Rapid Earthquake Magnitude Classification Using Single Station Data Based on the Machine Learning

Jingbao Zhu, Yueyong Zhou, Heyi Liu, Congcong Jiao, Shanyou Li, Tao Fan, Yongxiang Wei, and Jindong Song

***Abstract*— Magnitude is one of the fundamental parameters of earthquake and also one of the essential information for earthquake early warning (EEW). Rapidly and accurately determining the high-or low-magnitude event is important for mitigating the damage of earthquake hazard. To address the problem of predicting whether an earthquake event is high magnitude ($M \geq 5.5$) or low magnitude ($M < 5.5$), this letter proposes a machine learning magnitude classification framework (MCFrame) using single station data, which consists of feature extraction module and magnitude classifier, and we train the feature extraction module of MCFrame to learn the characteristics of high-magnitude ($M \geq 5.5$) and low-magnitude ($M < 5.5$) earthquake events, using strong-motion records collected from the Japanese Kyoshin network (K-NET) seismic network. Then, the extracted features are used as an input of magnitude classifier to classify earthquake magnitudes. Meanwhile, we analyze the impact of three different magnitude classifiers on the performance of MCFrame, which include the deep neural network (DNN), support vector machine (SVM), and random forest (RF). We show that within 5 s after the P-wave arrival, for these three different classifiers, the magnitude classification accuracy of MCFrame is close, and the MCFrame proposed in this work has a better performance than baseline models for magnitude classification. Additionally, for the MCFrame, the accuracy of high-magnitude events ($M \geq 5.5$) is more than 90%, and the accuracy of low-magnitude events ($M < 5.5$) is more than 99%. These results indicate that the MCFrame proposed in this work is significant for EEW.**

***Index Terms*— Earthquake, machine learning, magnitude classification, P-wave.**

I. Introduction

EARTHQUAKE early warning (EEW) system can provide warning information (such as magnitude, location, warning time, and so on) for users before the arrival of destructive earthquake waves; then, users can take timely measures to prevent and mitigate earthquake disasters. Currently, EEW system has been considered as one of the important tools for earthquake disaster reduction, and it has been operated and tested in many seismically active countries and regions in the world [1].

Predicting whether an earthquake is high or low magnitude is one of the essential tasks of the EEW system. Meanwhile, the magnitude is also one of the indicators used by the EEW system to judge whether to disseminating warning information to users. Additionally, Zollo et al. [2] indicated that the magnitude can be used as one of the parameters to predict whether the site near the station has earthquake potential damage. Some researches indicated that information related to final magnitude can be extracted from P-wave signal [3]. Colombelli et al. [4] found that the initial peak displacement (P_d) can be used to distinguish the large earthquake and small earthquake. Machine learning method uses neural networks to learn complex relationships and can be used to extract relevant features from data. In recent years, machine learning has been used in the field of seismology, and it allows for the exploration of and solution to several seismological problems via the extraction of relevant features from seismic waveforms, such as magnitude estimation [5], [6], [7], [8], earthquake detection and phase picking [9], [10], earthquake location [11], [12], earthquake data denoising [13], [14], peak ground motion prediction [15], and so on.

In this study, to quickly and accurately determine whether an earthquake event is of high or low magnitude, we propose a machine learning magnitude classification framework (MCFrame) based on a single station and explore the feasibility of machine learning method in earthquake magnitude classification. The MCFrame consists of a feature extraction module of a convolutional neural network-recurrent neural

Manuscript received 18 October 2023; revised 19 December 2023; accepted 20 December 2023. Date of publication 25 December 2023; date of current version 4 January 2024. This work was supported in part by the Scientific Research Fund of Institute of Engineering Mechanics, China Earthquake Administration under Grant 2021B07; in part by the National Natural Science Foundation of China under Grant U2039209, Grant 42304074, and Grant 51408564; in part by the Natural Science Foundation of Heilongjiang Province under Grant LH2021E119; in part by the Spark Program of Earthquake Science under Grant XH23027YB; and in part by the National Key Research and Development Program of China under Grant 2018YFC1504003. *(Corresponding authors: Heyi Liu; Jindong Song.)*

Jingbao Zhu, Heyi Liu, Shanyou Li, and Jindong Song are with the Key Laboratory of Earthquake Engineering and Engineering Vibration, Institute of Engineering Mechanics, China Earthquake Administration, Harbin 150080, China, and also with the Key Laboratory of Earthquake Disaster Mitigation, Ministry of Emergency Management, Harbin 150080, China (e-mail: zhujingbao@iem.ac.cn; liuhy@iem.ac.cn; shanyou@iem.ac.cn; jdsong@iem.ac.cn).

Yueyong Zhou and Yongxiang Wei are with the Fujian Earthquake Agency, Fuzhou 350003, China (e-mail: zhouyuey@hotmail.com; wyongx@fjdzj.gov.cn).

Congcong Jiao is with the Earthquake Administration of Shanghai Municipality, Shanghai 200062, China, and also with the Shanghai Sheshan National Geophysical Observatory, Shanghai 200062, China (e-mail: jiaocongcong1014@163.com).

Tao Fan is with the Hubei Key Laboratory of Earthquake Early Warning, Hubei Earthquake Agency, Wuhan 430071, China, also with the Wuhan Institute of Seismic Instruments Co., Ltd., Xianning 437100, China, and also with the Engineering Technology Research Center for Earthquake Monitoring and Early Warning Disposal of Major Projects in Hubei Province, Xianning 437100, China (e-mail: fant_cea@163.com).

This article has supplementary downloadable material available at https://doi.org/10.1109/LGRS.2023.3346655, provided by the authors.

Digital Object Identifier 10.1109/LGRS.2023.3346655

network-attention mechanism (CNN–RNN–AM) and a magnitude classifier. The feature extraction module of MCFrame is trained to learn the characteristics of high-magnitude and low-magnitude earthquake events from the early P-wave signal. The extracted features by feature extraction module of MCFrame are used as an input of magnitude classifier of MCFrame to predict whether an earthquake event is low or high magnitude. MCFrame is trained and tested on the Japanese strong-motion data with M3–M8. Meanwhile, we analyze the impact of three different magnitude classifiers on the performance of MCFrame, which include the deep neural network (DNN), support vector machine (SVM), and random forest (RF). We find that within the first few seconds after P-wave arrival, for these three different classifiers, the magnitude classification accuracy of MCFrame is close, and the MCFrame proposed in this work has a better performance than baseline models for magnitude classification. Meanwhile, to verify the robustness of MCFrame, we apply MCFrame to independent earthquake events, which are not in the training and test datasets, and the results show high accuracy for low- and high-magnitude classification using 3-s P-wave signal.

II. Data

The 129 513 three components of strong-motion acceleration data recorded by the Japanese Kyoshin network (K-NET) seismic network have been used in this study to provide the datasets for training and testing MCFrame. The earthquake catalog includes 2794 seismic events that occurred in Japan between 2007 and 2016 (see Table S1 in the Supplementary Material). The range of magnitude is from M3 to M8, and the focal depth is less than 30 km. The distribution of events and recording stations is shown in Fig. S1 (see the Supplementary Material). The acceleration record is integrated to obtain the velocity record, and then, the velocity record is integrated to obtain the displacement record. To remove the low-frequency drift, a 0.075-Hz high-pass Butterworth filter with four poles is used to filter the records after integration. The P-wave arrival time is selected manually from the vertical acceleration record. We randomly split the data into 70% and 30% subsets (training dataset and test dataset) for training and testing MCFrame.

III. Magnitude Classification Framework

A. Feature Extraction Module of MCFrame

We construct the CNN–RNN–AM feature extraction module [Fig. 1(a)] of MCFrame to extract the characteristics of high- and low-magnitude earthquake event from the early P-wave signal, which is mainly composed of a CNN subblock, an RNN subblock, an AM, dropout, and fully connected layer (FCL).

The CNN subblock [Fig. 1(b)] mainly consists of two convolutional layers (CLs), two batch normalization layers (BNLs), and two max pooling layers (MPLs). The two CLs have 25 and 50 filters with a kernel size of 4. The strides in each CL are 2. Meanwhile, each CL uses the rectified linear unit (ReLU) activation function. A BNL followed each CL. A BNL is followed by an MPL with a pool size of 2 and strides of 2.

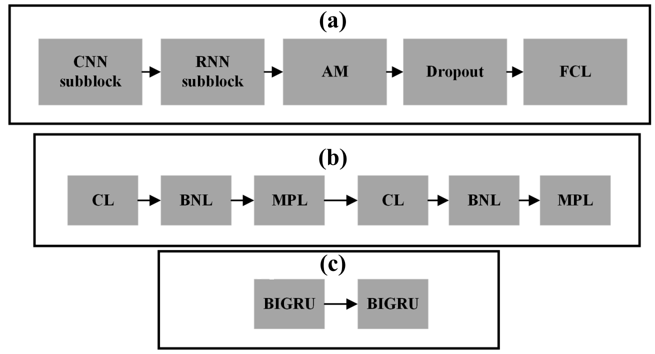

Fig. 1. Architecture of MCFrame. (a) CNN–RNN–AM feature extraction module of MCFrame. (b) CNN subblock of the feature extraction module. (c) RNN subblock of the feature extraction module.

The RNN subblock [Fig. 1(c)] mainly consists of two bidirectional gated recurrent unit (BIGRU) layers [16]. The two BIGRU layers have 50 and 25 units. The command "return_sequences" of each BIGRU layer is set as "true." The RNN subblock is followed by the AM. Meanwhile, to prevent overfitting, a dropout layer with a rate of 0.5 followed the AM. Next, we use an FCL with 100 neurons and the ReLU activation function to summarize the output of the dropout layer. Then, an FCL with one neuron and sigmoid activation function is followed to obtain the final output. Meanwhile, we use the binary cross entropy as the loss function in the training process. We train the feature extraction module of MCFrame using the training dataset. With the trained feature extraction module of MCFrame, we remove the last layers of the feature extraction module, and the FCL with 100 neurons can return vectors of 100 features. Simultaneously, the hyperparameters of the networks are set according to Saad et al. [17].

B. Magnitude Classifier of MCFrame

We use the 100 dimensional features extracted from the trained feature extraction module as the input of the magnitude classifier. Then, we train the magnitude classifier to predict whether the final magnitude is high or low. According to relevant studies [18], [19], a magnitude of 5.5 is the boundary between low- and high-magnitude events.

Based on MCFrame, in this letter, we analyze the impact of three different magnitude classifiers on the performance of MCFrame, which include the DNN [20], SVM [21], and RF [22]as follows.

1) The DNN classifier has three FCLs with 16 neurons, eight neurons, and one neuron. The first two FCLs use the ReLU activation function, and the last FCL uses the sigmoid activation function. The ReduceLROnPlateau (with a factor of 0.1 and a patience of 5) is used to optimize the learning rate (initial learning rate of 5×10^{-5}), the early stop uses a patience of 10, the batch size is 64, the number of epochs is 200, and the Adam optimizer and binary cross entropy loss function are used in the training process of the DNN classifier.

2) The SVM classifier uses the Gaussian radial basis function. There are two parameters (cost parameter and kernel parameter), which mainly affect the performance of the SVM classifier. To obtain a high accuracy on the test dataset, based

on three threefold cross-validations, we use the grid search method, with grids ranging from 10^{-3} to 10^{2}, to optimize these two parameters.

3) For the RF classifier, according to the research of Li et al. [23], to obtain high accuracy on the test dataset, based on the three threefold cross-validation, we use the grid search method to optimize the parameters of the max tree depth and the number of decision trees. The grid range of the parameter of the max tree depth is 10–40, and the grid range of the parameter of the number of decision trees is 10–100. Other settings of the SVM and RF classifiers are the default values in sklearn [24].

Meanwhile, the magnitude classifier outputs the probabilities for a single station corresponding to the likelihood of high magnitude ($M \geq 5.5$). If the probability for output is above 0.5, the station is assigned to the high-magnitude event class. If the probability for output is below 0.5, the station is assigned to the low-magnitude event class.

C. Input of MCFrame

In this letter, the input of MCFrame is composed of three seismic waveforms collected from a single station, which are calculated based on the three components of ground motion records after P-wave arrival. The three seismic waveforms are acceleration (a), velocity (v), and displacement (d) records, which are calculated as follows:

$$\boldsymbol{a} = \sqrt{\boldsymbol{a}_{\mathbf{ud}}^2(\boldsymbol{t}) + \boldsymbol{a}_{\mathbf{ew}}^2(\boldsymbol{t}) + \boldsymbol{a}_{\mathbf{ns}}^2(\boldsymbol{t})} \quad (1)$$

$$\boldsymbol{v} = \sqrt{\boldsymbol{v}_{\mathbf{ud}}^2(\boldsymbol{t}) + \boldsymbol{v}_{\mathbf{ew}}^2(\boldsymbol{t}) + \boldsymbol{v}_{\mathbf{ns}}^2(\boldsymbol{t})} \quad (2)$$

$$\boldsymbol{d} = \sqrt{\boldsymbol{d}_{\mathbf{ud}}^2(\boldsymbol{t}) + \boldsymbol{d}_{\mathbf{ew}}^2(\boldsymbol{t}) + \boldsymbol{d}_{\mathbf{ns}}^2(\boldsymbol{t})} \quad (3)$$

where $a_{\mathrm{ud}}(t)$, $a_{\mathrm{ew}}(t)$, and $a_{\mathrm{ns}}(t)$ are the acceleration time histories in the up and down, east and west, and north and south directions, respectively; $v_{\mathrm{ud}}(t)$, $v_{\mathrm{ew}}(t)$, and $v_{\mathrm{ns}}(t)$ are the velocity time histories in the up and down, east and west, and north and south directions, respectively; $d_{\mathrm{ud}}(t)$, $d_{\mathrm{ew}}(t)$ and $d_{\mathrm{ns}}(t)$ are the displacement time histories in the up and down, east and west, and north and south directions, respectively. Meanwhile, to remove the distance effect, we normalize the epicentral distance to a reference distance of 10 km, which is a common practice used in relevant study [25]. The size of the input is (N, 3), and N is the number of sampling points in the P-wave time window (PTW).

In this letter, we use accuracy to evaluate the result of magnitude classification, and accuracy is defined as follows:

$$\text{Accuracy} = \frac{N_{\mathrm{SC}}}{N_{\mathrm{SC}} + N_{\mathrm{FC}}} \quad (4)$$

where N_{SC} is the number of samples for successful classification and N_{FC} is the number of samples for false classification.

IV. Results

In this study, we test MCFrame on Japanese strong-motion data with M3–M8 earthquakes. Meanwhile, we analyze the magnitude classification of different PTWs of 1–5 s and the impact of three different magnitude classifiers on the performance of MCFrame. Meanwhile, to verify the robustness

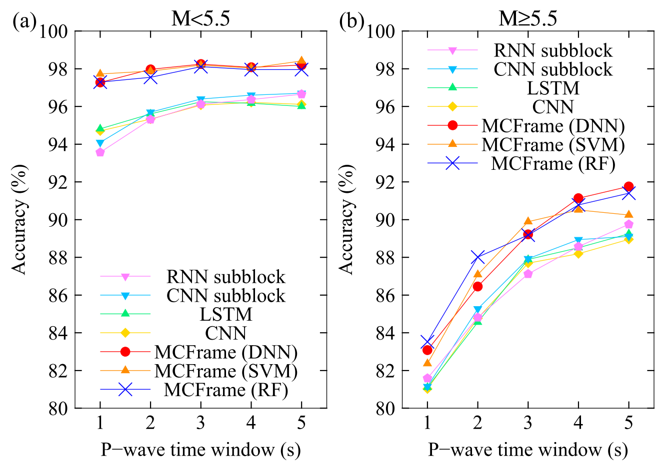

Fig. 2. Accuracy of the magnitude classification of MCFrame and baseline models on the test dataset. (a) Relationship between the accuracy of the low-magnitude event ($M < 5.5$) and the PTW. (b) Relationship between the accuracy of the high-magnitude event ($M \geq 5.5$) and the PTW.

of MCFrame, we apply MCFrame to independent earthquake events, which are not in the training and test datasets.

A. Test of MCFrame

We compare MCFrame with several baseline models for magnitude classification, such as the CNN model [6], the LSTM model [7], the CNN subblock model proposed in this work, and the RNN subblock model proposed in this work. To ensure an unbiased comparison, both the CNN model [6] and the LSTM model [7] were retrained using the same training dataset, inputs, and labels as MCFrame. This fundamentally constitutes a comparison between different network architectures, rather than a comparison between the methods proposed by each respective author. Furthermore, since the CNN model [6] and the LSTM model [7] are designed for magnitude prediction using regression, while our study focuses on magnitude classification, we modified the activation function in the last layer of the CNN and LSTM models to sigmoid activation. The CNN subblock model is the network architecture resulting from removing the RNN subblock from the CNN–RNN–AM feature extraction module of MCFrame, as shown in Fig. 1(a). The RNN subblock model is the network architecture resulting from removing the CNN subblock from the CNN–RNN–AM feature extraction module of MCFrame. For a fair comparison, both the CNN subblock model and RNN subblock model were trained using the same training dataset, inputs, and labels as MCFrame.

Fig. 2(a) shows that using the different classifiers (DNN, SVM, and RF), the accuracy of the low-magnitude event ($M < 5.5$) for MCFrame is close on the test dataset. Meanwhile, MCFrame has higher accuracy of low magnitude than the baseline models. For the MCFrame, at 1- and 3-s PTWs, the accuracy of the low magnitude exceeds 97% and 98%, respectively. Additionally, it can be seen from Fig. 2(b) that using these classifiers, the accuracy of the high magnitude ($M \geq 5.5$) for MCFrame is close on the test dataset. Meanwhile, MCFrame has higher accuracy of high magnitude than the baseline models. At 3-s PTW, the accuracy of the high

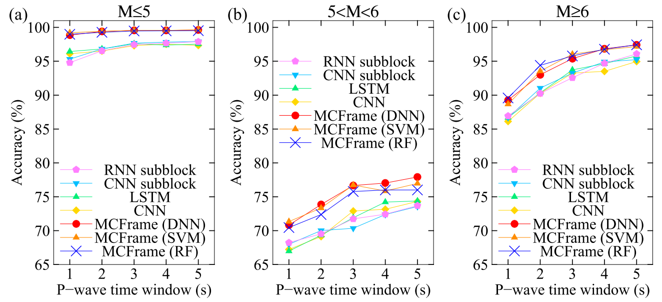

Fig. 3. Accuracy of MCFrame and baseline models for (a) $M \leq 5$, (b) $5 < M < 6$, and (c) $M \geq 6$.

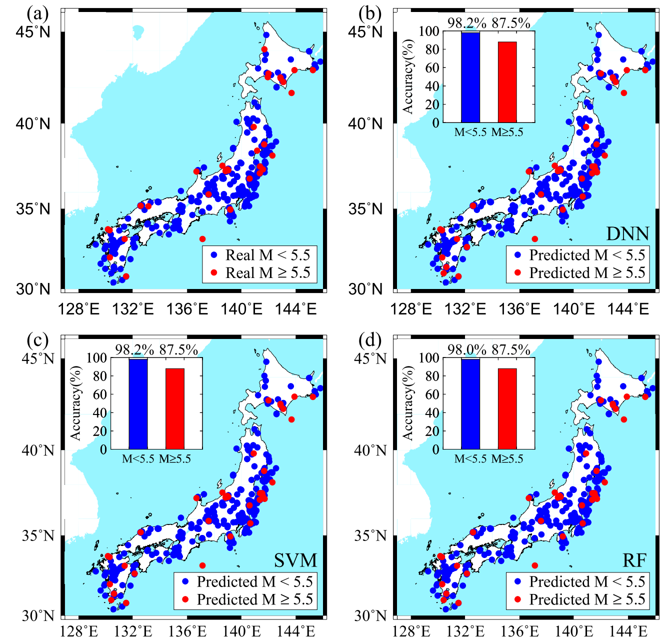

Fig. 4. At 3-s PTW, the accuracy of the magnitude classification of the MCFrame on the independent events. (a) Distribution of real low-magnitude earthquakes ($M < 5.5$) and high-magnitude earthquakes ($M \geq 5.5$). (b) Based on the DNN classifier, the prediction by the MCFrame. (c) Based on the SVM classifier, the prediction by the MCFrame. (d) Based on the RF classifier, the prediction by the MCFrame. The inset shows the accuracy of high-magnitude events and low-magnitude events.

magnitude ($M \geq 5.5$) for MCFrame exceeds 89%. Meanwhile, within 1–5-s PTWs, the accuracy of the high magnitude clearly increases with the time window increases after the P-wave arrival.

To explore the sensitivity of MCFrame to different magnitude ranges, for the test dataset, we analyze the magnitude classification accuracy of MCFrame in three different magnitude ranges, as shown in Fig. 3. We can see from Fig. 3(a) that compared with baseline models, for these three different magnitude ranges, MCFrame has higher magnitude classification accuracy. At 1-s PTW, for earthquakes with $M \leq 5$, using the different classifiers (DNN, SVM, and RF), the accuracy of $M \leq 5$ for MCFrame is close, which reaches 99%. As shown in Fig. 3(c), for earthquakes with $M \geq 6$, using the different classifiers (DNN, SVM, and RF), the accuracy of $M \geq 6$ for MCFrame is close. Meanwhile, the accuracy of $M \geq 6$ increases with increasing PTW. At 3- and 5-s PTWs, the accuracy of $M \geq 6$ for MCFrame reaches 95% and 97%, respectively. Interestingly, compared with Fig. 3(a) and (c), Fig. 3(b) shows that for earthquakes with $5 < M < 6$, there is a relatively low-magnitude classification accuracy. At 5-s PTW, the accuracy of $5 < M < 6$ for MCFrame is less than 78%.

B. Application of MCFrame to Independent Events

To verify the robustness of MCFrame, we apply MCFrame to 496 independent earthquake events (see Table S2 in the Supplementary Material) that occurred in Japan mainly between 2017 and 2019, which are not in the training and test datasets. Based on MCFrame, we analyze the accuracy of the magnitude classification of these independent events using a 3-s PTW, as shown in Fig. 4. Here, to obtain a reliable classification of an event, the average probability is obtained by averaging over all the available stations. If the average probability is more than 0.5, the event is assigned to the high-magnitude event class. If the average probability is less than 0.5, the event is assigned to the low-magnitude event class. Fig. 4 shows that when using a 3-s PTW, for these three classifiers, there is good agreement between the predicted magnitude event class and the real magnitude event class. Meanwhile, using these three classifiers, the accuracy of MCFrame for high-magnitude events reaches 87.5%. Additionally, for the DNN, SVM, and RF classifiers, the accuracies of MCFrame for low-magnitude events are 98.2%, 98.2%, and 98%, respectively.

V. Conclusion and Discussion

CNN excels in handling spatial information in images, while RNN is adept at managing temporal information. Combining the two allows the model to process multiscale information [5]. Furthermore, the attention mechanism enables the model to allocate varying attention weights to different parts of the input data. This empowers the model to focus on critical information relevant to the task, thereby enhancing its performance and generalization capabilities. In this letter, we propose a machine learning MCFrame, which utilizes a feature extraction module of a CNN–RNN–AM and magnitude classifier to predict whether an earthquake will be a high or low-magnitude event using the P-wave signal of the first few seconds after an earthquake occurs. Our results show that compared to baseline models, such as the CNN model, LSTM model, CNN subblock model, and RNN subblock model, MCFrame achieves higher accuracy in magnitude classification, which indicates that the MCFrame enhances the accuracy of magnitude classification and is significant for EEW. Furthermore, we analyze the impact of three different magnitude classifiers on the performance of MCFrame, which include DNN, SVM, and RF. Based on the extracted features from feature extraction module of MCFrame, the magnitude classification accuracy is close but not the same for these three different magnitude classifiers (DNN, SVM, and RF) of MCFrame. In other words, different classifiers did not have a significant impact on the magnitude classification accuracy for MCFrame. Meanwhile, using these three different magnitude classifiers, we infer that the slight difference in magnitude classification accuracy of MCFrame is due to the differences in parameter tuning for these three classifiers.

There is an interesting finding in Fig. 3. For earthquakes with $M \leq 5$, at 1-s PTW, the accuracy of magnitude classification reaches 99%. For earthquakes with $M \geq 6$, at 1-s PTW, the accuracy of magnitude classification is close to 90%. However, for earthquakes with $5 < M < 6$, there is relatively low accuracy. At 5-s PTW, for earthquakes with $5 < M < 6$, the accuracy of magnitude classification is less than 80%. Because the boundary between low and high magnitudes is M5.5 in this letter, we infer that the characteristics of high-magnitude events and low-magnitude events near the boundary are similar, which makes it difficult for MCFrame to classify high-magnitude events and low-magnitude events.

Early earthquake magnitude prediction research has usually involved using regression algorithms to estimate the final magnitude of event [8]. Moreover, Saad et al. [8] established a magnitude prediction model based on the vision transformer (ViT) network. In this study, a classification algorithm was employed to predict whether an event corresponds to high or low magnitudes. It is worth noting that earlier research has indicated that ViT networks based on regression algorithms can accurately predict low-magnitude events [8]. However, ViT networks tend to underestimate the magnitude of high-magnitude events. In future research, to further enhance the accuracy of EEW magnitude estimation, it might be beneficial to consider designing high-performance ViT magnitude estimation networks separately for low-magnitude and high-magnitude events. Meanwhile, integrating the ViT network with the MCFrame proposed in this letter could be explored. After an earthquake occurs, initially predicting whether the event is high or low magnitude, and then using the corresponding ViT network for magnitude estimation, would be a valuable approach.

Acknowledgment

All data are available in the main text or the supplementary materials. The National Research Institute for Earth Science and Disaster Prevention (NIED), Japan, provided the K-NET station strong-motion data (http://www.kyoshin.bosai.go.jp/).

References

[1] R. M. Allen and D. Melgar, "Earthquake early warning: Advances, scientific challenges, and societal needs," *Annu. Rev. Earth Planet. Sci.*, vol. 47, no. 1, pp. 361–388, May 2019.

[2] A. Zollo, O. Amoroso, M. Lancieri, Y.-M. Wu, and H. Kanamori, "A threshold-based earthquake early warning using dense accelerometer networks," *Geophys. J. Int.*, vol. 183, no. 2, pp. 963–974, Nov. 2010.

[3] Y. Wu and L. Zhao, "Magnitude estimation using the first three seconds P-wave amplitude in earthquake early warning," *Geophys. Res. Lett.*, vol. 33, no. 16, Aug. 2006, Art. no. L16312.

[4] S. Colombelli, A. Zollo, G. Festa, and M. Picozzi, "Evidence for a difference in rupture initiation between small and large earthquakes," *Nature Commun.*, vol. 5, no. 1, p. 3958, Jun. 2014.

[5] S. M. Mousavi and G. C. Beroza, "A machine-learning approach for earthquake magnitude estimation," *Geophys. Res. Lett.*, vol. 47, no. 1, Jan. 2020, Art. no. e2019GL085976.

[6] J. Zhu, S. Li, J. Song, and Y. Wang, "Magnitude estimation for earthquake early warning using a deep convolutional neural network," *Frontiers Earth Sci.*, vol. 9, May 2021, Art. no. 653226.

[7] A. Cofré et al., "End-to-end LSTM-based earthquake magnitude estimation with a single station," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[8] O. M. Saad, Y. Chen, A. Savvaidis, S. Fomel, and Y. Chen, "Real-time earthquake detection and magnitude estimation using vision transformer," *J. Geophys. Res., Solid Earth*, vol. 127, no. 5, May 2022, Art. no. e2021JB023657.

[9] O. M. Saad et al., "EQCCT: A production-ready earthquake detection and phase-picking method using the compact convolutional transformer," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4507015.

[10] S. M. Mousavi, W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza, "Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking," *Nature Commun.*, vol. 11, no. 1, p. 3952, Aug. 2020.

[11] O. M. Saad et al., "Machine learning for fast and reliable source-location estimation in earthquake early warning," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[12] T. Perol, M. Gharbi, and M. Denolle, "Convolutional neural network for earthquake detection and location," *Sci. Adv.*, vol. 4, no. 2, Feb. 2018, Art. no. e1700578.

[13] W. Zhu, S. M. Mousavi, and G. C. Beroza, "Seismic signal denoising and decomposition using deep neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9476–9488, Nov. 2019.

[14] O. M. Saad, Y. Chen, A. Savvaidis, W. Chen, F. Zhang, and Y. Chen, "Unsupervised deep learning for single-channel earthquake data denoising and its applications in event detection and fully automatic location," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 3209932.

[15] J. Zhu, S. Li, and J. Song, "Hybrid deep-learning network for rapid on-site peak ground velocity prediction," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 3230829.

[16] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.

[17] O. M. Saad, A. G. Hafez, and M. S. Soliman, "Deep learning approach for earthquake parameters classification in earthquake early warning system," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 7, pp. 1293–1297, Jul. 2021.

[18] M. Böse, T. H. Heaton, and E. Hauksson, "Real-time finite fault rupture detector (FinDer) for large earthquakes," *Geophys. J. Int.*, vol. 191, no. 2, pp. 803–812, Nov. 2012.

[19] J. Cheng, Y. Rong, H. Magistrale, G. Chen, and X. Xu, "Earthquake rupture scaling relations for Mainland China," *Seismolog. Res. Lett.*, vol. 91, no. 1, pp. 248–261, Jan. 2020.

[20] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[21] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 2008.

[22] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

[23] Z. Li, M. Meier, E. Hauksson, Z. Zhan, and J. Andrews, "Machine learning seismic wave discrimination: Application to earthquake early warning," *Geophys. Res. Lett.*, vol. 45, no. 10, pp. 4773–4779, May 2018.

[24] F. Pedregosa et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Jan. 2012.

[25] W. Kuang, C. Yuan, and J. Zhang, "Network-based earthquake magnitude determination via deep learning," *Seismolog. Res. Lett.*, vol. 92, no. 4, pp. 2245–2254, Jul. 2021.