** 8171–8185, doi: 10.1002/2014JB011383.

Yamashita, Y., H. Yakiwara, Y. Asano, H. Shimizu, K. Uchida, S. Hirano, K. Umakoshi, H. Miyamachi, M. Nakamoto, M. Fukui, *et al.*

(2015). Migrating tremor off southern Kyushu as evidence for slow slip of a shallow subduction interface, *Science* **348,** 676–679, doi: 10.1126/science.aaa4242.

Yoon, C. E., O. O'Reilly, K. J. Bergen, and G. C. Beroza (2015). Earthquake detection through computationally efficient similarity search, *Sci. Adv.* **1,** e1501057, doi: 10.1126/sciadv.1501057.

Masaru Nakano
T. Hori
R&D Center for Earthquake and Tsunami
Japan Agency for Marine-Earth Science and Technology (JAMSTEC)
Kanazawa, Yokohama 236-0001, Japan
mnakano@jamstec.go.jp
horit@jamstec.go.jp

D. Sugiyama
S. Tsuboi
Center for Earth Information Science and Technology
Japan Agency for Marine-Earth Science and Technology (JAMSTEC)
Kanazawa, Yokohama 236-0001, Japan
sugiyamad@jamstec.go.jp
tsuboi@jamstec.go.jp

T. Kuwatani[1]
Department of Solid Earth Geochemistry
Japan Agency for Marine-Earth Science and Technology (JAMSTEC)
Natsushima-cho, Yokosuka 237-0061, Japan
kuwatani@jamstec.go.jp

Published Online 3 January 2019

[1] Also at Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), Kawaguchi, Saitama 332-0012, Japan.