

The Authors applied a pre-trained convolutional network (AlexNet) to build a multi-station automatic classification system for volcanic seismic signals. They used data that were recorded by a temporal network installed in the Lascar volcano. The Authors designed and performed four experiments using different datasets with real data, synthetic data, and two different combinations of these (combined 1 and combined 2).

The authors obtain fairly good correct classification performances especially in some experiments (those in which they used a data augmentation technique).

I found the article interesting and worth of publishing in the Natural Hazards and Earth System Sciences journal. I have, however, some general and specific comments, which will require some revision of the article.

General comments:

My feeling is that with some changes in the pre-processing phase the performance can further improve. Furthermore, I find the use of spectrograms in image form rather original, however I do not think it is incorrect.

Specific comments:

The abstract must be improved because the reader does not understand which signals the Authors want to classify and what results they obtain in terms of method performance.

Figures

The labels of the figures must be larger (for example, Fig. 1, 2, 6, 7, 8, 9, 10, 11)

In Fig. 4 the letters a, b, c,

On lines 3, 48, 49 etc. "artificial data" should be replaced with "synthetic data".

Lines 69- 71 This paragraph does not seem clear to me. Please rephrase.

Lines 107 -111 In principle, in step 1 you should first remove the average, then filter (a 10th order filter seems a bit strong to me as a filter for a fairly high frequency band such as 1-10 Hz) and finally resample the series, however I think the ObsPy resampling routines apply an appropriate filter before resampling. Furthermore, looking at figure 2 you can see that tectonic events (TC) and tremor have a significant frequency content in the 10-20 Hz band (this is unusual for volcanic tremor which generally has lower frequencies, between 1 and 6 Hz), therefore, in the pre-processing step it would be more appropriate to filter in the 1 - 20 Hz frequency band.

Lines 117 -123 In step 3 to avoid the imbalance produced on the labeled data, the Authors could create a fifth dataset not applying the data augmentation technique, but selecting a subset of the data so that the amount of data that is selected for each class depends on the number of events in the less populated class; in particular the hybrid events HY class (213 events) that can be used as reference (the number of events for the other classes could perhaps be even slightly higher, for example 250 events for the other classes and 213 for hybrid events).

147 see lines 117 -123

Lines 152 -157 In this paragraph the authors should insert values of performance also in the text and not refer only to table 4.

Line 202 "This will improve the understanding and evaluation of the hazards and risks associated with the activity of volcanoes." I believe this last sentence is not necessary.