Volcanic Earthquake Classification using Transformer Encoder and Its Interpretability Evaluation

Y. Suzuki[1], Y. Yukutake[2], T. Ominato[2], M. Yamasaki[1] and Ahyi Kim[1*]

[1]Yokohama City University

[2]Earthquake Research Institute, University of Tokyo

Corresponding author: Ahyi Kim (ahyik@yokohama-cu.ac.jp)

Key Points:

- The transformer-encoder-based model precisely classifies volcanic earthquakes, matching or exceeding traditional methods.
- After training data screening, attention weights in the model focused on seismic waveform features similar to typical human analysis.
- A clear distinction in the waveform features via new criteria for data labeling is essential to boost the model's interpretability.

Abstract

Precisely classifying earthquake types is crucial for elucidating the relationship between volcanic earthquakes and volcanic activity. However, traditional methods rely on subjective human judgment, which requires considerable time and effort. To improve this, we developed a deep learning model using a transformer encoder for a more objective and efficient classification. Tested on Mount Asama's diverse seismic activity, our model achieved high F1 scores (0.876 for tectonic, 0.964 for low-frequency earthquakes, and 0.995 for noise), equivalent to or better than other methods. According to the attention weight visualization, our model focuses on critical seismic signal features for classification, similar to expert analysis. However, it has been demonstrated that removing subjective elements and employing standardized labeling of the training data based on waveform features are necessary to enhance the interpretability of the model. Additionally, the analyses suggest that stations near the volcanic crater are essential for a highly interpretative and accurate classification.

Plain Language Summary

Volcanoes can cause several small earthquakes, particularly prior to eruptions. Unlike regular earthquakes caused by the movement of the Earth's plates, these are linked to volcanic activity. Knowing the differences between these earthquake types is critical for predicting eruptions, but it is usually a complex task that takes considerable time for experts. To simplify this process, we created a computer program that learns from data to identify earthquakes more easily. We tested it on earthquakes from a volcano in Japan, and it worked very well, even better than the other methods. The program examines earthquake data and identifies key features to classify them, similar to experts. However, it is crucial to ensure that the data used for training the program are labeled appropriately. In addition, for best results, data from monitoring stations very close to the volcano's crater should be used.

1 Introduction

Active earthquake swarms are frequently associated with volcanic activity in volcanic regions. Monitoring these earthquakes is crucial for assessing current volcanic activities and providing insights into the physical processes of volcanic fluids, such as magma and hydrothermal fluids (Chouet et al., 1988; McNutt, 1996; Nishimura & Iguchi, 2011). Minakami et al. (1970) proposed that volcanic earthquakes could be classified into four types based on waveform characteristics, source location, and their relationship with surface activity: A-type (tectonic earthquakes), B-type (low-frequency earthquakes), volcanic tremors, and explosion earthquakes. Although the frequency of these four types of earthquakes varies by volcano and may be further subdivided or include other types, a typical pattern before eruptions begin is an increase in A-type earthquakes, followed by an increase in B-type earthquakes and volcanic tremors (Oikawa et al., 2006; Iguchi, 2013). However, the relationship between these earthquakes and their eruptions remains unclear. The precise classification of earthquake types is crucial for elucidating the relationship between volcanic earthquakes and volcanic activity.

Traditional methods classify volcanic earthquakes based on visual observations of the seismic wave amplitude, dominant frequency, duration, and related surface activity evaluated by experts, which are time-consuming and costly for human resources. In addition, the classification criteria may need to be standardized across observers and institutions.

Several deep learning models have been proposed for automatic volcanic earthquake classification, including support vector machines (Malfante et al., 2018) and convolutional neural networks (CNN) (Titos et al., 2018). CNN studies employed two-dimensional time-frequency representation data as inputs (Canário et al., 2020; Lara et al., 2021; Nakano et al., 2022).

Deep learning has significantly improved the accuracy of earthquake-type classification and reduced human and time costs; however, the opacity of the classification rationale has been highlighted. To mitigate the adverse effects of the black-box nature of deep learning, eXplainable AI (XAI) technologies have been developed to make the reasoning behind decisions and predictions made using deep learning understandable to humans and applicable in Earth sciences (Mohammadi et al., 2023). Therefore, we constructed a model using a transformer encoder and visualized attention weights to verify whether the process it follows for classifying volcanic earthquakes is similar to that of experts or if it possesses unique criteria for classification.

2 Methods

A classification model was constructed using only the encoder component of the transformer architecture (Vaswani et al., 2017). A vital feature of the transformer is its attention mechanism, which enables the calculation of relevance among all tokens (the smallest unit of data) within the input sequence. Furthermore, visualizing attention weights (AW) potentially reveal the decision-making rationale and serve a function in XAI. The encoder employs a self-attention mechanism that calculates the representation of each input position by considering its relationship with every other position in the sequence. To illustrate, given an input sequence $X = [x_1, x_2, \dots x_n]$, for any element x_i in this sequence, the associated weight matrices create vectors known as queries (Q), keys (K), and values (V). The model computes the similarity scores between the position of x_i and every other position x_j in the sequence, using the dot product of Q and K. These scores are typically normalized by the scaling factor $\sqrt{d_k}$, where d_k represents the dimensionality of K, ensuring that the values do not excessively increase. A softmax function is applied to these normalized scores to calculate the AW, highlighting the importance of the attention that each token should receive relative to the others. Tokens with similar Q and K pairs receive higher AW. These AW are then applied to V to produce the final output vector as follows:

$$Attention\,(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (1)$$

to construct a classification model for volcanic earthquakes (Figure 1). The model embeds inputs of three-component seismic waveforms through convolution and positional embedding and then processes them through three layers of the transformer encoder, outputting probabilities for A- and B-type events and noise through a linear transformation of the classification token. Specifically, waveform data with a sequence length of 3000 as an array with three channels and a one-dimensional convolution layer that converted a 50-frame with a 3-channel array into a 150-channel array were applied with a stride of 10.

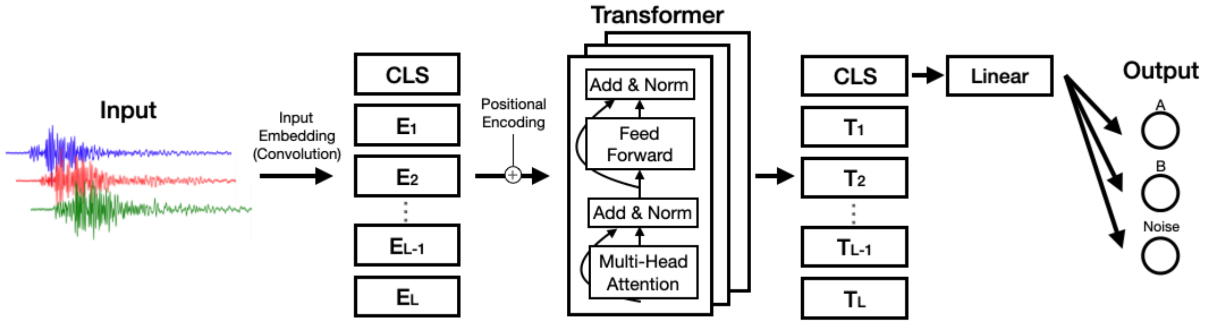

Figure 1. The network architecture. CLS represents the classification token, E represents embedding, T indicates the sequence encoded by the transformer encoder, the subscripts denote sequence numbers (1, 2, ⋯, L), and Linear indicates a linear layer. The model embeds inputs of three-component seismic waveforms through convolution and positional embedding, then processes them through three layers of transformer encoders. A linear transformation of the classification token outputs probabilities for A- and B-type earthquakes and noise. Within the transformer encoder, each sublayer (Self-Attention layer and Feedforward Network) is followed by a normalization layer, with residual connections added between the input and output.

Transformers are adept at capturing long-term features but require assistance with short-term features (Gulati et al., 2020). Convolutional layers were used to embed the inputs to compensate for this weakness. Convolutional layers apply a filter to the input data to create feature maps, enabling dimension reduction and extraction of local features.

Relative positional embedding was employed because absolute positional information was not necessary to classify the waveform data (Huang et al., 2018). Relative positional embedding utilizes an algorithm called Skew to create the matrix S^{rel} and adjust QK^T. Specifically, a dummy column vector of length N is padded after the rightmost column of the sequence, the matrix is flattened, and then a dummy row of length N-1 is padded before reshaping the matrix to size $(N+1, 2N-1)$. This matrix was sliced to retain only the first N rows and the last N columns, resulting in an $N \times N$ matrix. With relative positional embedding, Equation (1) was as follows:

$$Attention\ (Q, K, V) = softmax\left(\frac{QK^T + S^{rel}}{\sqrt{d_k}}\right)V \tag{2}$$

When performing classification with a transformer, adding a single data point for classification (classification token) at the beginning of the input embedding is common. The classification token was then extracted from the results encoded by the transformer encoder, and a linear transformation was performed to calculate the output probabilities for each type. Various methods exist for computing the classification tokens, such as random initialization or calculations from the mean of the outputs. We adopted an averaging approach across temporal dimensions.

Training was conducted with 100 iterations using a mini-batch size of 200, and a multiclass cross-entropy error was employed as the loss function. Considering several parameters in the transformer, which pose a risk of overfitting, the AdamW optimizer, Adam with weight decay, was used for optimization. The learning rate was set to 0.001, and the weight decay to 0.01.

3 Data

To validate the model mentioned above, earthquakes occurring at Mount Asama, where various low-frequency earthquakes are frequently observed, were selected for the analysis (Figure 2a). Mount Asama, an active andesitic volcano in central Japan, has witnessed most of its historical eruptions at its summit crater, with a diameter of approximately 400 m. Details of the historical volcanic activity were comprehensively described by Oikawa et al. (2006) and Takeo et al. (2022). Seismic observations have been conducted on Mount Asama for more than 100 years. Since the establishment of a modern observation network around the volcano in 2003, a network consisting of 30 seismometers, including 19 with broadband sensors, has been installed by institutions such as the Earthquake Research Institute (ERI), Japan Meteorological Agency (JMA), and National Research Institute for Earth Science and Disaster Resilience (NIED) (Figure 2a) up to 2017.

Volcanic earthquakes on Mount Asama have been classified based on their source locations and waveform characteristics (Minakami et al., 1970). Recent observations and research have identified several types of volcanic earthquakes on Mount Asama: B-type earthquakes located directly beneath the crater; A-type earthquakes situated slightly westward and more profoundly than the B-type distribution area; F-type earthquakes occurring on the flanks of the volcano away from the crater; and N-type earthquakes occurring within the B-type source area but characterized by decaying oscillatory waveforms (Oikawa et al., 2006). These are currently subdivided into 12 types. Among these, B-type earthquakes are the most frequent and tend to cluster, often preceding eruptions (Oikawa et al., 2006). Therefore, the precise identification of B-type earthquakes is crucial for understanding the seismic and volcanic activity of Mount Asama.

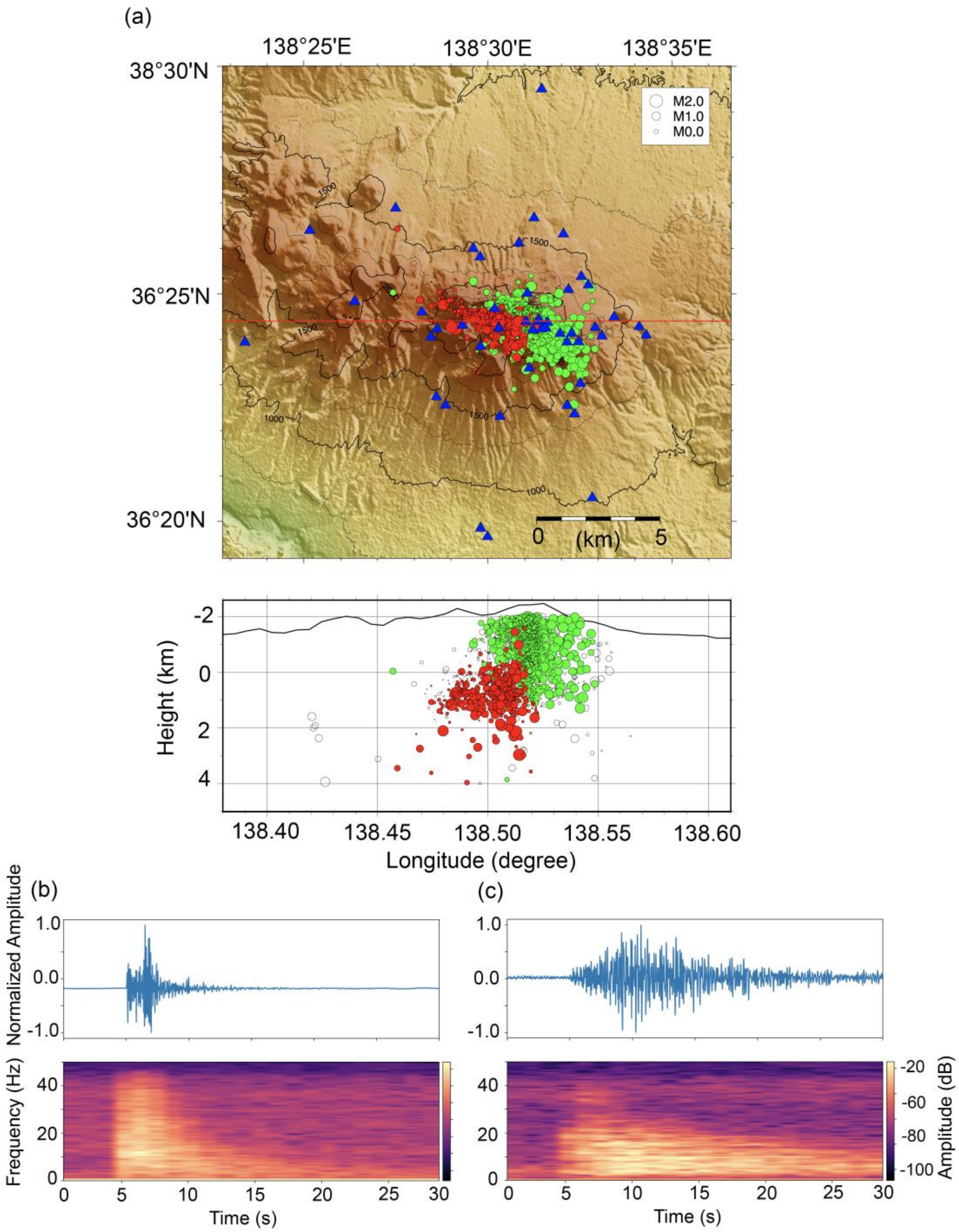

Figure 2. Distribution of earthquakes and stations and examples of input data. Earthquakes detected between January 2003 and October 2022, only with location determination errors less than 100 m in horizontal and depth directions, were plotted. (a) Distribution of earthquakes (circles) and stations (blue triangles) at Mount Asama. Red circles represent A-type, green circles represent B-type labeled events, and open circles represent earthquakes that do not belong to either type. The top image is a map view, and the bottom image is a cross-sectional view along the red line in the map view. Height is positive in the depth direction. (b) An example of an A-type earthquake. The top panel shows the seismic waveform, and the bottom image shows the spectrogram. (c) Same as b) but representing an example of a B-type earthquake.

Waveforms and spectrograms of A- and B-type earthquakes are shown in Figure 2b and 2c, respectively. A-type earthquakes have distinct P- and S-waves, whereas B-type earthquakes often have unclear P-wave onsets and indistinguishable S-waves. B-type earthquakes tend to have longer durations and lower dominant frequencies than type A earthquakes. Therefore, this study aimed to classify A- and B-type earthquakes and differentiate them from noise. We used a seismic catalog developed by the ERI Asama Volcano Observatory to construct the classification model between January 2003 and October 2022. The earthquake types in this catalog were manually classified based on visual inspection of waveforms. Figure 2a shows only those with horizontal and depth source errors within 100 m for 900 A-and 3191 B-type earthquakes. However, because the model training did not utilize information on the hypocenter locations, all detected earthquakes, including those with relatively large location errors, were used, totaling 1027 A-type earthquakes (8,318 waveforms) and 4090 B-type earthquakes (37,196 waveforms). The onsets of the P- and S-wave arrival times for these earthquakes were manually selected. Each dataset consisted of three channels (two horizontal and one vertical), and we decimated the waveforms of stations recorded at 200 Hz such that the sampling frequency of all waveforms was 100 Hz. Each waveform was initially extracted from a segment of 9001 points. For A- and B-type earthquakes, the starting point was randomly selected from 150 to 1000 points after the P-wave arrival time, and 3000 points were extracted for use. Noise data were obtained from areas with no earthquake signals.

4 Results

4.1. Model's performance

By incorporating 112,852 noise waveforms of the A- and B-type earthquakes previously described, 80% of the data were used for training and the remainder for validation to assess the model's performance. The model's training curve is shown in Figure S1. The model's performance was compared to that of a model consisting of four layers of CNN and two fully connected layers (Figure S2), as used by Nakano et al. 2022 (Table 1).

Table 1. Comparison of the performance of individual models

Model	This Study			CNN[a]			Under sampling[b]		
BACC	0.944			0.938			0.923		
Event Type	A	B	Noise	A	B	Noise	A	B	Noise
Precision	0.881	0.961	0.996	0.809	0.969	0.991	0.883	0.913	0.973
Recall	0.871	0.966	0.996	0.877	0.940	0.996	0.915	0.876	0.978
F1 score	0.876	0.964	0.995	0.842	0.955	0.993	0.899	0.894	0.976

The CNN model was created using the method described by Nakano et al. (2022).
[b] Models trained with the architecture developed in this study with the lowest number of training data.

The evaluation metrics calculated were the Balanced Accuracy (BACC), Precision, Recall, and F1-score, as expressed by the following equations:

$$BACC: \frac{\frac{TP}{TP+FN}+\frac{TN}{TN+FP}}{2} \tag{3}$$

$$Precision: \frac{TP}{TP+FP} \tag{4}$$

$$Recall: \frac{TP}{TP+FN} \tag{5}$$

$$F1\ Score: \frac{2TP}{2TP+FP+FN} \tag{6}$$

where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively. Compared to traditional methods, the classification performance evidently surpassed that of the other methods in all aspects, except for the recall of A-type and noise and the precision of B-type earthquakes (Table 1). The recognition accuracy for the A-type was lower than that for the B-type for both methods, likely owing to the smaller amount of training data available for the A-type. To address the imbalanced data issue, undersampling was performed to match the number of labels with the smallest dataset. Although undersampling slightly increased the A-type score, the overall performance decreased significantly (Table 1). Therefore, increasing the amount of data, even if imbalanced, can reduce the false negatives and positives, contributing to an overall improvement in model performance. Furthermore, because transformers have a scaling rule that improves their performance with increased data volume (Kaplan et al., 2020), increasing the amount of A-type data can lead to further improvements.

4.2. Model's Interpretability

Similar to the creation of many volcanic earthquake catalogs, the classification of earthquake types at Mount Asama is performed visually by humans based on the clarity of the P- and S-wave onsets and hypocenter depth. However, the decision-making process of deep learning models is generally considered a "black-box," making it difficult to discern the criteria used for their decisions. We attempted to visualize the decision-making criteria of our model by utilizing the AW calculated within the model.

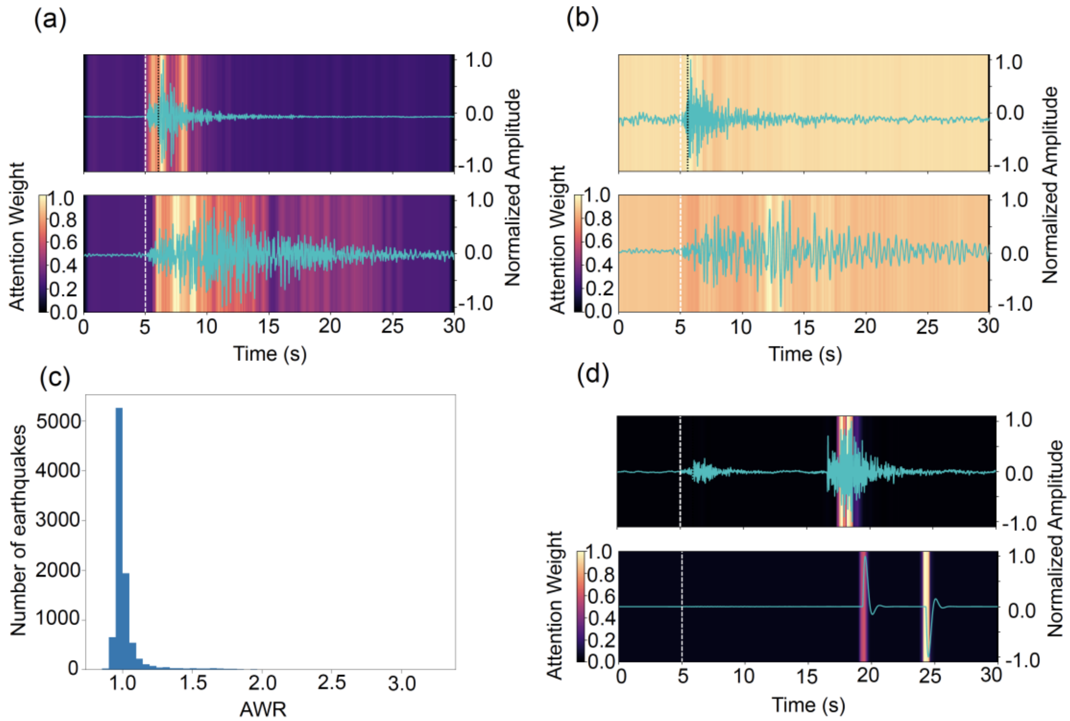

Figure 3. Prediction results of validation data. (a) An example where high AW is focused on the seismic signal. The upper figure shows a waveform classified as an A-type earthquake, and the lower figure shows one classified as a B-type earthquake. AW are normalized between 0 and 1.0. (b) Examples of dispersed AW. The upper figure shows a waveform classified as an A-type earthquake, and the lower figure shows one classified as a B-type earthquake. AW are normalized between 0 and 1.0. (c) The number of earthquakes relative to AWR. (d) Examples of inappropriate training data. The upper figure shows multiple earthquakes in a single dataset, and the lower figure shows a large non-earthquake signal. The white and black dotted lines in the waveforms represent the onset of P- and S-waves selected by experts, respectively.

We explored a method for directly visualizing the AW by averaging across the heads in the final layer and using only the AW related to the classification token. The visualization results showed that AW tended to concentrate mainly on the earthquake signal (Figure 3a). This suggests that our model focuses on features of seismic waveforms, such as how humans classify earthquakes. However, many cases were observed in which the AW was dispersed across the entire waveform despite high confidence in the classification (Figure 3b).

We extracted the AW for 0.1 s before and 5 s after the P-wave to investigate the specific types of waveforms where AW was concentrated or dispersed. We compared this with the AW from the start of the waveform to 1 s before the P-wave. This ratio is called the attention weight ratio (AWR). The visualization showed that many values were concentrated at approximately 1, indicating the dispersion of AW (Figure 3c). When considering the details of the data with an AWR below 1.2, some data contained multiple earthquakes, non-earthquake signals, inaccurate detection of P- and S-waves, and/or typical A-type earthquakes with prominent P- and S-waves labeled as B-type (Figures 3d and S2).

5 Discussions

The quality of training data, which hinders effective learning, is presumed to be a significant reason for the divergence in attention weights. Generally, B-type earthquakes are characterized by indistinct P- and S-waves. However, for Mount Asama, shallow earthquakes can be classified as B-type, regardless of their waveform characteristics. Therefore, we retrained the model by removing the data with S-wave detection from a B-type earthquake. The BACC was 0.943, and although there was no significant change, the AWR increased significantly (Figure 4a and 4b).
Regarding BACC, because the number of training data for the B-type decreased significantly to 27,405, we investigated the impact of substantially reducing the B-type training data. Training with a randomly sampled dataset from the entire B-type earthquake data to match the volume of data without detected S-waves resulted in a lower BACC of 0.920. The details of the performance comparisons are presented in Table S1. Therefore, S-waves in B-type earthquake data may act as noise in the learning process. However, we do not assert that earthquakes classified as B-type, which have waveforms similar to those of A-type earthquakes, are mislabeled simply because of their waveform similarity. Experts classified these earthquakes as B-type because of their features, which are distinct from those of other A-type earthquakes, such as extremely shallow epicenters. In our model, they were treated as noise because their characteristics differed from those of other B-type earthquakes, meaning that they could potentially be classified as either A-type or B-type, but approximately in between, or as a completely different type of earthquake. When comparing the spectra of the three types of earthquakes, the B-type with S-wave detection tended to have more prominent high frequencies compared to the B-type without S-wave detection, yet appeared to be more dominated by low frequencies than the A-type (Figure 4c). To clarify these distinctions, scrutinizing the waveforms and employing unsupervised learning techniques, such as clustering, to extract the distinct features of each earthquake type is essential.

After excluding the above data and apparently erroneous data, as shown in Figure 3d, events with low AWR were still observed (Figure 4b). A slight tendency for a higher AWR at shorter source-receiver distances was observed (Figure 4d). As many earthquakes occur beneath the crater, stations near the crater inevitably record more events. Therefore, we compared the proportion of data with a low AWR (<1.2) to the total data from all stations against the distance from the crater and found that the proportion of low AWR values increased with distance from the crater. Furthermore, by overlaying the distribution of all the data from each station with the distribution of data from sites with a prediction probability of over 99%, these distributions matched. Therefore, using stations near the crater is vital for enhancing both classification performance and interpretability (Figure 4e). This trend remained unchanged even after removing the waveforms of events with large location errors (Figure S3).

In this study, the model was specifically applied to data from Mount Asama. Applying this model to other volcanic regions presents several challenges. For instance, the model learns not only the characteristics of the seismic source but also incorporates features from the geological structure and the source-receiver distance. To adapt this model to other regions, it may be necessary to retrain it using data specific to the target volcano or to implement transfer learning techniques. Furthermore, as transformer models require extensive training data, maintaining high performance necessitates a dataset of size comparable to or greater than what was used in this study. If sufficient data cannot be prepared for the target volcano, data augmentation using seismic data from other volcanoes or synthetic data might be necessary.

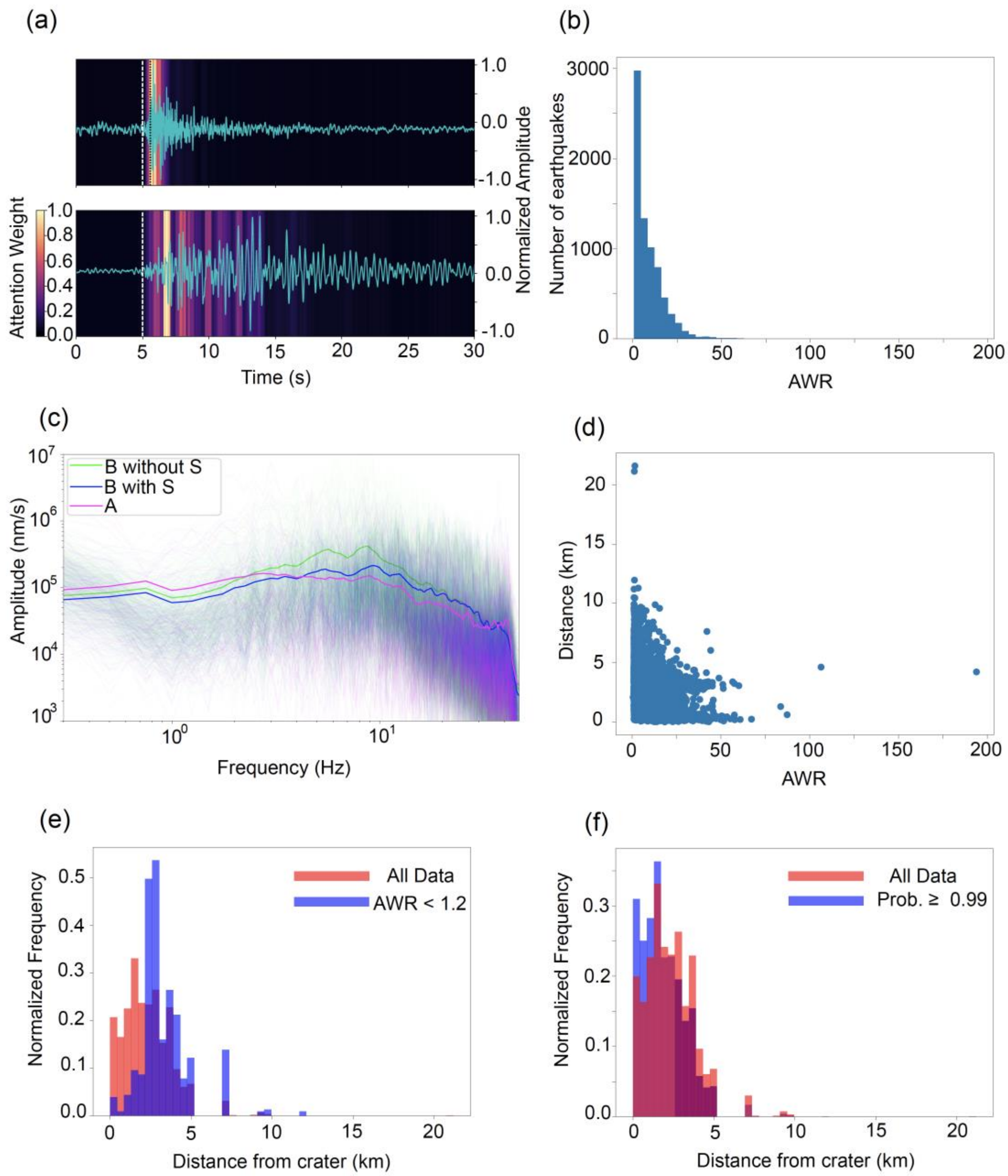

Figure 4. Evaluations using a model trained by excluding events with S-wave detection from B-type training data. (a) The same data is shown in Figure 3(b). (b) Number of earthquakes relative to AWR. (c) Comparison of spectra for A-type (magenta), B-type with S-wave detection (green), and B-type without S-wave detection (blue) recorded at KAE station (Lat: 36.407 Lon: 138.523, Figure 2a). Thin solid lines represent individual spectra, whereas thick solid lines indicate respective average values. (d) The source-receiver distance relative to AWR. (e) Histograms of distances from the crater (Lat.: 36.4, Lon.: 138.52) to each station, showing all test data (red) and data with AWR less than 1.2 (blue). The total data is normalized to sum up to one. (f) The horizontal axis and the red histogram are the same as in (e). Blue represents the data histogram

with a prediction probability value of over 99%. The total number of data is normalized to sum up to one.

6 Conclusions

In this study, we developed a volcanic earthquake classification model using a transformer encoder, demonstrating high performance comparable to or surpassing that of conventional methods. Data augmentation may mitigate this issue in the future using earthquakes from other regions. Interpretability using AW revealed that our model focused on waveform features, such as how humans classify earthquakes, although it did not elucidate any unique criteria beyond that.

AWR validation highlighted the importance of data near the crater in constructing a highly precise and interpretable model. Therefore, high-quality and consistent training data are required to enhance interpretability. Historically, earthquakes on Mount Asama have been classified by humans, but these classifications may be influenced by subjective biases or region-specific rules. However, this does not mean the labeling is incorrect; it may indicate a lack of data consistency for our model, which learns waveform patterns. Whether the earthquakes we excluded from the B-type accurately share equivalency with the A-type needs to be confirmed through waveform scrutiny and methods such as clustering to extract features of each earthquake type. Establishing new classification criteria can enable the application of deep learning classification models across various volcanoes. Additionally, the detection of new earthquake types can deepen our understanding of the relationship between earthquakes and volcanic activity.

Acknowledgments

The waveform data and earthquake catalogs used in this study were collected and created by the Asama Volcano Observatory at the University of Tokyo. Waveforms were recorded by stations operated by the ERI, NIED, and JMA. Maps were prepared using Geophysical Mapping Tools (Wessel et al., 2019), and other figures were created using Matplotlib, a Python module. This study was supported by JSPS KAKENHI, Grant No. 22K03752, 15KK0171, and the Next Generation Volcano Research and Human Resource Development Project of the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

Open Research

Waveform data were obtained from Hinet (https://hinetwww11.bosai.go.jp/auth/?LANG=en) and the Earthquake Research Institute, and the Japan Meteorological Agency via the JVDN system (https://jvdn.bosai.go.jp/app/pages/index.html?root=publicDataList&lang=en). The model developed in this study, training data, and seismic station lists were downloaded from out GitHub page (https://github.com/yugosuz/Volcanic-Earthquake-Classification-using-Transformer-Encoder).

References

Canário, J. P., Mello, R., Curilem, M., Huenupan, F., & Rios, R. (2020), In-depth Comparison of Deep Artificial Neural Network Architectures on Seismic Events Classification. *Journal of Volcanology and Geothermal Research*, *401*, 106881. doi:10.1016/j.jvolgeores.2020.106881.

Chouet, B. (1988), Resonance of a fluid-driven crack: radiation properties and implications for the source of long-period events and harmonic tremor. *Journal of Geophysical Research: Solid Earth*, *93*, 4375–4400. doi.org/10.1029/JB093iB05p04375

Gulati, A., Qin, J., Chiu, C., Parmar, N., Zhang, Y., Yu, J., Han, W., Wang, S., Zhang, Z., Wu, Y., & Pang, R. (2020), Conformer: Convolution-augmented transformer for speech recognition, *arXiv:2005.08100.* doi.org/10.48550/arXiv.2005.08100.

Iguchi, M. (2013), Magma Movement from the Deep to Shallow Sakurajima Volcano as Revealed by Geophysical Observations. *Bulletin of Volcanological Society of Japan*, *58*(1), 1–18.

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020), Scaling laws for neural language models. *arXiv:2001.08361*. doi.org/10.48550/arXiv.2001.08361.

Lara, F., Lara-Cueva, R., Larco, J. C., Carrera, E. V., & León, R. (2021), A deep learning approach for automatic recognition of seismo-volcanic events at the Cotopaxi volcano. *Journal of Volcanology and Geothermal Research*, *409*, 107142. doi:10.1016/j.jvolgeores.2020.107142.

Malfante, M., Mura, M.D., Mars, J.I., Metaxian, J.-P., Macedo, O., & Inza, A. (2018), Automatic classification of volcano seismic signatures, *Journal of Geophysical Research: Solid Earth*, *123*, 10,645–10,658.

McNutt, S. R. (1996). Seismic monitoring and eruption forecasting of volcanoes: A review of the state-of-the-art and case histories, *Monitoring and Mitigation of Volcano Hazards* (pp. 99–146). Berlin, Heidelberg: Springer Berlin Heidelberg. doi.org/10.1007/978-3-642-80087-0_3

Minakami, T., Utibori, S., Hiraga, S., Miyazaki, T., Gyoda, N., & Utsunomiya, T. (1970) Seismometrical Studies of Volcano Asama, Part I. Seismic and Volcanic Activities of Asama During 1934-1969. *Bulletin. Earthquake Research Institute*, *48*, 235–301.

Takeo, M., Aoki, Y., & Koyama, T. (2022), Recent volcanic activity at the Asama volcano and long-period seismic signals, *Proceedings of the Japan Academy, Series B*, *98*(8), 416–438. doi.org/10.2183/pjab.98.022

Mohammadi, A., Karimzadeh, S., Banimahd, S.A., Ozsarac, V., & Lourenco, P.B. (2023) The potential of region-specific machine-learning-based ground motion models: Application to Turkey. *Soil Dynamics and Earthquake Engineering*, *172*, 108008. doi.org/10.1016/j.soildyn.2023.108008.

Nakano, M., & Sugiyama (2022), D. Discriminating seismic events using 1D and 2D CNNs: applications to volcanic and tectonic datasets. *Earth Planets and Space*, *74*, 134. doi.org/10.1186/s40623-022-01696-1.

Nishimura, T. & Iguchi, M., (2011), Volcanic Earthquakes and Tremor in Japan, *Kyoto University Press*. ISBN: 4876989796, 9784876989799.

Oikawa, J., Ida, Y., & Tsuji H. (2006), Precise Hypocenter Determination for Volcanic Earthquakes Observed at Asama Volcano, Japan, *Bulletin of the Volcanological Society of Japan*, 51(2), 117–124.

Titos, M., Bueno, A., Garcia, L. & Benitez, C. (2018), A deep neural networks approach to automatic recognition systems for volcano-seismic events. *IEEE Journal of Selected Topics in*

Applied Earth Observations and Remote Sensing, *11*(5), 1533–1544. doi.org/10.1109/JSTARS.2018.2803198

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017), Attention is all you need. *Advances in Neural Information Processing Systems, 30*, 6000–6010.

Wessel, P., Luis, J. F., Uieda, L., Scharroo, R., Wobbe, F., Smith, W. H. F., & Tian, D. (2019), The generic mapping tools version 6. *Geochemistry, Geophysics, Geosystems, 20*(11), 5556–5564. doi.org/10.1029/2019GC008515.