IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 23 April 2024, accepted 8 May 2024, date of publication 14 May 2024, date of current version 21 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3401009

RESEARCH ARTICLE

A Ghost-Attention Network for Discriminating Tectonic and Non-Tectonic Events on a Small and Imbalanced Dataset

XINLIANG LIU[1], TAO REN[1], HONGFENG CHEN[2], (Member, IEEE), PENGYU WANG[1], AND FANCHUN MENG[1]

[1]Institute of Network Science and Big Data Technology, Software College, Northeastern University, Shenyang 110169, China
[2]China Earthquake Networks Center, Beijing 100029, China

Corresponding author: Tao Ren (chinarentao@163.com)

This work supported in part by the National Natural Science Foundation of China under Grant 62276058, Grant 61902057, and Grant 41774063; and in part by the Fundamental Research Funds for the Central Universities under Grant N2217003.

ABSTRACT Discrimination between tectonic and non-tectonic events is crucial to assess seismic hazards and manage associated risks. However, the discrimination process is challenging due to the imbalanced distribution of tectonic and non-tectonic events. In this paper, we propose a ghost-attention network (GA-Net) consisting of multiple ghost modules and convolutional block attention modules (CBAMs) to solve this problem. Ghost module allows the network to extract feature maps using cost effective operations, which are suitable for small and unbalanced training sets. CBAM emphasizes meaningful features along channel and spatial axes, effectively learning which information to emphasize or suppress. We train the proposed GA-Net using seismic data from Shaanxi Province in China. The evaluation shows that on the test set, the proposed GA-Net achieves sensitivity, specificity, and accuracy of 96.40%, 93.24%, and 95.75%, respectively. In addition, a comparative analysis with several state-of-the-art networks establishes its superiority. Our GA-Net exhibits robust generalized performance on several training sets with different imbalanced levels and an independent dataset.

INDEX TERMS Seismic event discrimination, ghost module, convolutional block attention module, imbalanced dataset.

I. INTRODUCTION

Monitoring seismic activities and seismic analysis are two main objectives in assessing seismic hazards and managing associated risks [1], [2]. Modern seismic networks record a large number of seismic events, including several non-tectonic events, such as explosions and collapses, which contaminate the seismic catalogue. These non-tectonic events can cause variations in seismicity patterns, potentially leading to misinterpretations as precursors to past or future major shocks [3]. Therefore, it is imperative to identify and eliminate such non-tectonic events from the seismic catalogue to facilitate the extraction of meaningful information about true Earth processes.

The associate editor coordinating the review of this manuscript and approving it for publication was Leimin Wang.

Over the past decade, the amount of data available to geoscientists has grown dramatically, so identifying all non-tectonic records through manual inspection is no longer a viable option. Big data in geosciences also introduces the challenge of extracting as much useful information as possible and gaining new insights from this data, simulations, and the interplay between the two [4], [5]. Machine learning techniques have proven to be effective tools for analyzing data in various fields of seismology, including event detection [6], phase picking [7], [8], [9], event discrimination [1], [2], [10], [11], [12], [13], [14], [15], geoscience variables estimation [16], [17], [18], [19], and earthquake forecasting [20], [21].

Concerning event discrimination, many different methods have been proposed to distinguish non-tectonic events from tectonic events based on machine learning techniques. They

can be classified into parameter-based and waveform-based methods. Methods based on parameters identify non-tectonic events using calculated related parameters, such as focal depth, first motion polarity, shear wave generation efficiency, P/S amplitude ratio, difference between local magnitude (M_L) and coda duration magnitude (M_c), complexity, and wavelet packet energy ratio [22], [23], [24], [25], [4], [26], [27], [28], [29], [30], [31], [32], [33]. The above discrimination techniques, typically based on manually selecting features, may lead to unreliable results in some cases.

Waveform-based methods are proposed to classify seismic events using automatic feature extraction facilitated by deep learning techniques. Convolutional neural network (CNN) is one of the powerful architectures used to extract well-represented features from input waveforms and is widely used to discriminate between tectonic and non-tectonic events [10], [11], [12], [13], [13], [34], [35]. However, CNN suffers from several drawbacks: 1) CNN usually requires a large number of parameters and floating point operations to achieve satisfactory accuracy, as redundancy in feature maps often ensures a complete understanding of the input data. Accordingly, CNN should be trained using a large amount of training data to avoid overfitting. 2) CNN does not pay attention to the spatial relationship between feature maps.

Ghost module [36] and convolutional block attention module (CBAM) [37] can overcome the shortcomings of CNN. Ghost module first generates some intrinsic feature maps and then uses linear operations to augment features and increase channels. As a result, the number of parameters and computational resources is reduced, easing the requirement for large amounts of training data. CBAM dynamically fine-tunes attention across channels and spatial dimensions, enhancing its focus on critical information within the input data.

In this paper, we use the full seismic waveform as input and propose a novel ghost-attention network (GA-Net) to discriminate between tectonic and non-tectonic events. GA-Net is built with multiple ghost modules and CBAMs. The ghost module is built to efficiently extract features with fewer parameters and calculations, which is suitable for small and unbalanced training sets. In addition, we incorporate CBAM to simultaneously capture the significance of various channels and spatial dimensions within the input data, thus enhancing the focus on the target information. The proposed GA-Net makes several contributions.

1) We propose a novel GA-Net to discriminate between tectonic and non-tectonic events, which are robust in handling imbalanced datasets and achieve promising classification accuracy.

2) The ghost module's integration efficiently preserves multiple feature maps, maintaining a balance between efficiency and computational cost, which achieves satisfactory accuracy on a small dataset.

3) Due to the lack of inter-channel and spatial information in the ghost module, CBAM is introduced to integrate the significance of different channels and spatial dimensions within the input data, thus improving classification performance. In addition, incorporating CBAM before the flattening operation allows the model to learn features that are critical for discrimination.

The rest of this paper is organized as follows. In Section II, the general framework of our GA-Net is demonstrated. Results and discussions are presented in Section III and Section IV, respectively. The conclusions are outlined in Section V.

II. METHODOLOGY

Given that the input is time series seismic waveforms, we use one-dimensional convolution operations to capture temporal features within seismic signals. This allows the model to discern patterns, trends, and correlations inherent in seismic waveforms, thus enhancing the discrimination performance between tectonic and non-tectonic events. However, traditional CNNs usually require a large number of parameters to achieve satisfactory accuracy, which leads to the need for extensive training data to prevent overfitting. Considering that the output feature maps of convolutional layers often contain a lot of redundancy, and some of them could be similar to each other [36], we integrate ghost modules and CBAMs to efficiently produce multiple feature maps with low computational cost and emphasize meaningful features along channel and spatial axes.

Fig. 1 (a) demonstrates the proposed GA-Net architecture for discriminating between tectonic and non-tectonic events. First, we use a standard convolutional layer with 16 filters to extract feature maps from the input waveform. Then, a series of constructed ghost bottlenecks (including two ghost modules) and CBAMs, with progressively increased channels, are applied to extract features efficiently and effectively. Following this, 1×1 convolutional layer and global average pooling layer (GAP) are used to integrate multi-channel feature maps, which contributes to both information integration and model parameter reduction. Before the flattening operation, we apply a CBAM to highlight features considered critical for the final classification task. Finally, a fully connected layer (FC) is used for final classification.

Specifically, the ghost bottleneck is made up of two stacked ghost modules. The first ghost module functions as an expansion layer, increasing the number of channels. The second ghost module reduces the number of channels to match the shortcut path that mitigates the degradation problem. The ghost module produces multiple feature maps at a low computational cost. As shown in Fig. 1 (b), instead of generating N redundant feature maps one by one with a large number of parameters, it is more efficient to obtain $N = C_2 + C_3$ feature maps as output from a ghost module. First, C_2 intrinsic feature maps are obtained by a primary convolution:

$$Y'_{H_2,W_2,C_2} = X_{H_1,W_1,C_1} * f'_{1,C_2}, \tag{1}$$

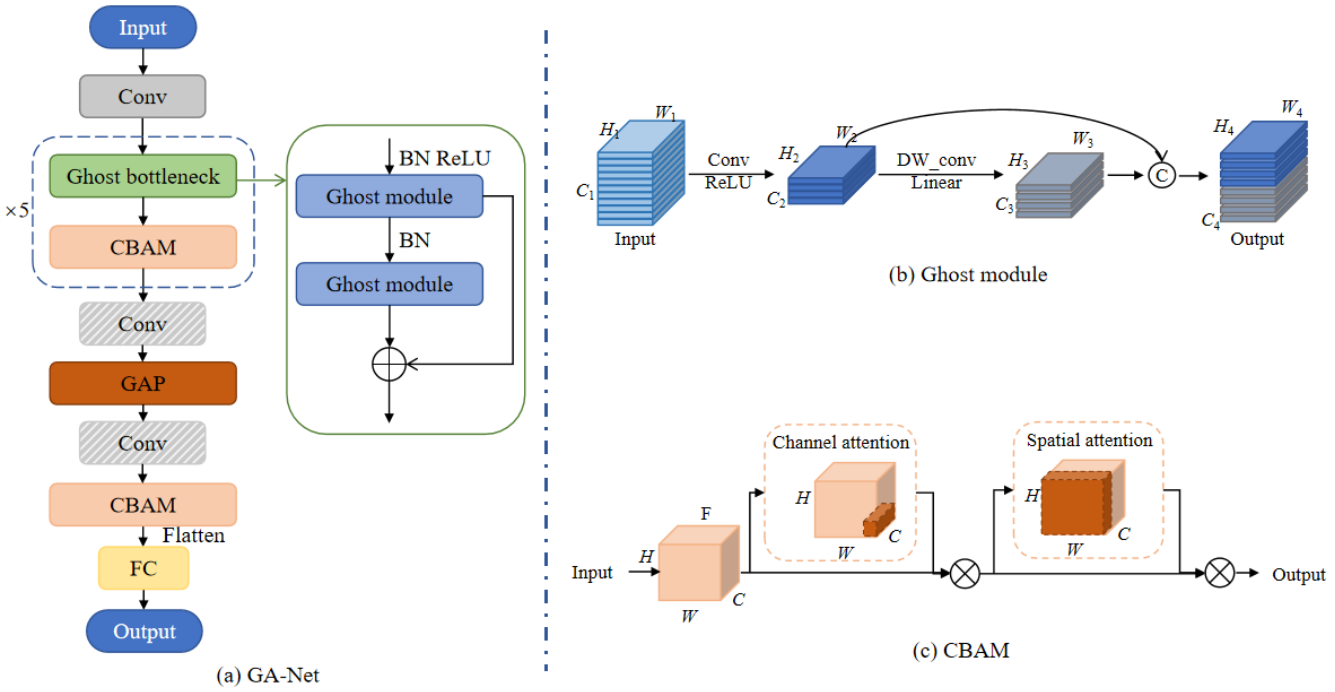

FIGURE 1. a) A schematic architecture of the GA-Net model used for discriminating between tectonic and non-tectonic events. The model input is a three-component waveform, and the output is the probability of the input waveform belonging to either the tectonic or non-tectonic event. (b) Details of the ghost module. (c) Details of the CBAM.

where f' is the utilized filter. The kernel size of f' is 1×1 with C_2 feature maps. Second, C_3 ghost feature maps are generated by a series of cheap linear operations on each intrinsic feature in Y', which is following:

$$y''_{i,j} = \Phi_{i,j}(y'_i), \forall i = 1, \ldots, m, j - 1, \ldots, n, \quad (2)$$

where y'_i is the i-th intrinsic feature map in Y', $\Phi_{i,j}$ is the j-th linear operation for generating the j-th ghost feature map $y''_{i,j}$. In principle, y'_i can have one or more ghost feature maps. Typically, linear transformations are implemented by FC. However, the connections between each input neuron and each output neuron result in quadratic parameter growth, thus challenging training with small datasets. In order to efficiently generate these ghost feature maps, we use depthwise convolution to independently process individual maps of intrinsic feature maps and facilitate mapping from C_2 intrinsic features to C_3 ghost features. The depthwise convolution operation is as follows:

$$Y''_{H_3,W_3,C_3} = Y'_{H_2,W_2,C_2} * f''_{3,C_3}, \quad (3)$$

where f'' is the depthwise convolutional filter. The kernel size is 1×3 with C_3 feature maps. Finally, to further obtain the desired N feature maps, we concatenate the C_2 intrinsic feature maps with the C_3 ghost feature maps.

Feature maps extracted by the ghost module undergo an initial convolution operation of 1×1 to generate intrinsic feature maps, resulting in a lack of spatial information. Subsequently, the intrinsic feature maps are subjected to a depthwise convolutional layer to produce ghost feature maps, which exhibit a lack of inter-channel information. Therefore, in order to address these limitations, we integrate CBAM after the ghost bottleneck to further extract both channel and spatial information. As shown in Fig. 1 (c), given an intermediate feature map as input, CBAM sequentially applies channel and spatial attention modules, then the attention maps are multiplied to the input feature map for adaptive feature refinement. Channel attention and spatial attention are calculated as follows:

$$F' = A_c(F) = \sigma(W_1(\omega(W_0(F_{avg}))) + W_1(\omega(W_0(F_{max}))), \quad (4)$$

$$F'' = A_s(F') = \sigma_s(f^{7\times7}([F'_{avg}; F'_{max}])), \quad (5)$$

where F, $A_c(F)$ and $A_s(F')$ are the input and two attention maps, respectively. σ and ω denote the sigmoid function and the ReLU function, respectively. W_0 and W_1 are the weights of shared multi-layer perceptron (MLP) with one hidden layer. F_{avg} and F_{max} are average pooling and max pooling operation, respectively. Through meticulous model design, our proposed GA-Net effectively extracts waveform features, yielding a satisfactory accuracy in discriminating between tectonic and non-tectonic events.

III. RESULTS

Our proposed GA-Net is based on Keras and utilizes Adam optimizer with a learning rate of 10^{-5} to optimize all

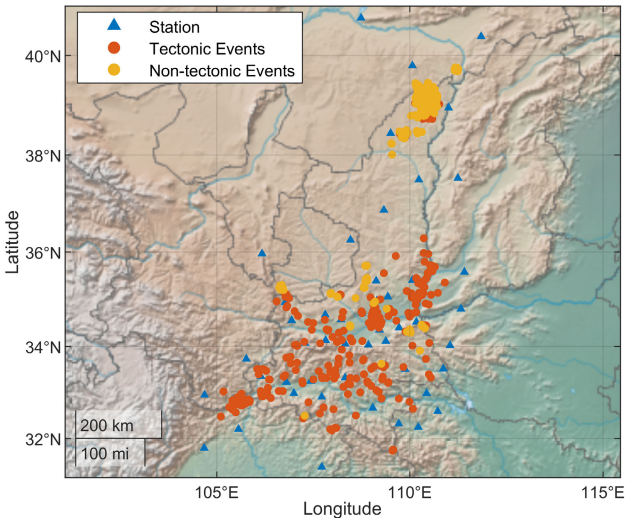

FIGURE 2. Distribution of the seismic stations, tectonic events, and non-tectonic events in the dataset.

parameters of the model. The epoch number is set as 15 to ensure the convergence. The platform is configured with a GeForce GTX 1060 for training and testing.

A. DATASET

Shaanxi Province is located at the intersection of the Qinghai-Tibet earthquake zone, the North China earthquake zone, and the South China earthquake zone. Thus, Shaanxi experiences active tectonic seismic activities. In addition, Shaanxi Province, rich in mineral resources, is prone to non-tectonic events, such as explosions and collapses. Seismic waveforms obtained from 49 stations are recorded by the Shaanxi Regional Seismic Network Center affiliated with the China Earthquake Network Center (CENC). Our dataset includes seismic events recorded from 2009 to 2018, with magnitudes ranging from M_L 1.6 to 4.0. Among them, there are 3359 tectonic records (646 tectonic events) and 1533 non-tectonic records (291 non-tectonic events). Each record has 3 channels with a sampling rate of 100 Hz, and the three-channel waveform is used as input for the proposed method. The distribution of seismic stations and seismic events is shown in Fig. 2.

To ensure the quality of waveforms, we implement the following procedures. First, we focus on waveforms within an epicentral distance of 200 km. Second, we specifically select a 100-second time window after the onset of the P-wave to maintain the completeness [10]. Third, a tenth degree polynomial and standardization are applied to waveforms. Finally, to simulate the model's actual generalization capability, the dataset is divided chronologically. Events from 2009 to 2016 are used for training, and events between 2017 and 2018 are used for testing. Due to the lower number of non-tectonic events, the imbalance in the collected dataset poses challenges in discriminating between these two types of events.

B. MEASUREMENTS

In general, accuracy is commonly used to describe the performance of classification models. However, relying solely on accuracy may mask the performance of the model on minority classes in imbalanced datasets. Sensitivity measures the model's ability to correctly identify positive class samples, while specificity measures its ability to correctly identify negative class samples. By simultaneously considering these two metrics, we can more accurately assess the model's performance across different classes. Therefore, sensitivity, specificity, and accuracy are utilized to quantify the performance in this study, which are defined as follows:

$$sensitivity = \frac{TP}{TP + FN}, \tag{6}$$

$$specificity = \frac{TN}{TN + FP}, \tag{7}$$

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \tag{8}$$

where TP represents the number of true positive samples (a positive sample of which the result of the model is also positive), FP represents the number of false positive samples (a negative sample of which the result of the model is positive), FN represents the number of false negative cases (a positive sample of which the result of the model is negative), TN represents the true negative samples (a negative sample of which the result of the model is also negative). In this paper, tectonic records are designated as positive samples, whereas non-tectonic records are designated as negative samples.

C. TESTING RESULTS

We evaluate the performance of the proposed GA-Net using the test set consisting of 1084 tectonic records (63 events) and 281 non-tectonic records (12 events). Figure 3 illustrates the performance of GA-Net on the test set. As shown in Fig. 3 (a), the proposed method correctly classifies 1307 waveforms (1045 tectonic records and 262 non-tectonic records). Sensitivity, specificity, and overall accuracy are 96.40%, 93.24%, and 95.75%, respectively. The receive operating characteristic curve (ROC) is shown in Fig. 3 (b). The area under the curve (AUC) is 0.99, indicating that the model has excellent discriminatory ability in classifying between tectonic and non-tectonic records.

Furthermore, we employ statistical analysis using the class probabilities outputted by the GA-Net. Fig. 4 illustrates that 83.5% of the tectonic records and 82.13% of the non-tectonic records are concentrated at the two extremes of the probability range (between 0-0.2 and 0.8-1.0). The probability values generated by the classification model tend to be closer to the extremes (0 or 1), suggesting a high level of confidence in its classifications.

We compare the proposed method to several state-of-the-art networks, including CapsNet [1], CNN [10], VGG [11], ResNet [11], GoogleNet [11], CNN [12], and CNN [13] with the same training and testing set. Table 1 shows the results of the GA-Net comparison with the benchmark networks.

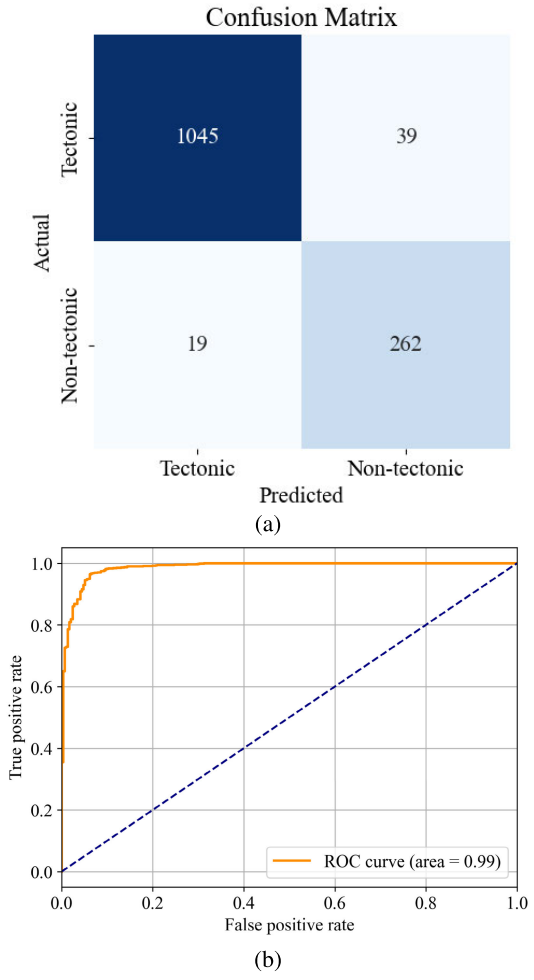

FIGURE 3. Results of the test set. (a) Confusion matrix; (b) ROC curve.

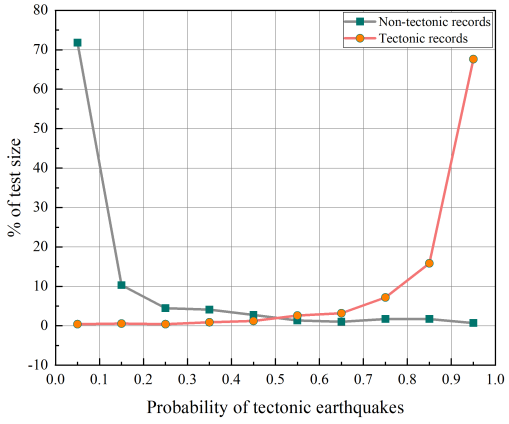

FIGURE 4. Confidence statistics of the test results.

ResNet and GoogleNet demonstrate accuracy in classifying tectonic records, but show inaccuracy in classifying non-tectonic records. This occurs due to the imbalance in the dataset, where tectonic records dominate the majority of the training set. Whereas the proposed GA-Net exhibits a slightly lower accuracy in identifying tectonic records compared to ResNet and GoogleNet, it achieves the highest accuracy in identifying non-tectonic records. That is, through careful model design, GA-Net effectively captures minority class characteristics when confronted with an imbalanced dataset. This advantage not only enhances the reliability of our model in practical applications, but also addresses some limitations in handling imbalanced datasets.

TABLE 1. Comparisons with state-of-the-art networks.

Method	Sensitivity	Specificity	Accuracy	Parameters
CapsNet [1]	90.31%	80.43%	88.28%	**40.32K**
CNN [10]	93.08%	90.04%	92.45%	1.36M
VGG [11]	93.08%	66.19%	87.55%	11.37K
ResNet [11]	98.61%	90.34%	96.85%	7.13M
GoogleNet [11]	**99.08%**	91.10%	**97.44%**	5.12M
CNN [12]	94.74%	89.68%	93.70%	5.20M
CNN [13]	91.42%	84.34%	89.96%	80.06K
GA-Net (ours)	96.40%	**93.24 %**	95.75%	1.57M

TABLE 2. The ablation study results.

Method	Sensitivity	Specificity	Accuracy
GA-Net (w/o GM)	93.91%	85.77%	92.23%
GA-Net (w/o CBAM)	95.02%	91.46%	94.29%
GA-Net (ours)	**96.40%**	**93.24 %**	**95.75%**

The number of parameters directly correlates with both memory usage and computational cost during both training and inference phases. As shown in Table 1, GA-Net utilizes significantly fewer parameters (1.57M) compared to ResNet (7.13M) and GoogleNet (5.12M), yet achieves comparable or superior performance metrics. This demonstrates that GA-Net not only requires less computational resources but maintains a high level of accuracy and stability in its predictions. The reduction in parameters translates to a more efficient model that is easier to train and deploy, particularly in environments where computational resources are limited.

D. ABLATION STUDY

To demonstrate the effectiveness of the ghost module and CBAM, the ablation study is conducted. As shown in Table 2, without the ghost module (w/o GM) and the CBAM (w/o CBAM), the model tends to learn features associated with the dominant class (tectonic records) while ignoring those relevant to the minority class (non-tectonic records). Both GA-Net (w/o GM) and GA-Net (w/o CBAM) exhibit overfitting to the dominant class, resulting in poor generalization performance in practical applications. However, the proposed GA-Net, incorporating the ghost module and CBAM, achieves the best results on the unbalanced dataset.

IV. DISCUSSION

To select the optimal network architecture, we tune the number of ghost bottlenecks stacked. As shown in Fig. 5, both the excessive and insufficient stacking of ghost bottlenecks can lead to suboptimal model performance. In the former case, constrained by limited training data, the model becomes overly complex, making it prone to overfitting noise and

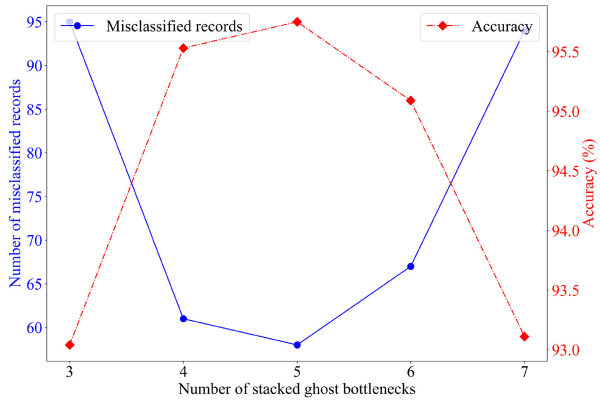

FIGURE 5. Results of different number of stacked ghost bottlenecks.

TABLE 3. Results of GA-Net with different CBAM locations.

	Misclassified records	Accuracy
After input (Input)	70	94.87%
After each ghost bottleneck (G-bneck)	66	95.16%
Before the flatten operation (Flatten)	71	94.80%
Input+G-bneck	67	95.09%
Input+Flatten	70	94.87%
G-bneck+Flatten (ours)	**58**	**95.75%**
Input+G-bneck+Flatten	78	94.29%

details within the data, thereby diminishing generalized performance. Conversely, in the latter scenario, insufficient model complexity hinders the capture of essential patterns and intricacies in the data, resulting in underfitting. Therefore, we use five stacked ghost bottlenecks to reduce the complexity of the model and achieve high classification accuracy.

A. THE LOCATION OF CBAM

We use CBAM to highlight meaningful features along channel and spatial axes and efficiently learn which information to highlight or suppress. To investigate the effect of placing CBAM in different locations. We perform experiments as shown in Table 3. The results show that placing the CBAM after each ghost bottleneck and before the flattening operation achieves the highest accuracy. Positioning the CBAM after each ghost bottleneck enhances attention to various channels and spatial features. This allows the model to dynamically learn and adapt to significant features within the input data, improving overall performance. In addition, applying CBAM before flattening allows for a more effective focus on features considered critical to the final classification task.

B. MISCLASSIFIED RECORDS ANALYSIS

To analyze misclassified records, we visualize some records in Fig. 6. From a waveform perspective, GA-Net has a better classification effect for records with relatively large amplitude P and S waves, which aligns with the established P/S amplitude ratio. In addition, GA-Net is robust in discriminating records with channel dropout and baseline deviation, as shown in Fig. 6 (b) and Fig. 6 (d). Fig. 6 (e) and Fig. 6 (f) represent misclassified tectonic and non-tectonic records, respectively. These records exhibit large epicentral distances, resulting in a low signal to noise ratio and poor waveform quality.

TABLE 4. GA-Net results across various imbalance handling techniques.

	Sensitivity	Specificity	Accuracy
Time shifting	94.37%	80.07%	91.43%
Noise addition	95.48%	91.46%	94.65%
Focal loss	94.83%	86.12%	93.04%
GA-Net(ours)	**96.40%**	**93.24%**	**95.75%**

In addition, we analyze the impact of different types of events on the GA-Net. We use the event discrimination criteria proposed in our previous work [10]. There are 225 tectonic and 50 non-tectonic events (8 explosions and 42 collapses) in the test set. We only misclassified three tectonic events and one explosion. All collapse events are properly discriminated against. This ability can be attributed to larger periods of collapse events compared to tectonic events, which is consistent with the perception of manual discrimination. The vast majority of misclassified events are located at the edge of the seismic observation network. Only a small number of stations with large epicentral distances can record these events, so these events are easily confused with shallow explosions.

C. EXPLORING IMBALANCE HANDLING STRATEGIES

For a more effective resolution of dataset imbalance issues, we conduct a detailed analysis with various techniques, including time shifting, noise addition, and focal loss [38]. As shown in Table 4, the limited diversity in the data generated by time shifting and noise addition results in the model still overly focusing on tectonic records. The introduction of focal loss aims to address the issue of foreground-background imbalance in segmentation tasks, characterized by significant differences between the foreground and background. However, in our study, distinguishing between tectonic and non-tectonic records, which exhibit high similarity, renders focal loss inadequate for resolving this particular challenge. In the future, we will explore more appropriate methods to address the imbalance issue at the data level.

D. MODEL PERFORMANCE ACROSS REGIONS

To assess the model's applicability across various geophysical contexts, we obtain data spanning the year 2019 from the CENC, focusing on the regions of Shanxi, Shandong, and Henan. These regions are selected for their abundant data availability and authorized access, ensuring a robust dataset for our analysis and reinforcing the credibility of our findings. As shown in Table 5, the testing results of GA-Net in Shanxi, Shandong, and Henan are excellent, with all metrics exceeding 90%, indicating its robustness.

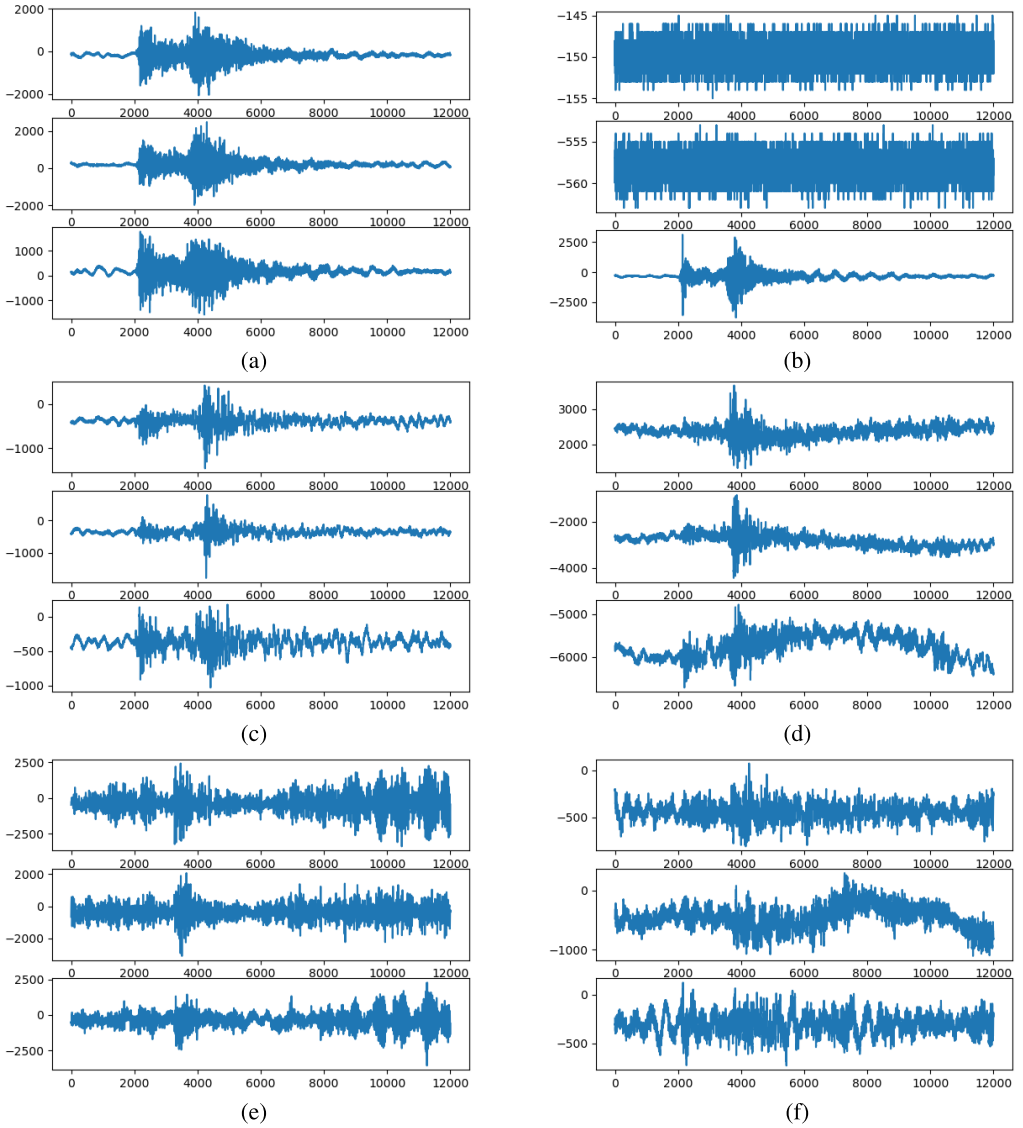

FIGURE 6. Discrimination analysis. (a) and (b) are correctly classified tectonic records; (c) and (d) are correctly classified non-tectonic records; (e) and (f) are misclassified tectonic and non-tectonic records, respectively.

TABLE 5. GA-Net testing results across different regions.

Region	Sensitivity	Specificity	Accuracy
Shanxi	93.59% (146/156)	95.24% (80/84)	94.16%
Shandong	91.28% (199/218)	97.01% (65/67)	92.63%
Henan	90.20% (46/51)	92.86% (26/28)	91.14%

E. IMBALANCE ANALYSIS EXPERIMENT

Due to the imbalanced distribution of tectonic and non-tectonic events, discrimination between these two categories may tend to overfit the majority class and neglect the minority class. To address this problem, we propose a GA-Net to reduce the number of parameters to prevent overfitting and increase focus on features that are critical for classification. To better validate the robustness of our GA-Net, we construct four training sets with varying levels of imbalance, including 1:1 (a balanced training set), 2:1, 5:1, and 10:1 ratios, employing two distinct strategies.

In strategy 1, we start by creating a balanced training set of randomly selected 200 tectonic and 200 non-tectonic recordings. Then, we construct the remaining three training scenarios by accumulating a certain number of randomly selected tectonic records into the previously created training scenario. For example, to create a training set with an imbalance of 2:1, we add another 200 tectonic records to the originally established balanced training set.

We report the corresponding sensitivity, specificity, and accuracy from four classifiers in Fig. 7 (a). With the accumulation of tectonic records, the imbalance in the training set gradually intensifies, leading to a significant increase in the sensitivity. Considering that the test set consists of 1084 tectonic records and 281 non-tectonic

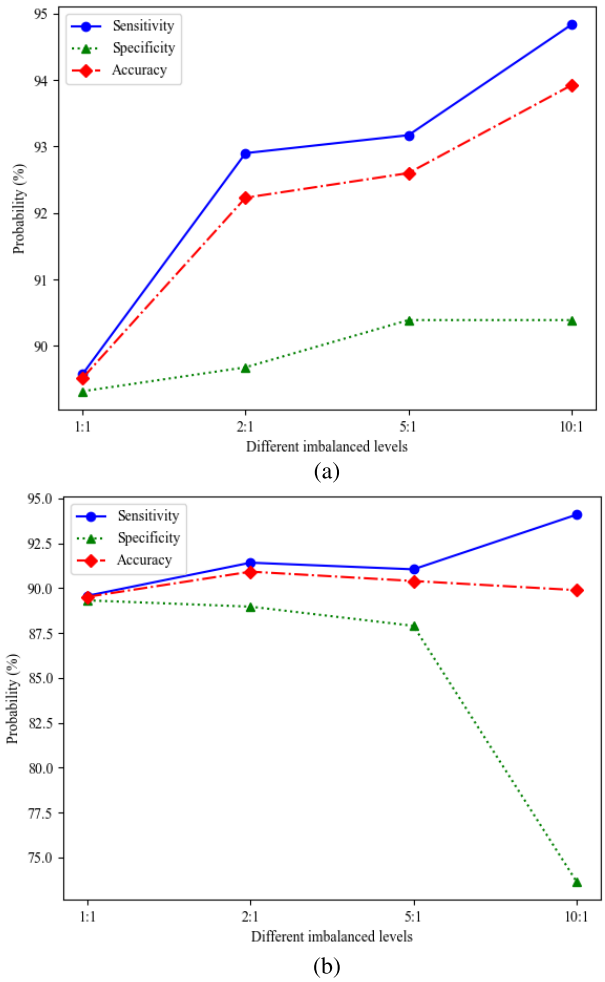

FIGURE 7. Results of two strategies for different levels of imbalance of the dataset. (a) Strategy 1: Increasing dataset; (b) Strategy 2: Constant dataset.

records, the accuracy curve exhibits a similar trend to sensitivity. It is worth noting that, although the number of non-tectonic records is unchanged, the accumulation of tectonic records allows GA-Net to learn the characteristics of tectonic records comprehensively, thus improving the accuracy of discriminating between tectonic and non-tectonic records. Moreover, the accuracy of the balanced training set with only 400 records yields more than 89%, and the specificity of the unbalanced training set of 10:1 still exceeds 90%. This indicates the robustness of our GA-Net with respect to limited and imbalanced training sets.

In strategy 2, we maintain a total dataset size of 400 and partition it according to various imbalance ratios. The corresponding sensitivity, specificity, and accuracy from four classifiers are shown in Fig. 7 (b). When the total dataset size remains constant, an increasing proportion of tectonic records correlates with an upward trend in sensitivity. For imbalance ratios of 2:1 and 5:1, specificity shows only a slight decrease. At an imbalance ratio of 10:1, where there are only 36 non-tectonic records in the training set, specificity still reaches 74.57%. In summary, the experiments conducted on two strategies for constructing imbalanced datasets demonstrate the stability of the proposed GA-Net in addressing small yet imbalanced datasets.

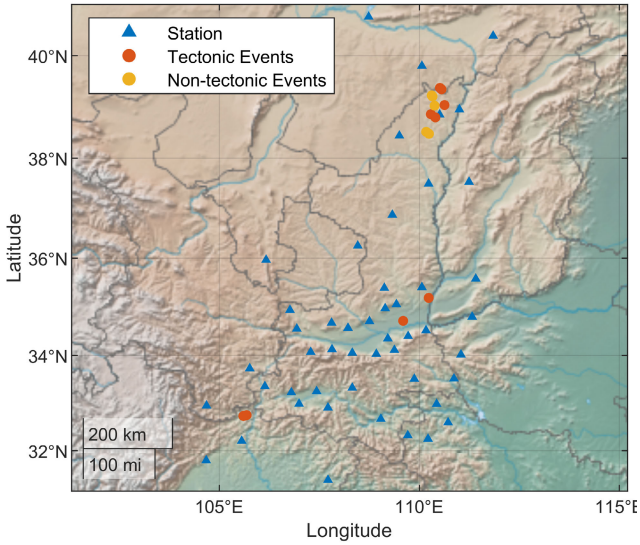

FIGURE 8. Distribution of the seismic stations, tectonic events, and non-tectonic events in the independent test set.

F. APPLICATION IN AN INDEPENDENT TEST SET

To evaluate the generalization capability of the proposed method, we apply the trained model to data recorded between January and February 2019, which is distinct from the dataset used for model training and testing. Accordingly, we extract 53 tectonic records (10 tectonic events) and 10 non-tectonic records (5 non-tectonic events). The distribution of seismic stations and seismic events is shown in Fig. 8. As a result, the proposed method misclassifies only 1 tectonic record and correctly classifies all non-tectonic records with an overall accuracy of 98.41%.

V. CONCLUSION

Event discrimination is important for better seismic analysis. In this paper, we use the full seismic waveform as input and propose a novel GA-Net to discriminate between tectonic and non-tectonic events. GA-Net includes multiple ghost modules and CBAMs. Specifically, the ghost module gives the network the ability to extract feature maps using inexpensive operations, making it particularly suitable for small and imbalanced training sets. CBAM emphasizes meaningful features along channel and spatial axes, and efficiently learns which information to emphasize or suppress.

The evaluation shows that on the test set, the proposed GA-Net achieves sensitivity, specificity, and accuracy of 96.40%, 93.24%, and 95.75%, respectively. In addition, the proposed method is compared with several state-of-the-art networks and achieves promising classification performance. Finally, we find that the proposed model achieves robust performance on an independent dataset, exhibiting high

classification accuracy for unseen data. Currently, the proposed GA-Net has been deployed for practical testing with the CENC. In the future, we will collaborate with the Shaanxi Regional Seismic Network Center to integrate this method into earthquake monitoring systems.

For future work, it is imperative to acknowledge the inherent imbalance in the classification problem of tectonic versus non-tectonic events. While our model has demonstrated efficacy, there remains ample room for exploring more effective feature extraction techniques tailored specifically for such imbalanced datasets. Moreover, future studies might involve a comprehensive analysis of the disparities in geological features across different regions and their impact on seismic event detection. By incorporating geological backgrounds into the model's training process, we can enhance its adaptability and performance across various geographical locations.

REFERENCES

[1] O. M. Saad, M. S. Soliman, Y. Chen, A. A. Amin, and H. E. Abdelhafiez, "Discriminating earthquakes from quarry blasts using capsule neural network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, 2022, Art. no. 8029605.

[2] Q. Kong, R. Wang, W. R. Walter, M. Pyle, K. Koper, and B. Schmandt, "Combining deep learning with physics based features in explosion-earthquake discrimination," *Geophys. Res. Lett.*, vol. 49, no. 13, Jul. 2022, Art. no. e2022GL098645.

[3] S. M. Mousavi, "Comment on 'recent developments of the middle east catalog' by Zare et al." *J. Seismol.*, vol. 21, pp. 269–271, Nov. 2017.

[4] K. J. Bergen, P. A. Johnson, M. V. de Hoop, and G. C. Beroza, "Machine learning for data-driven discovery in solid Earth geoscience," *Science*, vol. 363, no. 6433, Mar. 2019, Art. no. eaau0323.

[5] A. Karpatne, I. Ebert-Uphoff, S. Ravela, H. A. Babaie, and V. Kumar, "Machine learning for the geosciences: Challenges and opportunities," *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 8, pp. 1544–1554, Aug. 2019.

[6] O. M. Saad, Y. Chen, D. Siervo, F. Zhang, A. Savvaidis, G.-C.-D. Huang, N. Igonin, S. Fomel, and Y. Chen, "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.

[7] 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.

[8] E. Pardo, C. Garfias, and N. Malpica, "Seismic phase picking using convolutional networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7086–7092, Sep. 2019.

[9] W. Zhu, I. W. McBrearty, S. M. Mousavi, W. L. Ellsworth, and G. C. Beroza, "Earthquake phase association using a Bayesian Gaussian mixture model," *J. Geophys. Res., Solid Earth*, vol. 127, no. 5, May 2022, Art. no. e2021JB023249.

[10] X. Liu, T. Ren, H. Chen, and Y. Chen, "Classification of tectonic and non-tectonic seismicity based on convolutional neural network," *Geophys. J. Int.*, vol. 224, no. 1, pp. 191–198, Jul. 2020.

[11] L. Jia, H. Chen, and K. Xing, "Rapid classification of local seismic events using machine learning," *J. Seismol.*, vol. 26, no. 5, pp. 897–912, Oct. 2022.

[12] L. Barama, J. Williams, A. V. Newman, and Z. Peng, "Global nuclear explosion discrimination using a convolutional neural network," *Geophys. Res. Lett.*, vol. 50, no. 17, Sep. 2023, Art. no. e2021JB023249.

[13] J. Zhu, L. Fang, F. Miao, L. Fan, J. Zhang, and Z. Li, "Deep learning and transfer learning of earthquake and quarry-blast discrimination: Applications to Southern California and Eastern Kentucky," *Geophys. J. Int.*, vol. 236, no. 2, pp. 979–993, Dec. 2023.

[14] F. Miao, N. S. Carpenter, Z. Wang, A. S. Holcomb, and E. W. Woolery, "High-accuracy discrimination of blasts and earthquakes using neural networks with multiwindow spectral data," *Seismol. Res. Lett.*, vol. 91, no. 3, pp. 1646–1659, Mar. 2020.

[15] M. Musil and A. Plešinger, "Discrimination between local microearthquakes and quarry blasts by multi-layer perceptrons and Kohonen maps," *Bull. Seismol. Soc. Amer.*, vol. 86, no. 4, pp. 1077–1090, Aug. 1996.

[16] T. Ren, X. Liu, H. Chen, G. M. Dimirovski, F. Meng, P. Wang, Z. Zhong, and Y. Ma, "Seismic severity estimation using convolutional neural network for earthquake early warning," *Geophys. J. Int.*, vol. 234, no. 2, pp. 1355–1362, Aug. 2023.

[17] T. Okazaki, N. Morikawa, A. Iwaki, H. Fujiwara, T. Iwata, and N. Ueda, "Ground-motion prediction model based on neural networks to extract site properties from observational records," *Bull. Seismol. Soc. Amer.*, vol. 111, no. 4, pp. 1740–1753, May 2021.

[18] A. Chatterjee, N. Igonin, and D. T. Trugman, "A real-time and data-driven ground-motion prediction framework for earthquake early warning," *Bull. Seismol. Soc. Amer.*, vol. 113, no. 2, pp. 676–689, Dec. 2022.

[19] X. Zhang, W. Reichard-Flynn, M. Zhang, M. Hirn, and Y. Lin, "Spatiotemporal graph convolutional networks for earthquake source characterization," *J. Geophys. Res., Solid Earth*, vol. 127, no. 11, Nov. 2022, Art. no. e2022JB024401.

[20] K. Dascher-Cousineau, O. Shchur, E. E. Brodsky, and S. Günnemann, "Using deep learning for flexible and scalable earthquake forecasting," *Geophys. Res. Lett.*, vol. 50, no. 17, Aug. 2023, Art. no. e2023GL103909.

[21] J. Liu, T. Zhang, C. Gao, and P. Wang, "Forecasting earthquake magnitude and epicenter by incorporating spatiotemporal priors into deep neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5911413.

[22] D. Bowers and N. D. Selby, "Forensic seismology and the comprehensive nuclear-test-ban treaty," *Annu. Rev. Earth Planet. Sci.*, vol. 37, no. 1, pp. 209–236, May 2009.

[23] C. T. O'Rourke, G. E. Baker, and A. F. Sheehan, "Using *P*/*S* amplitude ratios for seismic discrimination at local distances," *B. Seismol. Soc. Amer.*, vol. 106, no. 5, pp. 2320–2331, Oct. 2016.

[24] M. L. Pyle and W. R. Walter, "Investigating the effectiveness of *P*/*S* amplitude ratios for local distance event discrimination," *Bull. Seismol. Soc. Amer.*, vol. 109, no. 3, pp. 1071–1081, Mar. 2019.

[25] R. Wang, B. Schmandt, and E. Kiser, "Seismic discrimination of controlled explosions and earthquakes near Mount St. Helens using *P*/*S* ratios," *J. Geophys. Res. Solid Earth*, vol. 125, no. 10, Oct. 2020, Art. no. e2020JB020338.

[26] M. L. Pyle and W. R. Walter, "Exploring the effects of emplacement conditions on explosion *P*/*S* ratios across local to regional distances," *Seismol. Res. Lett.*, vol. 93, no. 2A, pp. 866–879, Mar. 2022.

[27] K. D. Koper, M. M. Holt, J. R. Voyles, R. Burlacu, M. L. Pyle, R. Wang, and B. Schmandt, "Discrimination of small earthquakes and buried single-fired chemical explosions at local distances (<150 km) in the Western United States from comparison of local magnitude (M_L) and coda duration magnitude (M_C)," *Bull. Seismol. Soc. Amer.*, vol. 111, no. 1, pp. 558–570, Feb. 2021.

[28] J. R. Voyles, M. M. Holt, J. M. Hale, K. D. Koper, R. Burlacu, and D. J. A. Chambers, "A new catalog of explosion source parameters in the Utah region with application to ML–MC-based depth discrimination at local distances," *Seismol. Res. Lett.*, vol. 91, no. 1, pp. 222–236, Jan. 2020.

[29] W.-Y. Kim, V. Aharonian, A. L. Lerner-Lam, and P. G. Richards, "Discrimination of earthquakes and explosions in Southern Russia using regional high-frequency three-component data from the IRIS/JSP Caucasus network," *Bull. Seismological Soc. Amer.*, vol. 87, no. 3, pp. 569–588, Jun. 1997.

[30] W. Kim, P. G. Richards, D. Schaff, E. Jo, and Y. Ryoo, "Identification of seismic events on and near the North Korean test site after the underground nuclear test explosion of 3 September 2017," *Seismol. Res. Lett.*, vol. 89, no. 6, pp. 2120–2130, Sep. 2018.

[31] N. Orlic and S. Loncaric, "Earthquake—Explosion discrimination using genetic algorithm-based boosting approach," *Comput. Geosci.*, vol. 36, no. 2, pp. 179–185, Feb. 2010.

[32] M. M. Holt, K. D. Koper, W. Yeck, S. D'Amico, Z. Li, J. M. Hale, and R. Burlacu, "On the portability of ML–Mc as a depth discriminant for small seismic events recorded at local distances," *Bull. Seismol. Soc. Amer.*, vol. 109, no. 5, pp. 1661–1673, Oct. 2019.

[33] T. Wang, Y. Bian, Y. Zhang, and X. Hou, "Classification of earthquakes, explosions and mining-induced earthquakes based on XGBoost algorithm," *Comput. Geosci.*, vol. 170, Jan. 2023, Art. no. 105242.

[34] S. Yuan, J. Liu, S. Wang, T. Wang, and P. Shi, "Seismic waveform classification and first-break picking using convolution neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 272–276, Feb. 2018.
[35] L. Trani, G. A. Pagani, J. P. P. Zanetti, C. Chapeland, and L. Evers, "DeepQuake—An application of CNN for seismo-acoustic event classification in The Netherlands," *Comput. Geosci.*, vol. 159, Feb. 2022, Art. no. 104980.
[36] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More features from cheap operations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1577–1586.
[37] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 3–19.
[38] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

XINLIANG LIU received the B.S. degree in computer science and technology from Northeast Agricultural University, Harbin, China, in 2018. She is currently pursuing the Ph.D. degree with the School of Software Engineering, Northeastern University, Shenyang, China.

Her research interests include machine learning and its application in seismology.

TAO REN received the B.S., M.S., and Ph.D. degrees in engineering from Northeastern University, Shenyang, China, in 2003, 2005, and 2007, respectively.

He was a Postdoctoral Researcher in computer science, from 2009 to 2013. He is currently a Professor with Northeastern University. Recently, he is in charge of 20 projects, such as the National Natural Science Foundation of China. He has published more than 50 high-qualified academic papers in several high-ranking journals or conferences. Furthermore, he has published four books and been authorized more than 20 Chinese patents. His main research interest includes machine learning and its application in seismology.

HONGFENG CHEN (Member, IEEE) received the master's degree in solid geophysics from the Institute of Geophysics, China Earthquake Administration, Beijing, China, in 2006.

His research interests include earthquake monitoring, real-time seismology, artificial intelligence, and big data processing.

PENGYU WANG received the master's degree from Northeastern University, Shenyang, China, in 2020, where he is currently pursuing the Ph.D. degree with the Software College.

His research interest includes artificial intelligence.

FANCHUN MENG is currently pursuing the Ph.D. degree with the Software College, Northeastern University, Shenyang, China.

He is mainly engaged in machine learning in the field of seismology.