APPLIED GEOPHYSICS, Vol.21, No.4 (December 2024), P. 1–12, 9 Figures.
DOI:10.1007/s11770-024-1117-4

1D Convolutional Seismic Event Classification Method Based on Attention Mechanism and Light Inception Block

Yong-ming Huang[1]*, Yi Xie[1], Fa-jun Miao[2], Yong-sheng Ma[1], Gao-chuan Liu[3], Guo-bao Zhang[1], and Yun-tian Teng[3]

Abstract: Waveforms of artificially induced explosions and collapse events recorded by the seismic network share similarities with natural earthquakes. Failure to identify and screen them in a timely manner can introduce confusion into the earthquake catalog established using these recordings, thereby impacting future seismological research. Therefore, the identification and separation of natural earthquakes from continuous seismic signals contribute to the monitoring and early warning of destructive tectonic earthquakes. A 1D convolutional neural network (CNN) is proposed for seismic event classification using an efficient channel attention mechanism and an improved light inception block. A total of 9937 seismic sample records are obtained after waveform interception, filtering, and normalization. The proposed model can obtain better classification performance than other major existing methods, exhibiting 96.79% overall classification accuracy and 96.73%, 94.85%, and 96.35% classification accuracy for natural seismic events, collapse events, and blasting events, respectively. Meanwhile, the proposed model is lighter than the 2D convolutional and common inception networks. We also apply the proposed model to the seismic data recorded at the University of Utah seismograph stations and compare its performance with that of the CNN-waveform model.
Keywords: Attention mechanisms; Seismic classification; CNNs; Raw seismic waveform.

Introduction

In recent years, the field of seismology has attracted extensive research (Lu et al., 2019; Zhu et al., 2019a; Elsayed et al., 2022). The monitoring capability of seismic station networks has significantly improved with the increase in station density and the advancement of digital seismic observation systems, which can monitor all types of seismic events. The scientific nature of seismic hazard analysis and seismicity research will be significantly impacted if the natural seismic catalog is tainted due to seismic misidentification. The rapid identification of seismic events has great practical importance and is one of the hot topics of seismological research.

In the past, many scholars have studied seismic identification technology, and the identification methods can be mainly divided into two categories. One is based on the source mechanism with clear geophysical

Manuscript received by the Editor October 17, 2023; revised manuscript received June 13, 2024.
This research is supported by the Jiangsu Provincial Key R&D Programme 261 (BE2020116, BE2022154).
1. School of Automation, Southeast University, Jiangsu 210096, P. R. China
2. Seismological Bureau of Jiangsu Province, Jiangsu 210096, P. R. China
3. Institute of Geophysics, China Earthquake Administration, Beijing 100081, P.R. China
*Corresponding author: Yong-ming Huang (E-mail: huang_ym@seu.edu.cn).

significance. The other is centered on the numerical theory of statistical inference and signal analysis of waveforms. After years of research, various automatic earthquake identification methods have emerged, including those based on Akaike information criteria and on short-term average/long-term average (Zhao et al., 2019). Each method has its own pros and cons. Rational use of seismic recognition methods can classify seismic events and noise, preventing seismic event waveforms from being contaminated by noise fingerprints. In addition, similarity threshold methods (Yoon et al., 2015) can identify seismic events in continuous data.

With the rapid development of artificial neural networks, artificial intelligence has emerged. Thus, more scholars at home and abroad apply deep learning neural network techniques in the study of earthquakes. In China, Bian (2002) used the traditional BP neural network structure to classify seismic and blast events after training a feature dataset with log spectral amplitude values of P-wave and S-wave spectra. Ren et al. (2019) used a Bagging machine learning algorithm to classify natural and nonnatural earthquakes. A convolutional neural network (CNN) is a well-known deep learning method with powerful image recognition capability. The common CNN oil consists of a convolutional layer, a pooling layer, a fully connected layer, and an output layer, which learns the feature information of the input by training a large amount of data to recognize new input samples. Perol et al. (2018) proposed the first CNN ConvNetQuake using earthquake waveform data. Duan (2021) proposed a 13-layer CNN-Epq13 for natural and artificial blast earthquake identification. Chen Runhang et al. (Chen et al., 2018) used a CNN to classify seismic waveform signals from Mel frequency cepstral coefficient maps. Sugiyama et al. (2021) applied 3D convolution to seismic data for earthquake localization. Tian et al. (2022) identified natural earthquakes and blasts by multiple inputs in time and frequency domains. Ku et al. (2021) proposed a CNN with a spatial attention mechanism to classify seismic events, obtaining excellent results. Jin et al. (2024) proposed CNN_BAM that combines an improved visual geometry group with a convolutional block attention module for recognizing and classifying microseismic data. Numerous studies have shown that deep learning-based models perform effectively in the earthquake domain (Nakano et al., 2019, Trani et al., 2022, Lim et al., 2022, Wang et al., 2023).

CNNs are widely used in deep learning algorithms with excellent results. Moreover, the CNN model in earthquake studies performs efficiently, indicating that they have great potential for earthquake monitoring applications.

Table 1. Survey of seismic classification research. (√ denotes covered and × indicates not covered. CA + SI is channel attention that considers spatial information.)

Research	Raw waveform input	Light weight	CA + SI	Multidatasets
Perol et al., 2018; Duan, 2021; Tian et al., 2022; Trani et al., 2022	√	×	×	×
Chen et al., 2018; Sugiyama et al., 2021; Nakano et al., 2019	×	×	×	×
Ku et al., 2021	√	×	√	×
This paper	√	√	√	√

In the previous event classification methods, waveform data and spectrograms are usually used as input data. Nakano et al. (2022) compared the performance of 1D and 2D CNNs, and their experimental results show that producing the time–frequency domain representation is unnecessary for high-performance classification. Thus, we select 1D CNNs, which are simpler than 2D CNNs, to construct the network. Residual network (ResNet) (He et al., 2016) protects the integrity of the information by bypassing the input directly to the output. Consequently, the network becomes easier to optimize, alleviating the problem of vanishing gradients. The inception block allows for the aggregation of visual information at different sizes while initially performing dimensionality reduction on larger-sized matrices, facilitating the extraction of features from different scales (Liu et al., 2024). Therefore, to improve the classification accuracy and control the model complexity, the efficient channel attention network (ECA-Net) and light inception block are added to the 1D CNN model. The attention mechanism can render the feature map more discriminative, allowing the model to focus more on the essential features. The network structure is optimized using the inception structure for seismic waveform data, reducing the number of parameters.

Table 1 shows the results of seismic classification research. The contributions of our work are as follows:

(i) We proposed a 1D CNN model for seismic event classification using ECA-Net and light inception block. We modified the lightweight channel attention mechanism to enhance the spatial information. In addition, we optimized the commonly used multiscale convolutional network to achieve a lighter architecture.

(ii) We evaluated the proposed model with the collected dataset, and it achieved high test accuracy and low resource consumption.

(iii) The proposed model performed efficiently using the seismic data recorded at the University of Utah seismograph stations.

Methods

In this section, we focus on the model structure, including ECA-Net, light inception block, and the overall structure.

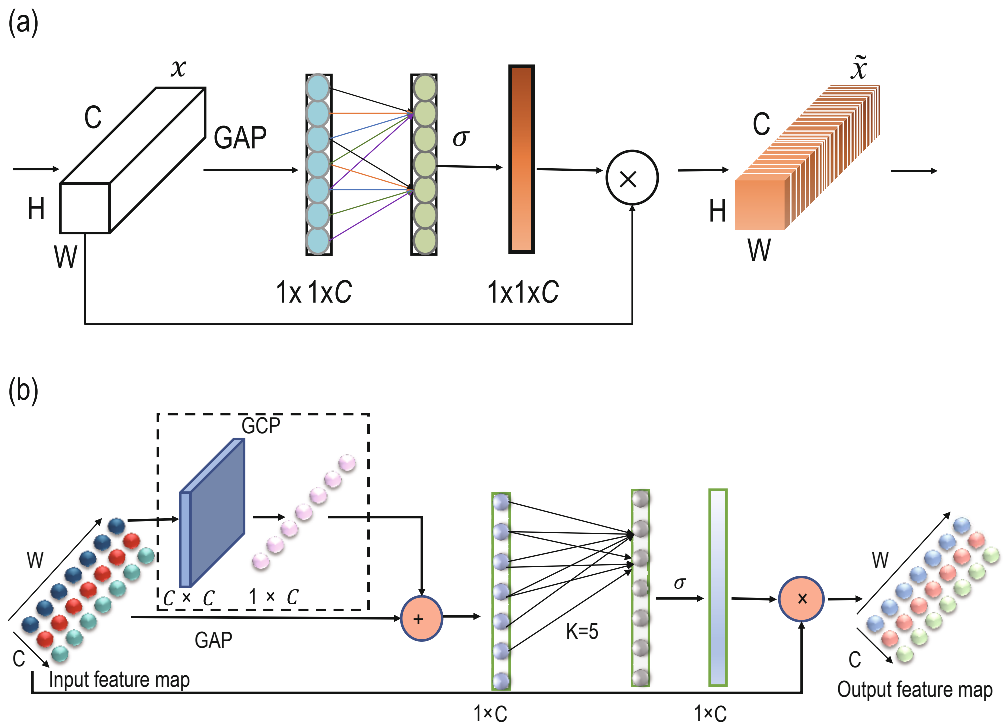

Fig.1. (a) ECA module (b) improved ECA module.

Attention

According to the importance of each input channel, SENet (Hu et al., 2020) can enhance the important channels and suppress the unimportant channels, constituting a channel attention mechanism. ECA-Net uses a 1D convolution with an adaptive kernel size of K (Fig. 1(a)), which represents the coverage of local cross-channel interactions instead of the fully connected layer in SENet. ECA-Net has been proven to perform favorably while benefiting from much lower model complexity (Wang et al., 2020).

The ECA-Net was modified to become more suitable for time series data. In the original ECA-Net, the global average pooling, where all feature values in the channels are assigned an identical weight, is used for the squeeze operation. We proposed to combine global average pooling and global covariance pooling (Dai et al., 2019) for capturing global information in the squeeze operation to improve ECA-Net (Fig. 1(b)), merging global statistical modeling. Global covariance pooling (GCP) that is used to aggregate the information of deep CNNs has achieved remarkable performance gains on various vision tasks (Fukui et al., 2016; He et al., 2016; Szegedy et al., 2015). GCP can select values that can represent the data distribution of the feature graph by calculating the covariance matrix (second-order information) of the feature graph. The covariance matrix describes the correlation between the feature graphs f_i (i=1,⋯,C) in the channel dimension. The correlation between channels is calculated in pairs to obtain the covariance matrix, which is defined as Equation (1), where conv(f_C, f_C) is the covariance calculation between paired feature graphs. The feature map vectors of all channels are merged into a feature map matrix X of size W×C. The covariance matrix is calculated using Equation (2), where *I* is an all-1 matrix with the same size as *X*.

$$\Sigma=\begin{bmatrix} conv(f_1,f_1) & \cdots & conv(f_1,f_C) \\ \vdots & \ddots & \vdots \\ conv(f_C,f_1) & \cdots & conv(f_C,f_C) \end{bmatrix}. \tag{1}$$

$$\Sigma=X^{T}\left(\frac{1}{W}-\frac{1}{W^{2}}\right)X. \tag{2}$$

The covariance matrix is orthogonally decomposed, and then the costandard deviation matrix $\Sigma^{1/2}$ of size C×C is obtained using Equation (3), where Λ is the diagonal matrix formed by the eigenvalues of Σ, and y_C represents the set of the costandard deviation of the feature graph of the c channel against the feature graph of other channels. A quantity that can better represent the data distribution of the feature graph of each channel is the average value of each component of its costandard deviation vector. The output of the global covariance pool is the tensor of C×1. Then, the result of the global average pooling is integrated to obtain the global information of the feature map.

$$\Sigma^{1/2}=U\Lambda^{1/2}U^{T}=[\ _{1},\ _{2},\cdots,\ ]. \tag{3}$$

The important features are considered, and those that are nonbeneficial to the current task are suppressed, facilitating the network's learning of the seismic data.

Light inception block

The inception model can reduce parameters while increasing the depth and width of the network. Moreover, the inception V1 model (Szegedy et al., 2015) uses a 1 × 1 convolution kernel for lift-dimensioning and simultaneous convolutional reaggregation at multiple sizes. Related improvements mainly include inception V2 (Ioffe, Szegedy, 2015), which incorporates batch normalization; inception V3 (Szegedy et al., 2016b), which factorizes convolutions with small filter size; and inception V4 (Szegedy et al., 2016a), which incorporates residual networks.

A general inception network (Fig. 2 (a)) and a residual network structure (Fig. 2 (b)) are used for feature extraction of 2D data. In this study, 1D CNNs, which are more suitable for time series data, are selected to replace 2D CNNs. Referring to the inception network, the pooling branch is subtracted from the cascade path, and a branch is added to improve the ability of the network to extract multiscale features and capture more seismic features. In addition, establishing a light inception block after changing the multibranch structure into a linear structure (Fig. 3) reduces the risk of overfitting caused by excessive parameters. The input data undergo 1×1 and two 1×3 1D convolutional kernel processes, and then the feature map is spliced to obtain $X_4 = C(X_1, X_2, X_3)$, where C represents the concatenation operation. After changing its feature dimension by using a 1×1 1D convolution kernel, the feature map is added to the input ($X_6 = X_5 + X_0$) and activated by RELU to output the multiscale feature information of the light inception block. The features of input data are extracted through the convolution kernel.

After aggregation, the approach is equivalent to combining the "collective knowledge" of the previous several layers, and then a convolution layer is used to extract the features. The input is connected with the

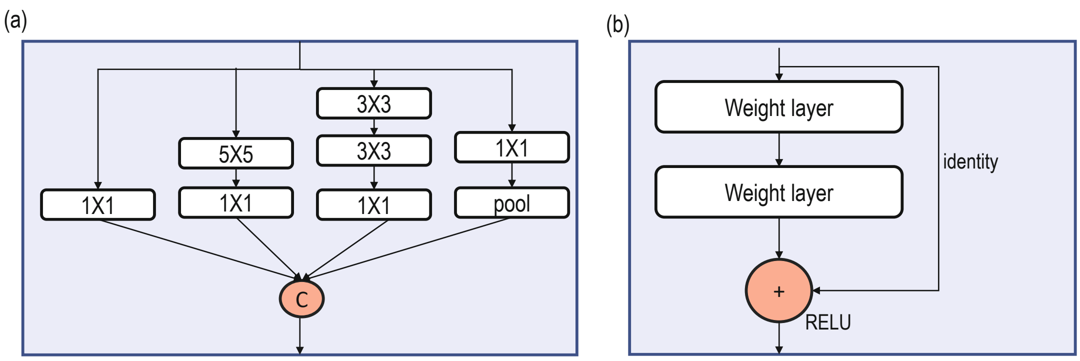

Fig.2. (a) Inception block (b) residual block.

output through identity mapping, and it is reinvested into the decision space after RELU activation, thereby accelerating the gradient propagation. The block combines residual network and inception network to maintain a low model complexity, extracts as many features of input information as possible, improves the retention rate of low-dimensional information, and enables easy optimization of the network.

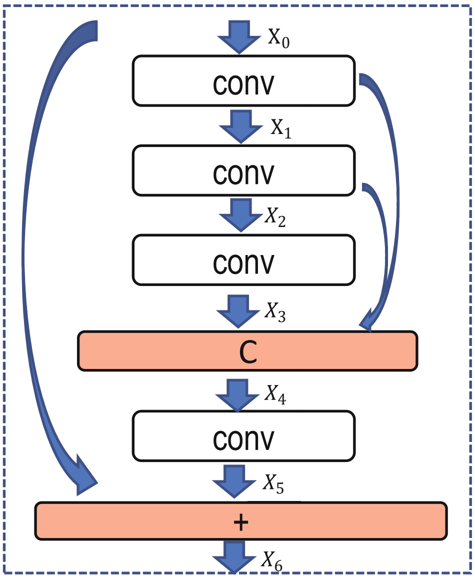

Fig.3. Light inception block.

Network structure

$$(\ ) = \sum_{=0} (\ - \ ) * (\ ) \tag{4}$$

The core of 1D CNN is to extract features along the temporal dimension of the seismic time series data. The size of the convolution kernel determines the range of the receptive field, and the fixed temporal neighborhood information can be sensed and encoded during feature extraction. The convolution is the fundamental layer that extracts feature maps by introducing filters with the size of L, as shown in Equation (4), where z represents the convolutional input, w denotes the filter with the size of L, and o indicates the output feature map.

Aiming at the problem of seismic event classification, the three-channel waveform is considered the input and the seismic classes as the output to train the CNN. The CNN extracts the inherent features of the waveform through the convolution kernel, sets multiple convolutional layers, extracts more features, integrates the above feature information, and finally outputs the probability of the seismic classes of the object by a similar voting method. Dropout regularization suppresses overfitting, potentially enhancing the stability and performance of the proposed model. In the training stage, the model uses the attention mechanism to suppress the irrelevant features but focuses on the significant features of the target. The approach does not require cutting the regions of interest between the networks, which is conducive to data processing. The features are then captured at multiple scales using lightweight inception blocks, and the combination of attention mechanisms and lightweight inception blocks assists the network in learning deep information about the data.

The network structure is shown in Fig. 4. A convolution layer with a width of 7 in the shallow layer of the convolutional network was replaced with a light inception block to improve the extraction ability of multiscale features. In addition to the light inception block, the other convolution sets include the convolution layer (the number of channels is 64, the length of the convolution kernel is 7, L2 regularization), the max pooling layer, and the dropout layer (dropout rate 0.2).

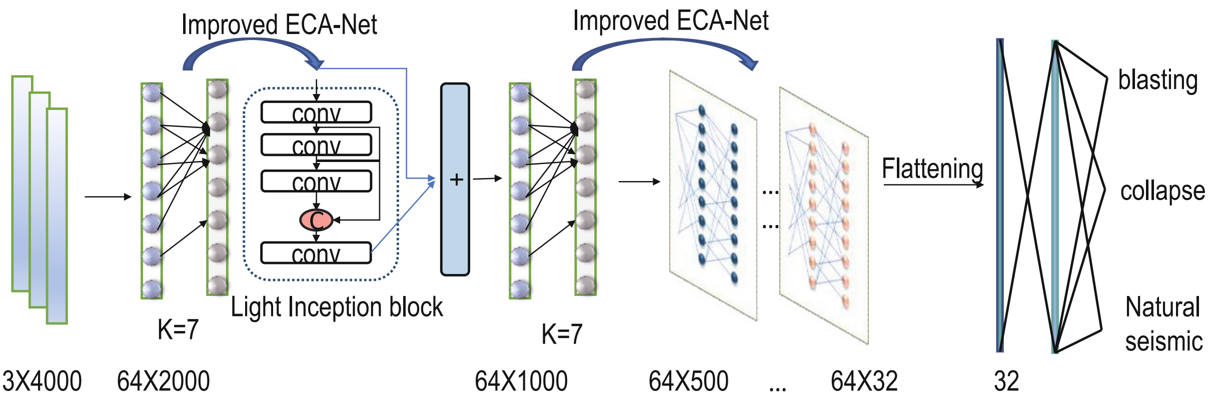

Fig.4. Seismic event classification of CNN structure.

These layers are activated by RELU through the flattened layer, followed by two fully connected layers (number of neurons 32, 3), and classified by softmax output seismic events.

Experiments

In this section, we performed experiments to investigate the effectiveness of the proposed CNN architectures with attention module and light inception block.

Dataset

In this section, we focus on the input data used in this study. The nonnatural seismic events in Jiangsu and its adjacent areas since 2015 are selected as input data, as well as the seismic events from the earthquake catalogs of China's capital area (Beijing, Hebei, Tianjin, Liaoning, and Shanxi) since 2008. Then, the arrival time, amplitude, and event type of the recorded waveforms within 200 km of the epicenter are labeled. In addition, natural seismic events with epicenter error less than 50 km and magnitude between 2.0 and 4.0 were collected (Fig. 5). A total of 9937 sets of three waveform samples were obtained (including 5337 natural seismic events, 2794 collapse events, and 1806 blasting events). The input data were divided into two independent sets, namely the test set and the training set, at a 1:4 ratio. Waveforms were sampled at 100 Hz on three channels corresponding to the three spatial dimensions, including BHE (oriented west–east), BHN (oriented north–south), and BHZ (oriented vertically).

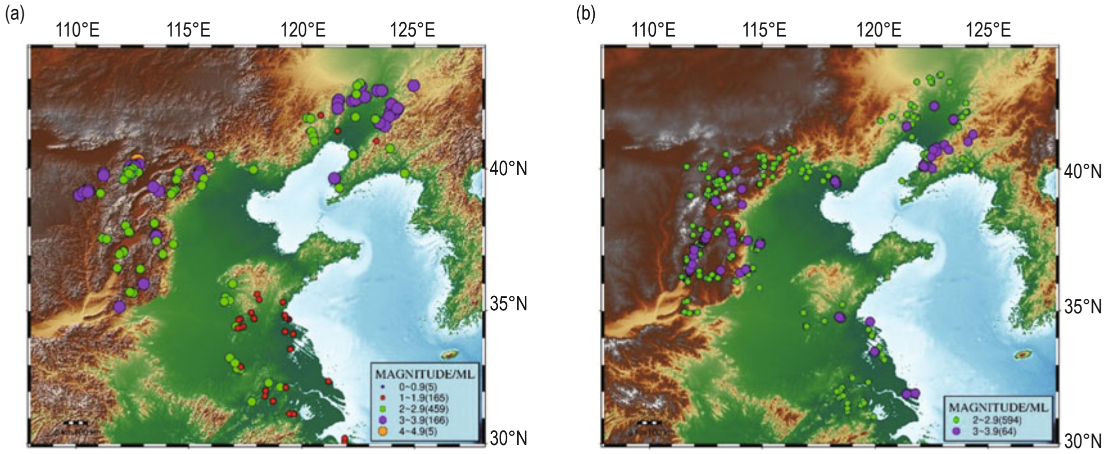

Fig.5. Map of (a) natural earthquakes and (b) unnatural earthquakes.

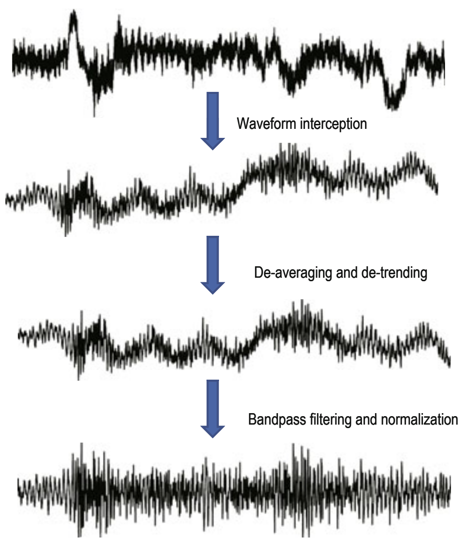

Fig.6. Pre-processing steps for input data, including five steps.

For each waveform, 5 seconds before and 35 seconds after the arrival of seismic P-wave were intercepted (40 seconds in total). After de-averaging, de-trending, and band-pass filtering of 1–30 Hz, these data are normalized to ensure that the input data were in the 0–1 range (the pre-processing steps were shown in Fig.6). Sub-headings should be typeset in boldface italic and capitalize the first letter of the first word only. Section number to be in boldface roman.

Performance evaluation

In this section, we discuss four topics: the selection of attention mechanisms, the ablation study, the evaluation of our methods compared with other classifier methods, and the evaluation of model complexity. The Adam algorithm was adopted in the training process, and the cross-entropy loss function was used to optimize the network parameters. The learning rate was set at 0.001, and the number of training samples in each batch was set at 32.

Selection of attention mechanisms

We initially discuss the selection of attention mechanisms. Soft attention mechanism includes channel domain attention (SENet, ECA-Net), spatial domain attention (SAM), and mixed domain attention (CBAM (SE + SAM)). To select a more suitable attention mechanism for seismic waveform classification, SENet (Hu et al., 2020), ECA-Net (Wang et al., 2020), SAM (Zhu et al., 2019b), and CBAM (SE + SAM) (Woo et al., 2018) were added to the model for experimental comparison. The classification accuracy of seismic signals by networks with different attention mechanisms is shown in Table 2.

Table 2. Accuracy of seismic event classification methods incorporating different attention mechanisms

Methods Indicators	Accuracy (%)	Recall (%)	F1 score
SAM	91.86	91.67	0.9179
CBAM	91.45	91.67	0.9157
SE	93.35	93.21	0.9328
ECA	93.67	93.78	0.9375

Fig. 7 and Table 2 illustrate that the channel attention mechanism is more suitable for seismic classification tasks. Compared with the SENet-integrated model, the accuracy and recall rates of the model integrated with the ECA-Net are increased by 0.32% and 0.57%, respectively. The addition of the ECA-Net improves the network classification effect.

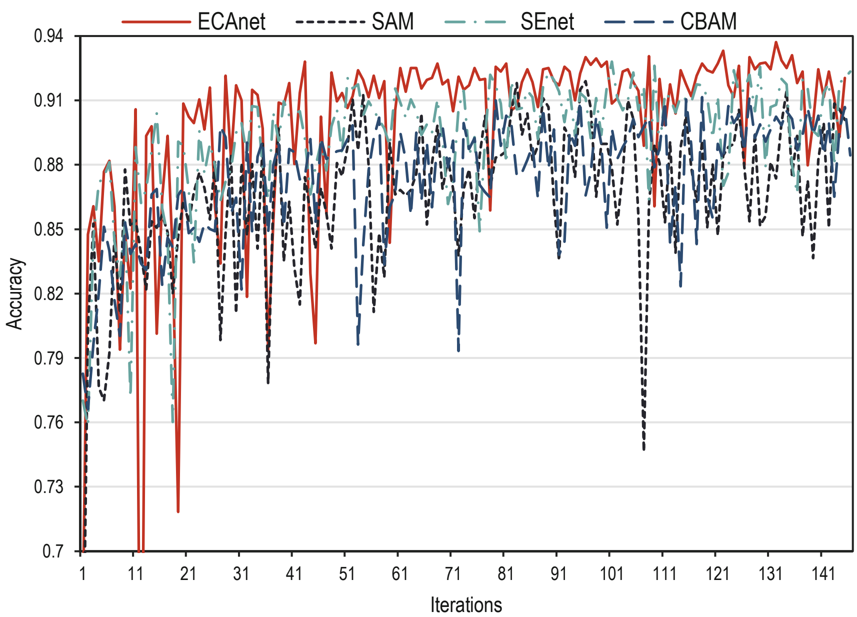

Fig.7. Performance comparison of different attention mechanisms.

Ablation study

Three sets of comparison experiments are conducted using the framework of a six-layer CNN, demonstrating that various structural improvements and fusion strategies can effectively enhance the segmentation effect of the algorithm. The results are shown in Table 3, where boldface is the optimal value of the same index. Fig. 8 shows the comparison of the training performance of the ablation study. Table 3 shows that the addition of the attention mechanism and light inception block improved the network accuracy to a certain extent and reduced the loss function. Moreover, the light inception network played a greater role in improving the performance of the proposed 1D CNN.

Table 3. Result of the ablation study using the proposed model as the backbone in terms of classification accuracy, recall, and F1 score.

Network Structure	Accuracy (%)	Recall (%)	F1 score
Without ECA-Net	92.69	91.67	0.9192
Without light inception block	91.14	91.01	0.91
Without GCP	94.67	93.78	0.9445
Model in this paper	96.79	96.87	0.9652

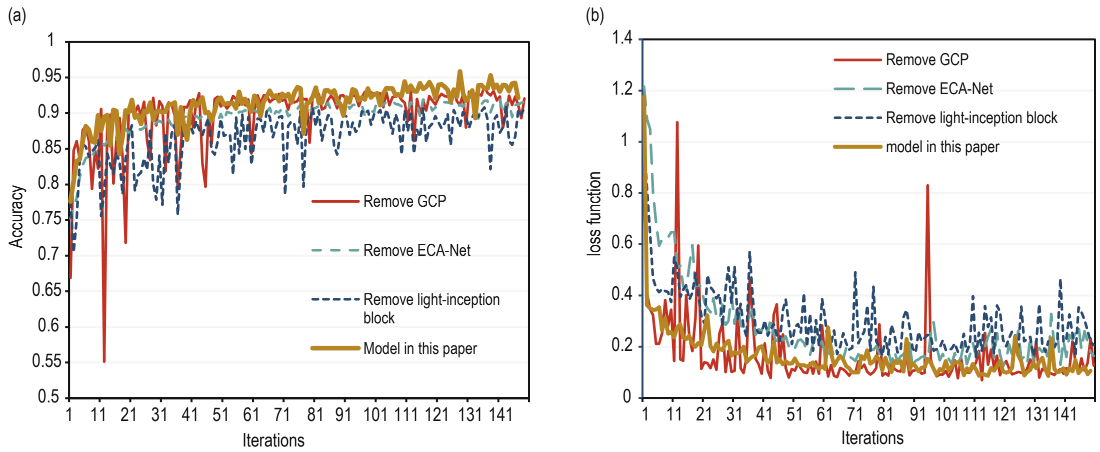

Fig.8. Comparison of the training performance in the ablation study: (a) accuracy curve and (b) loss function curve.

Evaluation of our methods compared with other classifier methods

To measure the performance difference between the proposed CNN model and other methods under the same experimental conditions and input data., the traditional classifier K-neighbor (KNN), support vector machine (SVM) (Tang et al., 2020), multilayer perceptron (MLP) (Laasri et al., 2013), CNN-waveform (Tian et al., 2022), Deepquake (Trani et al., 2022), and ConvNetQuake (Perol et al., 2018) are considered for reference. The accuracy pairs of each classification are shown in Table 4. The table illustrates that traditional classifiers SVM, MLP, and KNN have poor classification effects. Among them, our model improved by 2.9% in terms of overall accuracy compared with CNN-waveform; it also exhibited improvements of 0.28%, 5.25%, and 2.93% in single-class classification accuracy for natural earthquakes, blasting, and collapse. Our model performs more on the event classification of blasting and collapse than on natural earthquakes. The reason may lie in the fact that the attention mechanism and the lightweight inception module are more suitable for capturing the event features in blasting and collapse. In addition, the experimental results suggest that the recognition rate of natural seismic events is generally higher than that of blasting and collapse events. Thus, the large amount of natural seismic data cannot be ruled out.

Table 4. Comparison with other algorithms using the same dataset in terms of classification accuracy (NE is a natural earthquake, and AAR is the average accuracy rate).

Classifier methods	Year	NE	Blasting	Collapse	AAR
KNN	-	78.37	71.91	68.98	74.85
MLP (Laasri et al., 2013)	2013	74	68	51	70
ConvNetQuake (Perol et al., 2018)	2018	94.5	89.32	85.43	91.7
SVM (Tang et al., 2020)	2020	76.17	64.74	72.57	72.16
CNN-waveform (Tian et al., 2022)	2022	96.45	89.6	93.42	93.89
Deepquake (Trani et al., 2022)	2022	94.66	89.74	89.52	92.40
This paper	2024	96.73	94.85	96.35	96.79

Evaluation of model complexity

We compare the training parameters of each network. The number of network parameters can be used to evaluate the complexity and speed of the model. Network parameters can be divided into trainable parameters and untrainable parameters, which can reflect the memory size occupied.

The six-layer CNN 2D CNN, MLP, Deepquake (Trani et al., 2022), CNN-waveform (Tian et al., 2022), the SENet structure, and the inception V3 network structure of the neural network were added as references.

As shown in Table 5, the MLP network requires more parameters to be trained than the proposed CNN. The number of parameters and calculation amount is one-fifth of the input six-layer 2D CNN (224, 224, 3) and approximately one-tenth of the three-layer MLP network. The network in Fig. 2 (a) is used to replace the light inception block in this study, causing a parameter increase. Moreover, the SENet network requires more training parameters. A 1D CNN with a light inception block has a simpler network structure, less resource consumption, and lower requirements on the equipment.

Table 5. Comparison in terms of parameter, trainable parameter.

Network Structure	Parameter	Trainable parameter	Untrainable parameter
Six-layers 2D CNN	819404	819148	256
Three-layers MLP	1544963	1544963	0
Deepquake	301571	301571	0
CNN-waveform	8271331	8271331	0
SENet	166095	165839	256
Inception structure	263900	261660	256
This paper	160017	159889	128

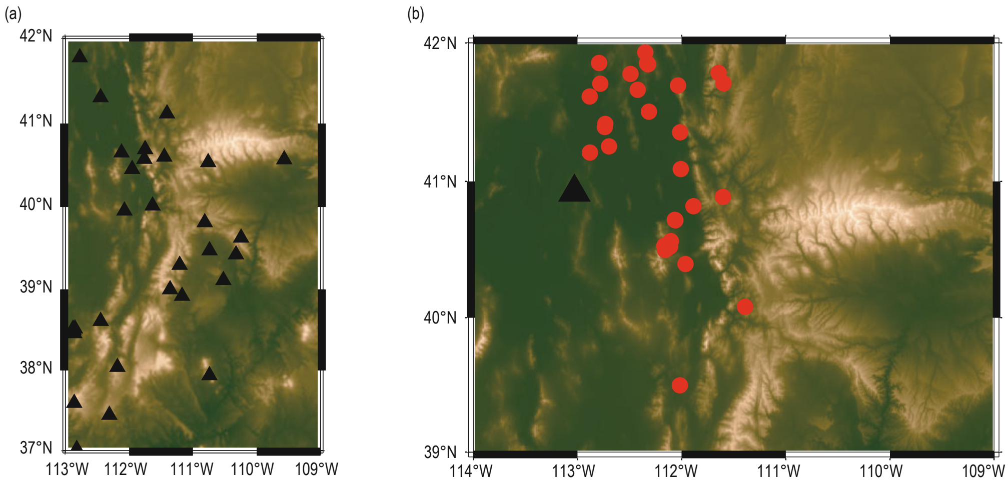

Fig.9. (a) Map of seismic monitoring stations (b) the BGU station and the distribution of earthquakes it monitored (triangles represent stations, red circles represent detected seismic events).

Discussion

We evaluated the classification performance of the proposed network using seismic data recorded at the University of Utah seismograph stations. Then, we compared it with the CNN-waveform model (Tian et al., 2022). Utah and its surrounding areas are considered the research areas (115°W-108°W, 36°N-43°N). According to the earthquake catalog provided by Linville (2019), we downloaded 2028 local seismic events and 1169 quarry blast events recorded by the University of Utah seismograph station during 2013–2017 from IRIS. The data acquired were from 39 stations (Fig. 9). The magnitude range is −9.99–3.48 ML, and the epicenter distance is 0.05–240 km. Based on the P-wave arrival times selected from the earthquake catalog, the waveforms of the P-wave arrival times and the 40s after arrival were intercepted and processed. Comparing the classification effectiveness of the network using the single station waveform of an event as input data, the accuracy of the proposed method and the CNN waveform model is 95.59% and 92.07%, respectively. The datasets of both methods are from the same region, whereas the accuracy of the proposed method is slightly higher. The proposed network also has better results on other datasets with certain generalizations.

Conclusion

In this study, a 1D CNN model integrating the attention mechanism and light inception block is built. Then, the three-channel seismic waveform data, after pre-processing, are directly fed into the network, thereby realizing the classification of the natural seismic, collapse, and blasting events. Our methods are evaluated and compared with other classifier methods, and the results show that the proposed model has a relatively simple structure, excellent classification performance, fewer training parameters, less computation, and relatively low requirements for equipment. Meanwhile, the addition of an attention mechanism and light inception block enhances the performance of CNNs. The proposed model performed efficiently when evaluated using the seismic data recorded at the University of Utah seismograph stations.

A number of recommendations for future research are provided. The neural network has a strong dependence on the training data set. Subsequently, more data must be collected, and the data set samples must be improved and optimized by adding collapse and blasting events and selecting more representative samples, etc., and the difference between varying source mechanisms, propagation paths, and absorption effects of the same source leads to distinct waveform signals. Therefore, more data sets can effectively improve the generalization ability of the system. Subsequently, we can further add the data of interference, noise, or shock and other categories, expand the categories of recognition, and re-clean the labels of the data set to reduce the risk of low recognition accuracy caused by label errors. In addition, given the data set limitations, the generalization ability of the proposed 1D CNN model must be tested and studied.

Acknowledgments

This research was financially supported by the Jiangsu Provincial Key R&D Programme 261 (BE2020116, BE2022154). Thanks to Jiangsu Seismograph Network for providing the seismic data.

References

Bian, Y., 2002, Application of Genetic BP network to discriminating earthquakes and explosions: Acta Seismologica Sinica, **24**(5), 516−524.

Chen, R., Huang, H.andChai, H., 2018, Study on the discrimination of seismic waveform signals between earthquake and explosion events by convolutional neural network: Progress in Geophysiscs, **33**(4), 1331−1338.

Dai, T., Cai, J., Zhang, Y., et al., 2019, Second-Order Attention Network for Single Image Super-Resolution, 11057−11066.

Duan, G., 2021, Research on identification of natural earthquake and artificial blasting based on convolutional neural network: Progress in Geophysiscs, **36**(4), 1379−1385.

Elsayed, H. S., Saad, O. M., Soliman, M. S., et al., 2022, Attention-Based Fully Convolutional DenseNet for Earthquake Detection: Ieee Transactions on Geoscience and Remote Sensing, **60**.

Fukui, A., Park, D. H., Yang, D., et al., 2016, Multimodal compact bilinear pooling for visual question answering and visual grounding: arXiv preprint arXiv:160601847.

He, K., Zhang, X., Ren, S., et al., 2016, Deep Residual Learning for Image Recognition:2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770−778.

Hu, J., Shen, L., Albanie, S., et al., 2020, Squeeze-and-Excitation Networks: IEEE Transactions on Pattern Analysis and Machine Intelligence, **42**(8), 2011−2023.

Ioffe, S.andSzegedy, C., 2015, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift: Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume **37**, 448–456 , numpages = 449.

Jin, S., Zhang, S. C., Gao, Y., et al., 2024, Microseismic

Event Recognition and Transfer Learning Based on Convolutional Neural Network and Attention Mechanisms: Applied Geophysics.

Ku, B., Kim, G., Ahn, J. K., et al., 2021, Attention-Based Convolutional Neural Network for Earthquake Event Classification: Ieee Geoscience and Remote Sensing Letters, **18**(12), 2057−2061.

Laasri, E. H. A., Akhouayri, E.-S., Agliz, D., et al., 2013, Seismic Signal Classification using Multi-Layer Perceptron Neural Network: International Journal of Computer Applications, **79**, 35−43.

Lim, J., Jung, S., JeGal, C., et al., 2022, LEQNet: Light Earthquake Deep Neural Network for Earthquake Detection and Phase Picking: Frontiers in Earth Science, **10**.

Linville, L., Pankow K, Draelos T, 2019, Deep Learning Models Augment Analyst Decisions for Event Discrimination: Geophysical Research Letters, **46**(7), 3643−3651.

Liu, W., Huang, Y., Lu, Y., et al., 2024, Seismic events classification based on Grams angle field and multi-scale residual neural network: Acta Seismologica Sinica, **46**(1), 69−80.

Lu, Y., Huang, Y., Xue, W., et al., 2019, Seismic data processing method based on wavelet transform for de-noising: Cluster Computing, **22**.

Nakano, M.andSugiyama, D., 2022, Discriminating seismic events using 1 D and 2D CNNs: applications to volcanic and tectonic datasets: Earth Planets and Space, **74**(1).

Nakano, M., Sugiyama, D., Hori, T., et al., 2019, Discrimination of Seismic Signals from Earthquakes and Tectonic Tremor by Applying a Convolutional Neural Network to Running Spectral Images: Seismological Research Letters, **90**(2A), 530−538.

Perol, T., Gharbi, M.andDenolle, M., 2018, Convolutional neural network for earthquake detection and location: Science Advances, 4(2), e1700578.

Ren, T., Lin, M. N., Chen, H. F., et al., 2019, Seismic event classification based on bagging ensemble learning algorithm: Chinese Journal of Geophysics-Chinese Edition, **62**(1), 383−392.

Sugiyama, D., Tsuboi, S.andYukutake, Y., 2021, Application of deep learning-based neural networks using theoretical seismograms as training data for locating earthquakes in the Hakone volcanic region, Japan: Earth Planets and Space, **73**(1).

Szegedy, C., Ioffe, S., Vanhoucke, V., et al., 2016a, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning: AAAI Conference on Artificial Intelligence, **31**.

Szegedy, C., Vanhoucke, V., Ioffe, S., et al., 2016b, Rethinking the Inception Architecture for Computer Vision:2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818−2826.

Szegedy, C., Wei, L., Yangqing, J., et al., 2015, Going deeper with convolutions:2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1−9.

Tang, L. L., Zhang, M.andWen, L. X., 2020, Support Vector Machine Classification of Seismic Events in the Tianshan Orogenic Belt: Journal of Geophysical Research-Solid Earth, **125**(1).

Tian, X., Wang, M. J., Zhang, X., et al., 2022, Discrimination of earthquake and quarry blast based on multi-input convolutional neural network: Chinese Journal of Geophysics-Chinese Edition, **65**(5), 1802−1812.

Trani, L., Pagani, G. A., Zanetti, J. P. P., et al., 2022, DeepQuake - An application of CNN for seismo-acoustic event classification in The Netherlands: Computers & Geosciences, **159**.

Wang, Q. L., Wu, B. G., Zhu, P. F., et al., 2020, ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks:2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 11531−11539.

Wang, T. T., Bian, Y. J., Zhang, Y. X., et al., 2023, Using Artificial Intelligence Methods to Classify Different Seismic Events: Seismological Research Letters, **94**(1), 1−16.

Woo, S., Park, J., Lee, J.-Y., et al., 2018, CBAM: Convolutional Block Attention Module:ECCV, Berlin, Heidelberg, 3−19.

Yoon, C. E., O'Reilly, O., Bergen, K. J., et al., 2015, Earthquake detection through computationally efficient similarity search: Science Advances, **1**(11), e1501057.

Zhao, M., Chen, S.andYuen, D., 2019, Waveform classification and seismic recognition by convolution neural network: Chinese Journal of Geophysics-Chinese Edition, **62**(1), 374−382.

Zhu, W., Mousavi, S. M.andBeroza, G. C., 2019a, Seismic Signal Denoising and Decomposition Using Deep Neural Networks: IEEE Transactions on Geoscience and Remote Sensing, **57**(11), 9476−9488.

Zhu, X., Cheng, D., Zhang, Z., et al., 2019b, An Empirical Study of Spatial Attention Mechanisms in Deep Networks:2019 IEEE/CVF International Conference on Computer Vision (ICCV), 6687−6696.

Huang Yong-Ming, Ph.D., Associate Professor, Assistant Dean of the School of Automation and Head of the Department of Automation. He received his B.S. degree in Automation from Harbin Engineering University in 2005 and his M.S. and Ph.D. degrees from Southeast University in 2008 and 2012, respectively. He is currently working at Southeast University, where he is engaged in the research of seismic electromagnetic disturbance data acquisition and processing, earthquake prediction, earthquake early warning and other directions.