Unsupervised deep learning on seismic data to detect volcanic unrest

EGU2020-18631
https://doi.org/10.5194/egusphere-egu2020-18631

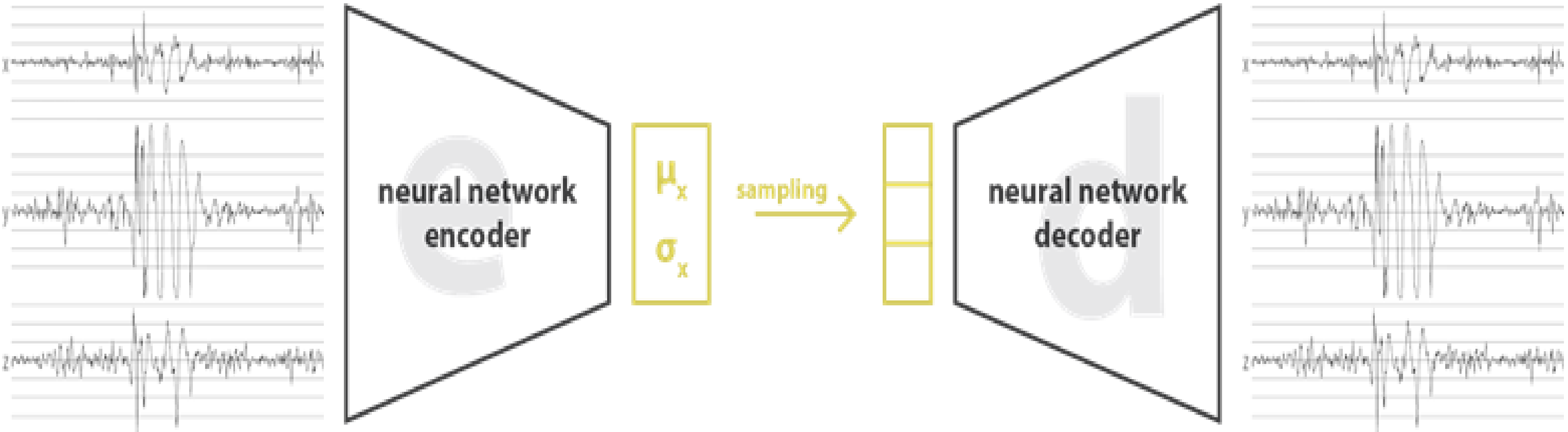
Flavio Cannavo' [1], Andrea Cannata [1,2], Simone Palazzo[3], Concetto Spampinato[3], Demian Faraci[3], Giulia Castagnolo[3], Isaak Kavasidis[3], Chiara Montagna[4], and Simone Colucci[4]

1 Istituto Nazionale di Geofisica e Vulcanologia – Osservatorio Etneo, Catania, Italy
2 Dipartimento di Scienze Biologiche, Geologiche e Ambientali - Sezione di Scienze della Terra, Università degli Studi di Catania, Catania, Italy
3 PeRCeiVe Lab – Department of Electrical, Electronics and Computer Engineering – University of Catania, Catania, Italy
4 Istituto Nazionale di Geofisica e Vulcanologia – sezione di Pisa, Italy

The significant efforts of the last years in new monitoring techniques and networks have led to large datasets and improved our capabilities to measure volcano conditions. Thus nowadays the challenge is to retrieve information from this huge amount of data to significantly improve our capability to automatically recognize signs of potentially hazardous unrest.

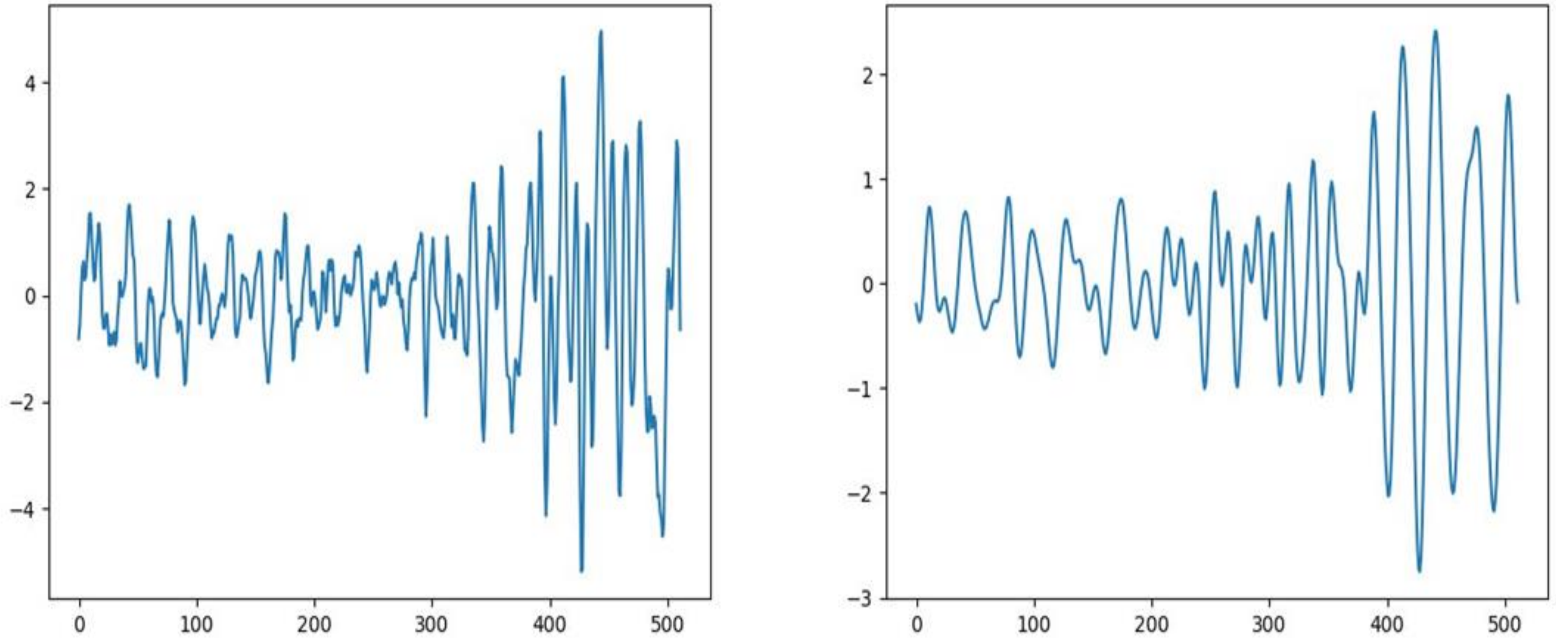
Unrest detection from unlabeled data is a particularly challenging task, since the lack of annotations on the temporal localization of these phenomena makes it impossible to train a machine learning model in a supervised way. The proposed approach, therefore, aims at learning unsupervised low-dimensional representations of the input signal during normal volcanic activity by training a variational autoencoder (VAE) to compress, reconstruct and synthesize input signals. Thanks to the internal structure of the proposed VAE architecture, with 1-dimensional convolutional layers with residual blocks and attention mechanism, the representation learned by the model can be employed to detect deviations from normal volcanic activity.

we developed a tool based on the employment of a *variational auto-encoder*, which consists of an *encoder* network, that extracts the low-dimensional representation of the input signal, and a *decoder* network, which receives such representation and reconstructs the original signal.



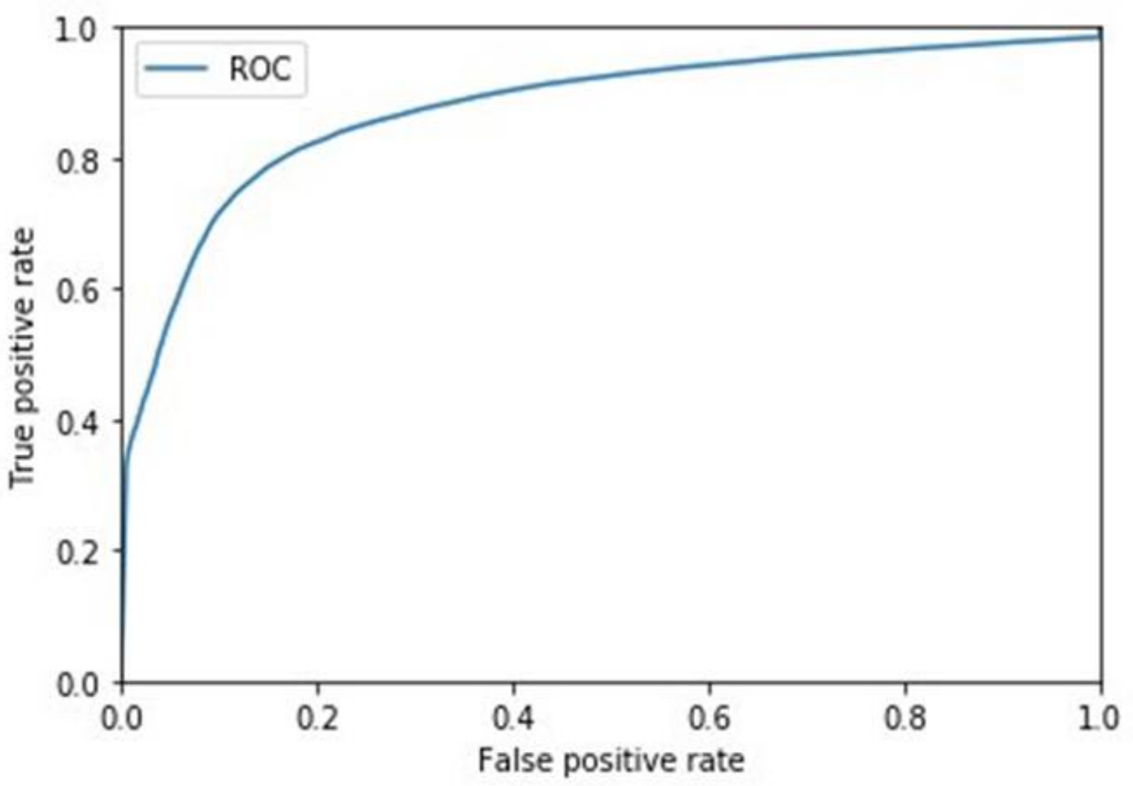
The proposed approach, thus, consists in training a variational auto-encoder on normal activity *only*, with the objective of teaching the model the typical structure and patterns of normal behaviour; then, the model can be employed for *anomaly detection*, i.e. for the identification of paroxystic activity, by thresholding the L_2 distance between the input signal and the reconstructed signal. Indeed, we expect the model to be able to accurately reconstruct normal activity, having been trained on those data, but not paroxystic activity and, in general, any volcanic activity that deviates from normality.

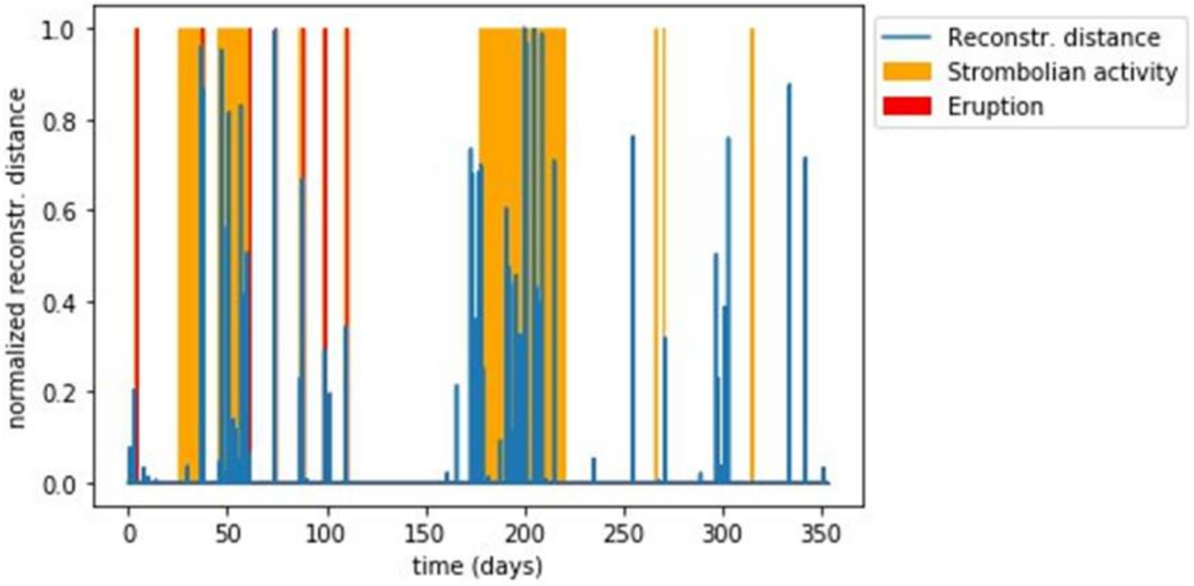
We employ the dataset of seismic waveforms collected at Etna in 2011 and 2012. The resulting model was able to achieve a satisfactory reconstruction accuracy, qualitatively. An example is shown below.



Left: original signal. Right: reconstructed signal. Notice how the reconstruction is able to capture the general trend of the signal, while removing most high-frequency noise.

Quantitatively, we evaluated the accuracy of the mode in distinguishing between normal activity and Strombolian/eruptive activity. Figure shows the corresponding ROC curve, computed for different thresholds of the L_2 distance between each input signal and its reconstruction.





It is possible to notice that increase in the reconstruction distance correlates well with the presence of Strombolian or eruptive activity, as shown below.