https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

Multi-station automatic classification of seismic signatures from the Lascar volcano database

Pablo Salazar[1,2,3], Franz Yupanqui[4], Claudio Meneses[4], Susana Layana[1,2,3], and Gonzalo Yáñez[5]

[1]Núcleo de Investigación en Riesgo Volcánico-Ckelar Volcanes, Universidad Católica del Norte, Antofagasta, Chile.
[2]Departamento de Ciencias Geológicas, Universidad Católica del Norte, Chile.
[3]Centro de Investigación para la Gestión Integrada del Riesgo de Desastres (CIGIDEN), Santiago, Chile.
[4]Departamento de Ingeniería de Sistemas y Computación, Universidad Católica del Norte, Chile.
[5]Departamento de Ingeniería Estructural y Geotécnica, Pontificia Universidad Católica de Chile, Chile.

Correspondence: Pablo Salazar (pasalaz@ucn.cl)

Abstract. This study was aimed to build a multi-station automatic classification system for volcanic seismic signatures. This system was based on a probabilistic model made using transfer learning, which has, as the main tool, a pre-trained convolutional network named AlexNet. We designed four experiments using different datasets with data that was real, artificial, and two different combinations of these (combined 1 and combined 2). The experiment presented the highest scores when a process of data augmentation was introduced into processing sequence. Thus, the lack of real data in some classes (imbalance) dramatically affected the quality of the results, because the learning step (training) was over-fitted to the more numerous classes. To test the model stability with variable inputs, we implemented a k-fold cross-validation procedure. Under this approach, the results were more than optimal, considering that only the percentage of recognition of the tectonic events (TC) class was partially affected. The most valuable benefit of using this technique was that the use of volcano seismic signals from multiple stations provided a more generalisable model which, in near future, can be extended to multi-volcano database systems. The impact of this work is significant in the evaluation of hazard and risk by monitoring the dynamic evolution of volcanic centres, which is crucial for understanding the stages in a volcano's eruptive cycle.

1 Introduction

1.1 The problem of monitoring

The task of detecting the seismic activity of an active volcano and the subsequent characterisation (classification) of these events are, in many cases, the most time-consuming in observatories worldwide. This is because of the massive amount of data that is collected daily in a continuous record, by a single seismic network with a few stations. In this context, big data analysis tools have become an attractive option for reaching levels of processing that have never been achieved using traditional techniques.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022


Natural Hazards and Earth System Sciences Discussions

This study proposed the use of machine learning and transfer learning techniques to automatically classify volcanic seismic events. Determining the type of volcanic event in a continuous seismic record (time series) will facilitate the construction of a model of evolution associated with the dynamics of a volcanic centre. This should create better understanding and evaluation of the hazards and risks associated with volcanic activity, improving efforts made in this matter (e.g., de Natale et al. , 2019; Magrin et al. , 2017; Rapolla et al. , 2010). The novelty of this approach is in its use of previously trained deep convolutional networks, such as AlexNet, in a scenario that considers the information recorded by a network of multiple stations. It permits variability in the input of data and improves the generalisation of the system. The generalisation of the system directly impacts the performance of the training models of pattern recognition for each class and creates the possibility of applying the generated models in a common multi-volcano database system in the near future.

1.2 A summary of methods used in Volcano Seismic Recognition

Among the classification techniques, methods based on a probabilistic approach and hidden Markov models (HMMs) are most relevant. The advantages of HMMs include the possibility of managing data with different durations, computational efficiency, and an elegant interpretation of results based on Bayes' theorem (Carniel, 2014). Several studies have been performed on volcanic systems using this technique with different approaches. Continuous HMMs were used for the simultaneous detection and classification of continuous volcanic responses (Beyreuther and Wassermann , 2008), whereas discrete HMMs were applied to analyse and classify events as described by Ohrnberger (2001). Other applications for HMMs were considered in the works of Bebbington (2007), who used the method to analyse a catalogue of flank eruptions recorded at Mt. Etna. Hidden semi-Markov models were applied by Beyreuther and Wasserman (2011) using time dependence to improve the performance of the method. Beyreuther et al. (2012) also introduced state clustering to improve the time discretisation in induced seismicity experiments. Contrastingly, Bicego et al. (2013) used an HMM method, based on a hybrid generative-discriminative classification paradigm, in pre-triggered signals recorded at the Galeras Volcano in Colombia.

Other classification techniques, such as artificial neural networks, provide an efficient approach for the classification of not only seismic events, but also time slices of continuous signals, such as volcanic tremors (Carniel, 2014). The multi-layer perceptron (MLP) is often used for the analysis of seismic signals recorded at volcanoes. Esposito et al. (2013) applied the MLP technique for landslide recognition, while Esposito et al. (2014) utilized MLP to estimate the possible trend of the seismicity level in Campi Flegrei (Italy). Self-organising maps (SOM), another class of artificial neural networks, have been used to analyse very long period events at the Stromboli volcano (Esposito et al. , 2008), as well as volcanic tremors at the Etna volcano (Langer et al. , 2009, 2011), Raoul Island volcano (Carniel et al. , 2013a), and Ruapehu volcano (Carniel et al. , 2013b). Furthermore, self-organising maps with time-varying structures (SOM-TVS) have been applied to volcanic signals to achieve improvements in relation to SOMs (Araujo and Rego , 2013).

Notably, the support vector machine (SVM) approach developed by Vapnik (1995), which is based on linear discrimination, should be mentioned. For a two-class problem, SVM uses linear elements for discrimination, i.e., lines, planes, or hyperplanes. Masotti et al. (2006, 2008) used this technique in analysing volcanic tremor data recorded at Mt. Etna in 2001. Langer et al. (2009) applied this approach to compare several supervised and unsupervised pattern-classification techniques. Ceamanos et

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022


Natural Hazards and Earth System Sciences Discussions

al. (2010) built a multi-SVM classifier for remote-sensing hyperspectral data. The simultaneous application of SVM and MLP was also performed by Giacco et al. (2009), who used the two methods to discriminate between explosion quakes, landslides, and tremors recorded at the Stromboli volcano.

Thus, numerous studies have been conducted to develop an automated system for the detection and classification of volcanic signals. The early systems consisted of classifiers that used data from a single station to design different approaches (Masotti et al. , 2006; Beyreuther and Wassermann , 2008; Rouland et al. , 2009; Langer et al. , 2011; Bicego et al. , 2015). However, after some years, the systems evolved into more complex algorithms that facilitated the building of models using the information from a few stations or channels (Z, E, N). Nevertheless, they did not use the data from all the possible stations in the network, instead their results were based on one station or channel that was used as a pattern (Álvarez et al. , 2012; Esposito et al. , 2013; Carniel et al. , 2013b; Cortés et al. , 2014, 2015; Curilem et al. , 2014a, b; Bicego et al. , 2015). Interestingly, the work of Curilem et al. (2016), based on station-dependent classifiers, shows the possibility to mix information from different stations to create models that enable the classification of events at different stations, despite the fact that experiments were performed with a reduced database.

1.3 Supervised machine learning as strategy for automatisation

Volcano-seismic recognition is an automated system that allows us, at an early stage, to build probabilistic models for each volcanic event or class from classified data determined previously by an expert geophysicist. The models obtained are later used over continuous seismic records for automatic and unsupervised classification.

As previously mentioned, pattern recognition and automatic classification require the previous classification of seismic signals into different classes, making this one of the most important, but also one of the most time-consuming, tasks when accomplished daily by a human operator.

The study was aimed to present a novel approach that considered a supervised machine-learning strategy (transfer learning) using AlexNet, a previously trained deep convolutional neural network, to create a multi-station automatic classification system for volcanic seismic signatures.

2 Methodological testing site

2.1 Seismic monitoring of Lascar volcano

The Lascar volcano (23°22′ S, 67°44′ W; 5.592 m a.s.l.) is located in northern Chile, 270 km NE from Antofagasta and 70 km SE from San Pedro de Atacama, on the western border of the Altiplano-Puna 'plateau' (Figure 1). Lascar is considered the most active volcano in the Central Andean Volcanic Zone (de Silva and Francis , 1991). It is a compound elongated strato-volcano, comprised of two truncated western and eastern cones (Gardeweg et al. , 1998) that host five nested craters aligned ENE-WSW. The Lascar volcano has been seismically monitored by the CKELAR-VOLCANES group using a temporal network of 11 three-component stations (Shallow Posthole Seismometers, Model F72-2.0). These short-period 2 Hz seismometers were

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022


Natural Hazards and Earth System Sciences Discussions

monitored continuously at 200 Hz from March to October 2018 in this first step of processing. Notably, only the Z channel was considered in building our database; the use of the other channels is reserved for future studies.

2.2 Lascar's database

Lascar's database corresponds to a catalogue of 6,145 seismic events, from which only 3,947 can be classified as volcanic events. The others, based on the distance to the hypocentres, are mainly tectonic events (not directly related to volcanic activity) recorded by Lascar's network during the period of observation. To guarantee the reliability of the database regarding volcanic activity, all observations were manually segmented, labelled, and checked from the continuous seismic record by CKELAR-VOLCANES experts. Our analysts were able to recognise five classes of events (Figure 2): the hybrid events (HY) corresponded to 213 signals; long-period catalogue (LP) corresponded to 577 events; the number of tremors (TR) was 471; volcano-tectonic (VT) activity was recorded in 2,686 events; and, finally, tectonic events (TC), not related to volcanic activity, were in 2,198 events (Tables 1 and 2).

3 Methods

We proposed a framework based on transfer learning, consisting of a previously trained deep convolutional neural network called AlexNet (Krizhevsky et al. , 2012). The main reasons for adopting this approach were, on one hand, to avoid the steps of an extensive search and selection of features, and, on the other hand, to deal with the limited number of labelled data. The data considered as input were spectrograms of the labelled dataset (waveform) from each class of seismic event. The line of processing consists of five steps: the first step describes the process applied in the time series (seismogram); the second step details obtaining spectrograms from the labelled seismograms; the third step consists of using the data augmentation technique to increase the number of data in the training and test datasets; the fourth step indicates how the prediction model is built; and, finally, the fifth step is performed to estimate the model's performance. In the following paragraphs, each step is explained in detail.

Step 1 (pre-processing): The observations were recorded with the same type of sensors, but at times the sample ratio should have been varied because of technical problems. Thus, many of the stations recorded at 200 Hz, but a few stations recorded at 500 Hz or 100 Hz. Considering this, we decided to resample the entire time series at 100 Hz. Thereafter, we removed the mean and normalised the signals. Following this process, we applied a 10th order Butterworth bandpass filter between 1 and 10 Hz. The entire pre-processing was implemented using ObsPy (Beyreuther et al. , 2010).

Step 2 (spectrogram): The spectrograms were calculated by applying a Short-Time Fourier Transform, using the formula:

$$S[x(t)](n,k) = \left| \sum_{m=0}^{N-1} x(m) \cdot w(m-n) \cdot e^{i2\pi mk} \right|^2 \quad (1)$$

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022


Natural Hazards and Earth System Sciences
Discussions

where $x(t)$ and $y(t)$ represent the seismic signal and short-time fourier transform sliding window, respectively. The sliding window size was set to 1 s with a 95% overlap. The frequency interval used to calculate the spectrograms ranged from 1 to 10 Hz. All spectrograms were transformed to RGB images of 224 $\times$ 224 pixels.

Step 3 (data augmentation): Due to the imbalance produced on the labelled data, we decided to apply the data augmentation technique to generate a balanced number of different classes of seismic events (Table 2). Considering all the techniques for data augmentation, a time stretch was chosen. This transformation was implemented by rotations around the frequency axis that were as random as possible, between 5-25% of the length of the signal, which were applied at the beginning, middle, or ending of the spectrogram as appropriate (Figures 3–5). The amount of data that was created for each class depended on the number of events in the most populated class; in this particular case the VT class was used as reference. The new signals generated were processed using the same procedure applied to the original signals (see Steps 1 and 2).

Step 4 (AlexNet): AlexNet is a deep convolutional neural network proposed by Krizhevsky et al. (2012) to classify the 1.2 million high-resolution images in the ImageNet LSVRC-2010 contest into 1,000 different classes. AlexNet was used in our study as a pre-trained deep convolutional neural network for spectrogram recognition mainly because the spectrogram can be easily represented as an RGB image of 224 $\times$ 224 pixels.

Step 5 (model performance): This step was considered to evaluate the performance of the model and, thus, validate the classification. This step was executed through a set of tests composed of signals that were not considered in the training step. We considered the following measures from the TorchMetrics in PyTorch:

i) Accuracy:

$$Accuracy = \frac{1}{N}\sum_{1}^{N} 1(y_i - \hat{y}_i) \tag{2}$$

where y_i and $\hat{y}_i$ are the tensor of the target values and the tensor of the predictions, respectively.

ii) F1-score:

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{3}$$

iii) Recall:

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

where TP and FN represent the number of true positives and false negatives, respectively.

iv) Precision:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

where TP and FP represent the number of true positives and false positives, respectively.

v) Cohen's kappa:

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \tag{6}$$

where p_0 is the empirical probability of agreement and p_e is the expected agreement when both annotators assign labels randomly. Note that is estimated using a per-annotator empirical prior over the class labels.

4 Results

We designed four experiments to test the model's ability to classify the data: (a) real data was used to build the models and test them using a real database without the application of augmentation processes; (b) artificial data was used to build the models and test them using artificial data created by augmentation processes (see Step 3); (c) combined 1 used exclusively artificial data to build models and real data to test them; and (d) combined 2 used real and artificial data to build the models, and exclusively real data was used to test them.

Tables 3 and 4 present the statistics and metrics of the different experiments. It is evident that in the case of the unbalanced database, in experiment (a), the performance of the experiment was inferior; however, when the data augmentation process was applied, in experiments (b–d), the results reached a particularly good percentage of the metric parameters. We want to highlight the results of the experiment, where the real data were used exclusively in the test (without being included in the model building). In this case, the metrics showed the second highest ranking, preceding only the experiment that exclusively included artificial data.

In analysing the complete training and validation phase, through the loss and accuracy epoch, the process was less effective when the (a) database was used (Figure 6), whereas the best performance was achieved by the process using the (b) database (Figure 7).

To evaluate the accuracy of the classification process for each experiment, we computed the confusion matrix. In the case of experiment (a), we noticed that classes with less data were negatively affected. In this case, a high percentage of HY events were confused by TC events and, thus, the TR events were also mistaken for TC events (Figure 8). Conversely, the classification process of experiment (b) performed well, and no confusion in the class recognition was found (Figure 9). For experiment (c), the performance was high, considering that the confusion of classes was practically minimal (Figure 10). Experiment (d) also presented a high score, and the confusion matrix indicated a small problem in the recognition of TC and VT classes, in which some are confused by TR and TC, respectively (Figure 11).

To test whether the probabilistic model, built with both real and augmented data, is stable under the variability of the input data, a k-fold cross-validation procedure was implemented (Table 5). The results showed that the metrics for the different classes, as in the previous experiments, were high. There was only one class (TC) affected (with more variability), but good scores were always maintained, which in the worst case was over 67%.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022


Natural Hazards and Earth System Sciences
Discussions

5 Discussion

The analyses of this study clearly showed the impact of imbalance in the database and how the process of machine learning was conditioned to build probabilistic models of classifications for the different classes. The model building process was notably influenced by classes with more events, which in this case were TC and VT. Under these conditions, an over-fitted model was built in the training phase, which largely coincided with the recognition of these two classes, to the detriment of the less numerous ones.

Conversely, the process of data augmentation facilitated a balance in the data of the different classes, which directly impacted the performance of the probabilistic model, reaching optimal scores over different test datasets. This can be explained by the fact that data augmentation provides an efficient process for searching the features of each class during the training process. Although the frequency content was the main classification characteristic to differentiate the different volcanic event types, the choice of stretch as a data augmentation method along the time axis was a successful strategy that did not reduce effectiveness. The artificial production of data using stretches in frequency should provoke overlap (on the dimensional map) between the different classes of volcanic events when these are analysed using the spectrogram image.

Another important topic is the technique used to build the probabilistic model. Transfer learning with a pre-trained large neural network, such as AlexNet, facilitated model building in less time. Instead of manually selecting the best features, this task was performed by learning features from the pre-trained models of AlexNet, thereby dramatically saving time in the process. This point was verified when the model built using AlexNet was validated using only real data in experiment (c), where high scores in the metrics of the experiment were achieved.

6 Conclusions

From the experiments implemented in this study, the following conclusions were drawn. First, the usefulness of a multi-station framework to build a probabilistic model of the Lascar database allowed us to obtain a more generalisable model, thereby avoiding the bias associated with the choice of a particular station to retrieve the features. This fact led us to believe that we are very close to obtaining high-score results, with the AlexNet tool playing a key role in reaching the challenge of building a multi-volcano probabilistic model to classify the seismic events. Second, data augmentation plays a key role as the main factor to improve the metrics of the experiments, thereby providing a built model validated by real data. Last, and perhaps the most relevant, the proposal based on transfer learning provided an efficient feature retrieval process using learning features. The performance of this approach was clearly superior when compared with an exhaustive process of evaluation for a list of an hundred statistical features sent to the system, as it is a process of designed features. Our approach has a high impact when the time process matters, as is the case in early warning systems for volcanic activity, and provides a more generalisable model of prediction. The acquisition of more generalisable models creates a good opportunity to develop a multi-volcano probabilistic model for volcanoes worldwide. This will improve the understanding and evaluation of the hazards and risks associated with the activity of volcanoes.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

Data availability. The data are registered by DOI: 10.5281/zenodo.6001869. These can be founded in the link:
https://zenodo.org/record/6001870#.YkGXKffQ-Xl

Competing interests. The author has declared that there are no competing interests.

Acknowledgements. This work has been funded by ANID/FONDECYT-INICIACIÓN/11190190, Antofagasta Regional Government, FIC-R project, code BIP N°30488832-0 and by Research Center for Integrated Disaster Risk Management (CIGIDEN), ANID/FONDAP/15110017. Our acknowledgment to CKELAR-VOLCANES team: Gabriel Ureta, Javier Valdés, Christian Ibaceta, Felipe Aguilera, Felipe Rojas, Diego Jaldín, Álvaro Vergara, Alfredo Esquivel, José Pablo Sepúlveda, Manuel Inostroza, Cristóbal González. To Greg Wait for his priceless course of training in volcanic signal.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022


Natural Hazards and Earth System Sciences Discussions

References

Álvarez, I., García, L., Cortés, G., Benítez, C., and De la Torre, A.: Discriminative feature selection for automatic classification of volcano-seismic signals, *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 151–155, 2012.

Araujo, AFR. and Rego, RLME.: Self-organizing maps with a time-varying structure. *ACM Comput. Surv.* 46, 1, Art. 7, 38 pp. DOI: http://dx.doi.org/10.1145/2522968.2522975, 2013.

Bebbington, M.: Identifying volcanic regimes using Hidden Markov Models, *Geophys. J. Int.*, 171, 921–942, doi: 10.1111/j.1365-246X.2007.03559.x, 2007.

Beyreuther, M. and Wassermann, J.: Continuous earthquake detection and classification using discrete Hidden Markov Models, *Geophys. J. Int.*, 175, 1055–1066, 2008.

Beyreuther, M., Barsch, R., Krischer, L., Mejies, T., Behr, Y. and Wassermann, J.: Obspy: A Python Toolbox for Seismology. *Seismological Research Letters*, 81(3): 530 – 533. doi: https://doi.org/10.1785/gssrl.81.3.530, 2010.

Beyreuther M., Wassermann, J.: Hidden semi-Markov model based earthquake classification system using weighted finite-state transducers. *Nonlinear Process Geophys.* 18(1):81–89, 2011.

Beyreuther, M., Hammer, C., Wassermann, J., Ohrnberger, M. and Megies, T.: Constructing a Hidden Markov Model based earthquake detector: application to induced seismicity, *Geophys. J. Int.*, 189, 602–610, 2012.

Bicego, M., Londoño-Bonilla, J.M., Orozco-Alzate, M.: Volcano-seismic events classification using document classification strategies. *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 9279, pp. 119–129, 2015.

Bicego, M., Acosta-Munoz, C., Orozco-Alzate, M.: Classification of seismic volcanic signals using hidden-Markov-model-based generative embeddings. *IEEE Trans. Geosci. Remote Sens.* 51(6):3400–3409, 2013.

Carniel, R.: Characterization of volcanic regimes and identification of significant transitions using geophysical data: a review. *Bull. Volcanol.* 76 (8), 1–22, 2014.

Carniel, R., Barbui, L., Jolly, AD.: Detecting dynamical regimes by Self-Organizing Map (SOM) analysis: an example from the March 2006 phreatic eruption at Raoul Island, New Zealand Kermadic Arc. *Boll. Geofis. Teor. Appl.* 54(1):39–52, 2013a.

Carniel, R., Jolly, AD., Barbui, L.: Analysis of phreatic events at Ruapehu volcano, New Zealand using a new SOM approach. *J. Volcanol. Geotherm. Res.* 254:69–79, 2013b.

Ceamanos, X., Waske, B., Benediktsson, J.A., Chanussot, J., Fauvel, M., Sveinsson, J.R.: A classifier ensemble based on fusion of support vector machines for classifying hyperspectral data. *Int. J. Image Data Fusion* 1(4):293–307, 2010.

Cortés, G., García, L., Álvarez, I., Benítez, C., de la Torre, T., Ibáñez, J. Parallel system architecture (PSA): An efficient approach for automatic recognition of volcano-seismic events. *J. Volcanol. Geotherm. Res.* 271, 1–10, 2014.

Cortés, G., Benitez, M.C., Garcia, L., Alvarez, I., Ibanez, J.M.: A comparative study of dimensionality reduction algorithms applied to volcano-seismic signals. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* http://dx.doi.org/10.1109/JSTARS.2015.2479300, 2015.

Curilem, M., Vergara, J., San, Martin C., Fuentealba, G., Cardona, C., Huenupan, F., Chacón, M., Khan, S., Hussein, W., Becerra, N.: Pattern recognition applied to seismic signals of the Llaima Volcano (Chile): an analysis of the events' features. *J. Volcanol. Geotherm. Res.* 282, 134–177, 2014a.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Curilem, M., Huenupan, F., San, Martin C., Fuentealba, G., Cardona, C., Franco, L., Acuña, G., Chacón, M.: Feature analysis for the classification of volcanic seismic events using support vector machines. In: Gelbukh, A., et al. (Eds.), MICAI 2014, Part II, Lecture Notes in Artificial Intelligence, LNAI 8857. Springer International Publishing Switzerland, pp. 160–171, 2014b.

Curilem, M., Huenupan, F., Beltrán, D., San Martin, C., Fuentealba, G., Franco, L., Cardona, C., Acuña, G., Chacón, M., Salman Khane, M.S., Becerra Yoma, N.: Pattern recognition applied to seismic signals of Llaima volcano (Chile): An evaluation of station-dependent classifiers. *J. Volcanol. Geotherm. Res.* 315, 15–27, 2016.

de Natale, G., Petrazzuoli, S., Romanelli, F., Troise, C., Vaccari, F., Somma, R., Peresan, A., Panza, G.F. (2019). Seismic risk mitigation at Ischia island (Naples, Southern Italy): An innovative approach to mitigate catastrophic scenarios, *Engineering Geology*, Volume 261, 105285, ISSN 0013-7952, https://doi.org/10.1016/j.enggeo.2019.105285.

de Silva, S., Francis, P.: Volcanoes of Central Andes, Springer-Verlag, New York, 1991.

Esposito, A.M., Giudicepietro, F., D'Auria, L., Scarpetta, S., Martini, M., Coltelli, M., Marinaro, M.: Unsupervised neural analysis of very long period events at Stromboli volcano using the self-organizing maps. *Bull. Seismol. Soc. Am.* 98:2449–2459. doi:10.1785/0120070110, 2008.

Esposito, A., D'Auria, L., Giudicepietro, F., Peluso, R., Martini, M.: Automatic recognition of landslides based on neural network analysis of seismic signals: an application to the monitoring of Stromboli volcano (Southern Italy). *Pure Appl. Geophys.* 170:1821–1832, 2013.

Esposito, A., D'Auria, L., Angelillo, A., Giudicepietro, F., Martini, M.: Predictive analysis of the seismicity level at Campi Flegrei volcano using a data-driven approach. In "Recent advances of neural network models and applications", Proceedings of the 23rd Workshop of the Italian Neural Networks Society (SIREN), May 23–25, Vietri sul Mare, Salerno, Italy. Series "Smart Innovation, Systems and Technologies" Vol. 26, Springer, pp 133–145, 2014.

Krizhevsky, A., Sutskever, I. and Hinton, G.: ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems.* 25. 10.1145/3065386, 2012.

Gardeweg, M., Sparks, R., Matthews, S.: Evolution of Lascar volcano, Northern Chile, *J. Geol. Soc.*, 155 (1998), pp. 89-104, 1998.

Giacco, F., Esposito, A.M., Scarpetta, S., Giudicepietro, F., Matinaro, M.: Support vector machines and MLP for automatic classification of seismic signals at Stromboli Volcano. Proceedings of the 2009 Conference on Neural Nets WIRN09, pp. 116–123, 2009.

Magrin, A., Peresan, A., Kronrod, T., Vaccari, F., Panza, G.F.: Neo-deterministic seismic hazard assessment and earthquake occurrence rate, *Engineering Geology*, Volume 229, Pages 95-109, ISSN 0013-7952, https://doi.org/10.1016/j.enggeo.2017.09.004, 2017.

Masotti, M., Falsaperla, S., Langer, H., Spampinato, S., Campanini, R.: Application of support vector machine to the classification of volcanic tremor at Etna, Italy. *Geophys. Res. Lett.* 33, 2006.

Masotti, M., Campanini, R., Mazzacurati, L., Falsaperla, S., Langer, H., Spampinato, S.: TREMOrEC: a software utility for automatic classification of volcanic tremor. *Geochem. Geophys. Geosyst.* 9, Q04007. doi:10.1029/2007GC001860, 2008.

McFee, B., Humphrey, E.J., and Bello, J.P.: A software framework for musical data augmentation, in 16th International Society for Music Information Retrieval Conference, ser. ISMIR, 2015.

Langer, H., Falsaperla, S., Masotti, M., Campanili, R., Spampinato, S., Messina, A.: Synopsis of supervised and unsupervised pattern classification techniques applied to volcanic tremor data at Mt. Etna, Italy. *Geophys. J. Int.* 178:1132–1144. doi:10.1111/j.1365-246X.2009.04179.x, 2009.

Langer, H., Falsaperla, S., Messina, A., Spampinato, S., Behncke, B.: Detecting imminent eruptive activity at Mt Etna, Italy, in 2007–2008 through pattern classification of volcanic tremor data. *J. Volcanol. Geotherm. Res.* 200:1–17, 2011.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022


Natural Hazards and Earth System Sciences
Discussions

Ohrnberger, M.: Continuous automatic classification of seismic signals of volcanic origin at Mt. Merapi, Java, Indonesia. Ph.D. thesis, Universität Potsdam, Germany, 2001.

Rapolla, A., Paoletti, V., Secomandi, M.: Seismically-induced landslide susceptibility evaluation: Application of a new procedure to the island of Ischia, Campania Region, Southern Italy, *Engineering Geology*, Volume 114, Issues 1–2, Pages 10-25, ISSN 0013-7952, https://doi.org/10.1016/j.enggeo.2010.03.006, 2010.

Rouland, D., Legrand, D., Zhizhin, M., Vergniolle, S.: Automatic detection and discrimination of volcanic tremors and tectonic earthquakes: an application to Ambrym Volcano, Vanuatu. *J. Volcanol. Geotherm. Res.* 181 (3–4), 196–206, 2009.

Vapnik, V.: The Nature of Statistical Learning Theory. Springer Verlag, 1995.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

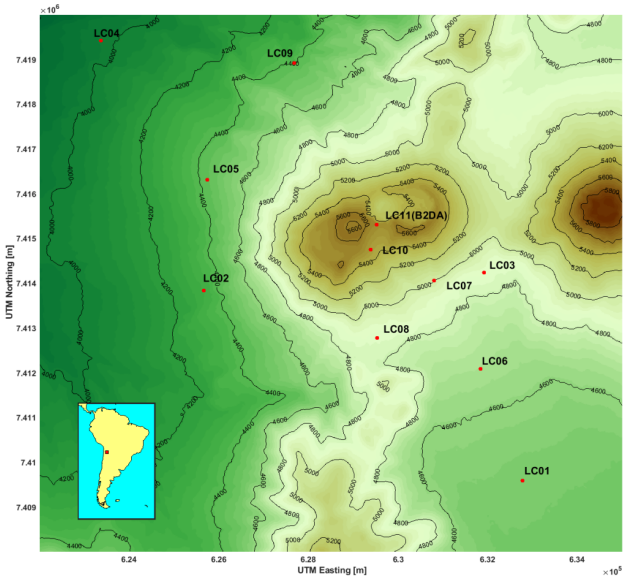

Figure 1. Location map of the Lascar volcano experiment. The red dots indicate the position of the short-period seismic stations and the black bold text represents the corresponding names. The LC10 and LC11 (B2DA) are located eastward, near the crater of the Lascar volcano. The DEM data is a product of ASTER Global Digital Elevation Model version 3 (ASTGTM v003), these can be downloaded directly by OPeNDAP link (https://lpdaac.usgs.gov/tools/opendap/).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

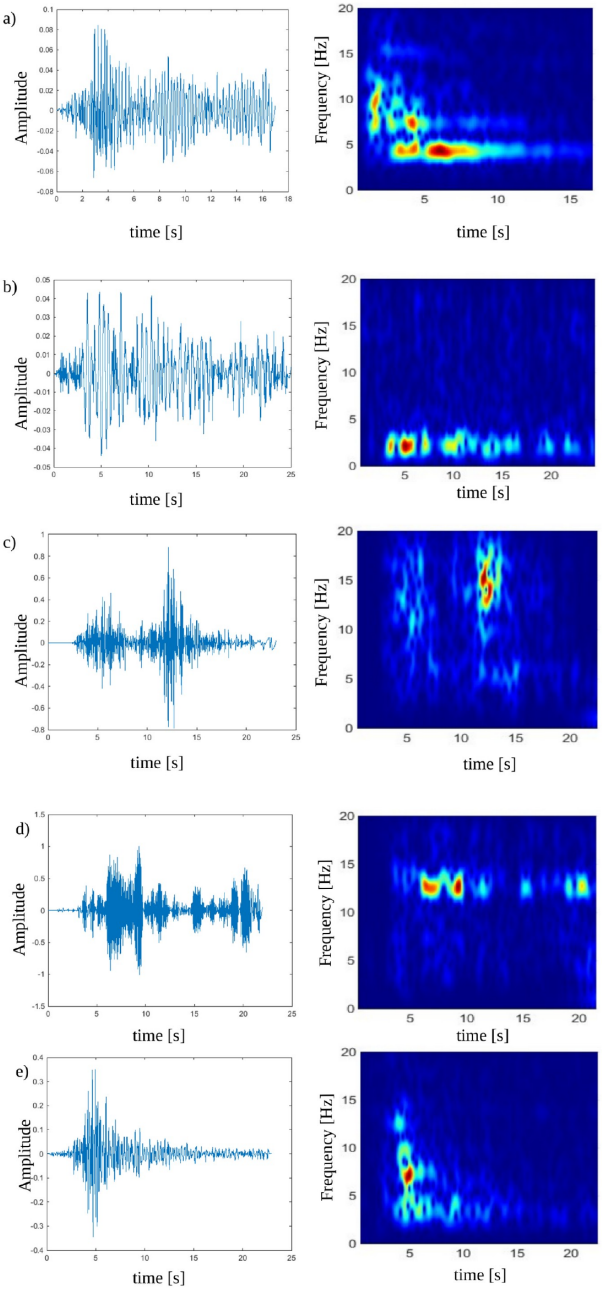

Figure 2. Examples of time series (left) and spectrograms (right) for the different classes in the Lascar database: a) hybrid events (HY), b) long period (LP), c) tectonic events (TC), d) tremors (TR), and e) volcano-tectonic (VT).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

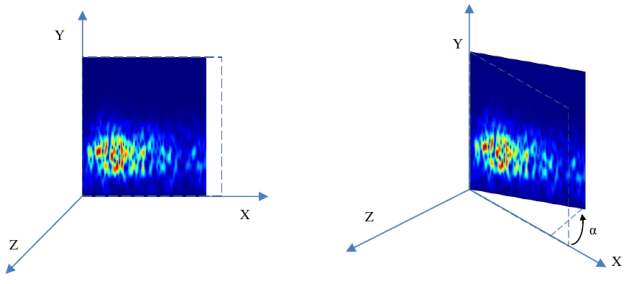

Figure 3. Transformation by rotations; the time stretching is produced when we change the angle of rotation around the frequency axis.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

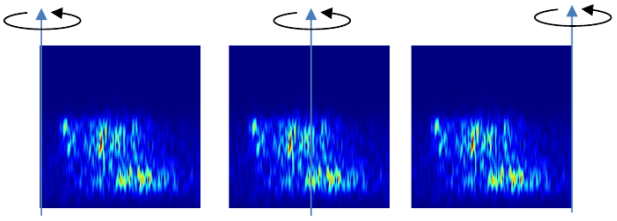

Figure 4. Three possibilities of rotations, around a left, central, and right axis, respectively.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

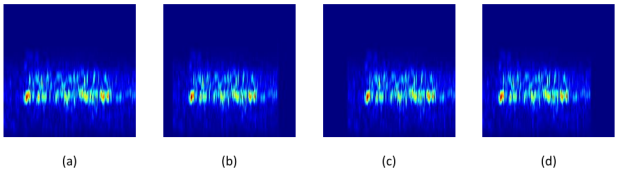

Figure 5. Examples of how the time stretching is produced by rotations: (a) original spectrogram, (b) rotation of 19% around the central axis, (c) rotation of 23% around a right axis, (d) rotation of 24% around a left axis.

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

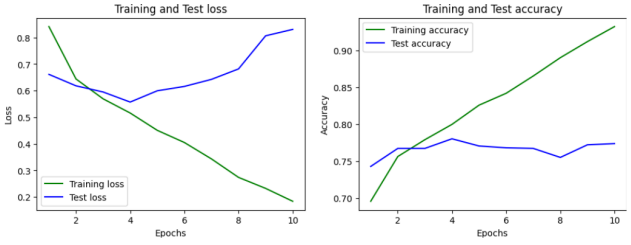

Figure 6. Training and validation performance scores of the transfer learning AlexNet for experiment (a).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

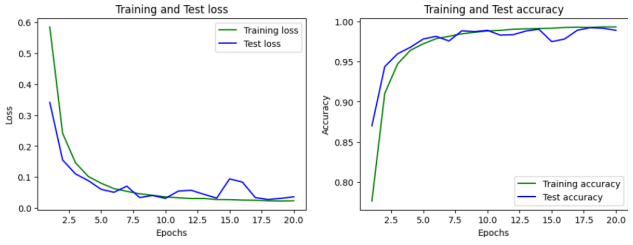

Figure 7. Training and validation performance scores of the transfer learning AlexNet for experiment (b).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

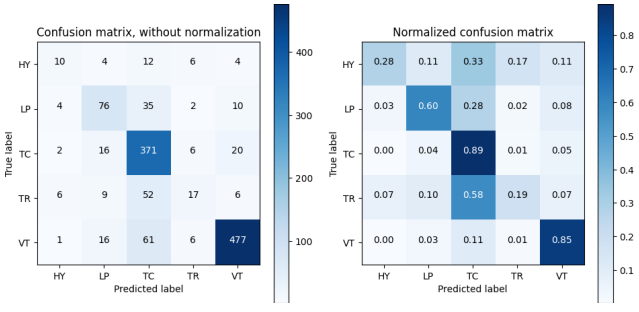

Figure 8. Confusion matrices, without normalisation (left) and normalised (right), for experiment (a).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

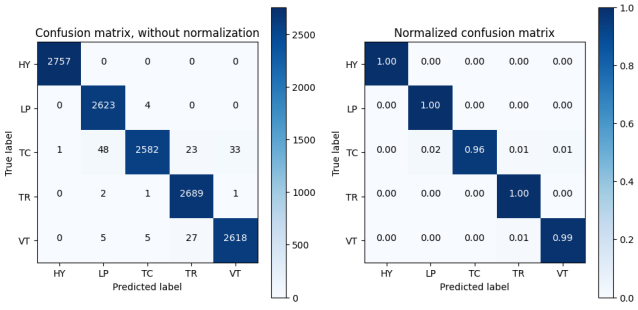

Figure 9. Confusion matrices, without normalisation (left) and normalised (right), for experiment (b).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

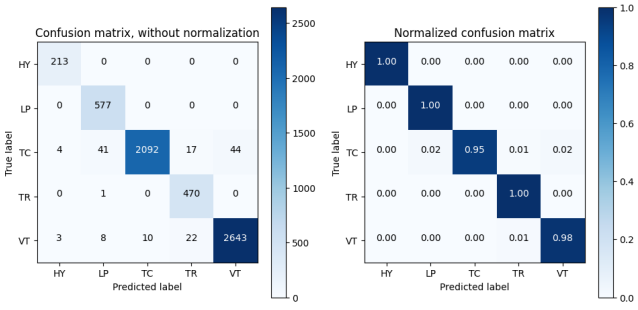

Figure 10. Confusion matrices, without normalisation (left) and normalised (right), for experiment (c).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions
Open Access
EGU

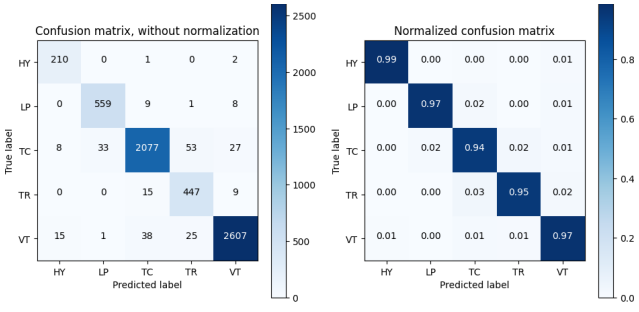

Figure 11. Confusion matrices, without normalisation (left) and normalised (right), for experiment (d).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

Table 1. Contribution of each station to Lascar's database.

Classes/Stations	LC01	LC02	LC03	LC04	LC05	LC06	LC07	LC08	LC09	LC10	B2DA
HY	0	0	0	7	45	4	0	3	12	136	6
LP	11	0	4	2	142	28	5	60	23	195	107
TC	9	0	11	97	284	36	157	127	144	1171	162
TR	1	0	5	29	21	4	61	13	35	292	10
VT	2	0	18	52	84	6	206	14	41	2249	14

This experiment considered only the Z channels (vertical).

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

Table 2. Statistics of the seismic records, based on the majority of the records.

Classes/Events	Quantity %	Quantity N°	Maximum	Minimum	Mean [s]	Median [s]	Mode [s]	Standard deviation [s]
HY	3	213	265	10	57.2	59	40	35.3
LP	9	577	600	9	82.3	67	25	63.5
TC	36	2198	670	9	78.5	65	20	56.2
TR	8	471	216	9	46.2	31	20	33.8
VT	44	2686	189	5	19.6	17	10	12.3
Total	100	6145	670	5	49.9	28	20	50

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

Table 3. Statistics related to the transfer learning experiments using AlexNet for hybrid events (HY), long period (LP), tectonic events (TC), tremors (TR), and volcano-tectonic (VT) classes.

Corpus Experiments	Data Augmentation	Balanced data	Epoch N°	Total data 100% N°	Train data 80% N°	Test data 20% N°	Time to build the model [minutes]
(a)	No	No	10	6145	4916	1229	4.6
(b)	Rotation 5–25%	Yes	20	67095	53676	13419	164.9
(c)	Rotation 5–25%	Yes	–	–	–	6145	0.3
(d)	Rotation 5–25%	Yes	–	–	–	6145	0.3

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

Table 4. Performance of the transfer learning experiments.

Experiments/Metrics	Accuracy	F1	Recall	Precision	Cohen's kappa	Class Accuracy				
						HY	LP	TC	TR	VT
(a)	56.2	57.8	56.2	62.9	65.2	27.8	59.8	89.4	18.9	85
(b)	98.9	98.9	98.9	98.9	98.6	100	99.8	96.1	99.9	98.6
(c)	98.7	97.2	98.7	95.8	96.4	100	100	95.2	99.8	98.4
(d)	96.4	94.6	96.4	92.9	94.0	98.6	96.9	94.5	94.9	97.1

https://doi.org/10.5194/nhess-2022-164
Preprint. Discussion started: 14 June 2022

Natural Hazards and Earth System Sciences
Discussions

Table 5. Performance of the transfer learning experiments applying k-fold validation.

k-fold/Metrics	Accuracy	F1	Recall	Precision	Cohen's kappa	Class Accuracy				
						HY	LP	TC	TR	VT
1	89.2	89.0	89.2	89.2	86.7	99.3	95.2	75.3	84.8	91.3
2	90.1	90.1	90.1	90.4	87.5	90.1	99.6	89.9	84.7	86.6
3	89.3	89.0	89.3	89.1	86.6	89.3	98.3	97.3	71.8	90.7
4	89.8	89.5	89.8	90.1	87.2	89.8	97.3	98.8	67.8	90.6
5	87.9	87.7	87.9	87.8	84.6	87.9	93.1	87.8	77.7	89.1
6	88.6	88.4	88.6	88.7	85.6	88.6	99.6	91.3	81.5	77.9
7	87.6	87.7	87.6	88.1	84.8	87.6	96.6	87.5	82.7	80.6
8	89.8	89.6	89.8	89.7	87.1	89.8	98.6	91.8	75.2	89.9
9	88.7	88.7	88.7	89.0	85.9	88.7	98.9	89.6	78.5	83.5
10	90.4	90.2	90.4	90.2	87.4	90.4	96.0	91.2	79.9	94.9