

Machine Learning for Automatic Classification of Volcano-Seismic Signatures

Marielle MALFANTE, Mauro DALLA MURA, Jérôme I. MARS
 Univ. Grenoble Alpes, CNRS, GIPSA-Lab,
 38000 Grenoble, France
 Email: marielle.malfante@gipsa-lab.fr

Jean-Philippe MÉTAXIAN
 Univ. Savoie Mont Blanc, ISTerre, CNRS, IRD CNRS,
 73376 Le Bourget du Lac, France.

Abstract—The evaluation and prediction of volcanoes activities and associated risks is still a timely and open issue. The amount of volcano-seismic data acquired by recent monitoring stations is huge (e.g., several years of continuous recordings), thereby making machine learning absolutely necessary for their automatic analysis. The transient nature of the volcano-seismic signatures of interest further enforces the need of automatic detection and classification of such events. In this paper, we present a novel architecture for automatic classification of volcano-seismic events based on a comprehensive signal representation with a large feature set. To the best of our knowledge this is one of the first attempts to automatize the classification task of these signals. The proposed approach relies on supervised machine learning techniques to build a prediction model.

signal processing domains, such as music, environmental noises, bioacoustic, transient or speech representation, classification. (ii) Furthermore, features are extracted from several representations of the signal, such as signal energy, time, frequency, frequency of frequency. (iii) the design of an automatic supervised classification procedure apt to process large datasets of seismic signals.

This paper is organized as follows. Section II recalls related works. Section III presents the proposed architecture. Experimental settings, results and results analysis are exposed in Section IV. Finally conclusions and perspectives are reported in Section V.

I. INTRODUCTION

With the progressive ease to gather data, the scientific community has been increasingly interested into methods and tools for automatic classification. In particular, the so-called Big Data phenomenon has highly impacted monitoring with an ever increasing number of observatories devoted to collect, process and mine recorded signals for extracting relevant information from the environment. This is the case in Geophysics where the number of seismic observatories is increasing due to the greater attention given to the monitoring of seismic events, since they are precursors of volcanic eruptions, tsunamis or earthquakes. The human and economical impact that such events have is dreadful, making monitoring and risks evaluation a societal concern. In the case of volcanoes monitoring, dedicated observatories have started to record, store and transmit continuous signals for only a few years. Tools to automatically process existing databases and analyze real-time recordings are still rare and not fully operational.

The aim of this paper is to present an efficient and new architecture for the automatic classification of seismic events produced by a volcanic activity. The main contributions of this work are: (i) the evaluation of a large feature set used to describe the signals of interest: considered features have been gathered from an extensive state of the art in various

II. RELATED WORKS

To the best of our knowledge, there are no established procedures for the classification of volcano-seismic data. However, several attempts of automatic classification have been proposed in the literature addressing acoustic signals in other scientific domains (e.g., speech processing). The typical architecture relies on performing first a feature extraction step, which provides a description of the acquired signals and then a learning phase that is the core of the automatic classification procedure [1].

This work mainly focuses on the first step. Specifically, we consider for the classification of volcano-seismic data, an extensive set of features proposed in several fields such as seismic [2], acoustics (environmental, bio-acoustics, animal [3], [4], [5], [6] or anthropic [7] ambient and/or landscape noises), speech and speech analysis [8], [9], [10] and music signals [11]. More in details, we take into account features such as those proposed in the recent work [2] where seismic waves are represented by few classically features such as duration time, statistical descriptors (skewness, kurtosis, statistics ratios) and fundamental frequencies. In classification of transient sonar sounds, the work in [7] considers more than 20 features. Those features are mainly composed of descriptors of the signal shape (rate of attack and decay, time of the main peak), statistical

moments (mean skewness and kurtosis) and of signal power (peak power, average and power standard deviation). In bioacoustic, similar features have been also used in [3] and in [4] respectively to distinguish boats from whales and for bird species identification. Descriptors based on entropy are used in [5] for frog sounds classification. Finally, similar features are also used in music classification. We can mention [6] in which music genres are discriminated using entropy, centroid, centroid ratio, bandwidth or silence ratio or [11] who recognize orchestral instruments with centroid, skewness, kurtosis or centroid velocity. We collect and adapt a large part of these features to our dataset.

As already mentioned, the automatic classification of seismo-volcanic waves is still an almost unexplored field. This is in part due to the fact that seismo-volcanic signals are in some aspects similar to speech signals, but at the same time incompatible with speech dedicated features since their frequency and distribution range is very different. Seismo-volcanic signals are also difficult to classify due to their extreme variability (for example, in the time duration or bandwidth of the events, and for the similarities between signals belonging to different semantic classes). However, the nature of these data makes reasonable to consider jointly several features defined for other acoustic or transient signals, since only features defined on a specific domain (e.g., speech or animal sounds) are not expected to be adapted to describe seismo-volcanic signals. In the following we present the features that were considered in this analysis as well as the proposed architecture for automatic classification.

III. PROPOSED ARCHITECTURE FOR AUTOMATIC CLASSIFICATION

A. Supervised Machine Learning

We briefly recall the background on supervised machine learning [12] in order to introduce the proposed technique. Machine Learning is a field of Artificial Intelligence that aims to build a model that separates data into various groups, referred as classes. Those methods are relatively new but currently widely spread [13], [14] and used in a wide range of applications including medical imaging [15] or image processing [16], [13] or hyperspectral imaging. Supervised Machine Learning refers to a set of methods that build a prediction model from a labeled dataset: set of observations whose thematic class is known (usually after a manual labeling). From those set of observation, a prediction model is built, which can then be applied to predict the class of series of unlabeled samples. In the literature, different algorithms can be used to build a supervised prediction model such as Random Forest (RF) [17], Support Vector Machine (SVM) [18] or Neural Network (NN) [19].

B. Proposed Classification Architecture

Here, we present the different steps of the proposed architecture.

1) *Preprocessing*: As we deal with potentially large datasets (e.g., the one we consider in this work collects acquisitions over 6 years) and a large number of classes, we propose at first to perform a preprocessing of the continuous recordings to obtain a formatted dataset of observations. We need to ensure that the signals that we consider belonging to each class have roughly an uniform duration (i.e., number of samples), see Table II for details. This dataset will be used for training and testing our models.

The patterns of the classification problem are obtained by manually segmenting the dataset into smaller chunks (of potentially different duration). Each extracted segment is then assigned to a specific class which corresponds to the nature of the underlying physical event (reference class). We propose also a preprocessing to carry out a normalization (of the observations) in time and in frequency, in order to be able to perform an analysis that is invariant to the energy of the signals. We perform this operation, because seismo-volcanic signals are widely variable over the 6 years of recording. Signals belonging to the same class might be characterized by both a high and low energy.

2) *Features Extraction*: Features extraction leads to a representation of each observation in a corresponding feature space in which an automatic decision rule can be established. More specifically, descriptors are computed from the observations in order to obtain a representation of the acquired signals that is more compact (reducing the dataset dimension is compulsory and in general meaningful) with respect to considering all the samples in each observation as features. This can be justified by several arguments. First, the representation used in input of the learning algorithm should make data as separable as possible. Using a feature space that are discriminative leads to a separates data into their classes support this assumption. Secondly, learning algorithms are affected by the curse of dimensionality [20] which states that the number N of needed examples to train a prediction model increases with d , the dimension of the data. Representing the data in a feature space of reduced dimension with respect to the original dimension allows training models to use a reduced number of observations.

Finally, learning algorithms such as SVM or RF do not capture ordered representation: information such as global trend and general shape of a signal are lost. Therefore, using waveform or spectra in input of a learning algorithm is not advised.

Using features that explicitly seize this information generally produced better results. For this application, we propose to extract features from several representations of the observations. Especially, from an observation $x(t)$, we extract features from three different domains which are as follows:

- the time domain, in which features are extracted from the signal itself $x(t)$ or from its energy $E(x(t))$ and they are able to describe the waveform and its shape.
- the frequency domain $X(f)$ to describe the spectral

content of the observation. The Welch spectral density estimator is used to compute $X(f)$, since it allows to decrease the signal dimension and also reduces the noise and smooths the signal.

$$X(f) = \text{Welch}(x(t)) \quad (1)$$

- the frequency of frequency domain from $\mathcal{X}(q)$. Representing an observation from the Fourier transform of $X(f)$ can underline the periodicity that can be found in the spectra, and so the harmonicity of the original observation. This space of visualization is mostly used for speech processing and is referred as the *cepstral* space.

$$\mathcal{X}(q) = \mathcal{TF}\{X(f)\} \quad (2)$$

Extracting the same features from various representations of a same observation allows to underline some specificities of the signal. Using this strategy can greatly improve results and is coherent with the manual procedure that experts use conventionally. As we said before, the features extracted from the various representations are combined in order to describe an observation in different domains. We considered several features that were defined in the state of the art in different fields (see Table I for a list). Those features belong to different types. Some of them are Shannon and Rényi entropies, shape descriptors such as rate of attack or decay, bandwidth measurements, maximum amplitude or ratios between mean and minimum values for example. Various statistical moments including standard deviation, skewness and kurtosis are also used, respectively describing spreading, asymmetry and flatness of an observation. By using all these features, we notice that each observation can be eventually represented by the feature vector of dimension $D = 79$ (detailed in Table I).

3) *Learning*: Once features have been extracted, labeled feature vectors are given as input to a learning algorithm that will produce a prediction model to be used on the unlabeled dataset. In this work, we used Support Vector Machine (SVM) for its robustness [8]¹. From the labeled dataset of N signals, a model is then learned on $\alpha \cdot N$ examples with $0 < \alpha < 1$ and test it with the remaining $(1 - \alpha) \cdot N$ examples. Classically, this process is repeated 50 times with different random trials of examples to ensure the results stability (this validation process is known as cross validation).

IV. EXPERIMENTAL RESULTS

A. Experimental material

The dataset used in this work was collected on Ubinas, a stratovolcano in southern Péru (16 22' S, 70 54' W), which has an altitude of 5676m and it is close to the city of San Juan de Tarucani. Ubinas is considered to be the Péru most active and dangerous volcano and is therefore

¹Similar results not presented here are obtained if other machine learning algorithms are used (like Random Forest or Neural Networks).

TABLE I

FEATURE SET FOR A GENERIC NUMERICAL SIGNAL $s[i]_{i=0}^n$ COMPOSED OF n DISCRETE SAMPLES (OBSERVATIONS MIGHT NOT HAVE THE SAME DURATION) AND WHICH CAN BE DEFINED IN DIFFERENT SPACE AS THE TIME $x(t)$, ENERGY $E(x(t))$, FREQUENCY $X(f)$ OR CEPSTRAL $\mathcal{X}(q)$.

Feature	Definition	Reference
Length	$n = \text{card}(s)$	[7]
Mean	$\mu_s = \frac{1}{n} \sum_i s[i]$	[7]
Standard deviation	$\sigma_s = \sqrt{\frac{1}{n-1} \sum_i (s[i] - \mu_s)^2}$	
Skewness	$\frac{1}{n} \cdot \sum_i \left(\frac{s[i] - \mu_s}{\sigma_s} \right)^3$	[2]
Kurtosis	$\frac{1}{n} \cdot \sum_i \left(\frac{s[i] - \mu_s}{\sigma_s} \right)^4$	[2]
Average	$\frac{\sum s(u)}{n}$	[7]
Central Energy	$\bar{i} = \frac{1}{E} \sum_i E_i i$	[7]
RMS bandwidth	$RM S_i = \sqrt{\frac{1}{E} \sum_i i^2 E_i - \bar{i}^2}$	[7]
Mean skewness	$\sqrt{\frac{\sum_i (i - \bar{i})^3 E_i}{E \cdot RM S_i^3}}$	[7]
Mean kurtosis	$\sqrt{\frac{\sum_i (i - \bar{i})^4 E_i}{E \cdot RM S_i^4}}$	[7]
Shannon entropy	$-\sum_j p(s_j) \log_2(p(s_j))$	[6], [5]
Rényi 'entropy'	$\frac{1}{1-\alpha} \log_2 \left(\sum_j p(s_j)^\alpha \right)$	[5]
Rate of attack	$\max_i \left(\frac{s[i] - s[i-1]}{n} \right)$	[7]
Rate of decay	$\min_i \left(\frac{s[i] - s[i+1]}{n} \right)$	[7]
Specific values	Ratios, min, max, mean, etc.	[7], [2]

closely monitored especially by the Instituto Geofísico del Péru. In this study, we used data recorded over 6 years (between 2000 and 2006) by a 1 Hz vertical sensor at station UB1 with a sampling rate of 100Hz and transmitted in real time. Our study is based on 6 different signal classes defined by the geophysicists (they are reported in Table II). For each class, a maximum of 800 examples per class is considered for the learning phase. If more than 800 examples are available, the first 800 ones have been selected in a chronological order. It is also worth noticing that the time window in which those signals are recorded can vary significantly, aspect that should be addressed by a preprocessing of the data. Since the duration of the recorded phenomena of interest can be extremely variable, we have chosen to consider a maximum observation window of 40 seconds. In the experiments we considered 3125 labeled signals.

B. Experimental settings

This section reports the different experimental settings. Considering different window lengths for the Welch estimator or using a logarithmic scale do not have a significant influence on the results. Once computed, the dynamic of the feature vectors has been normalized (in the interval $[0, 1]$) before being considered as input to a Support Vector Machine, with a Gaussian Radial Basis Function kernel of cost $C = 10$ and Gaussian parameter $\gamma = 0.01$. Those values were obtained to optimize results

TABLE II

UBINAS DATASET. FOR EACH CLASS, CLASS NAME (REFERENCE), A DESCRIPTION, NUMBER OF OCCURRENCES N_i , WINDOW LENGTH $\Delta_{t,i}$ NEEDED TO GATHER THE N_i SIGNALS AND MEAN LENGTH ARE GIVEN.

Ref.	Description	N_i	$\Delta_{t,i}$ (in days)	Mean length
LP	Long period	800	201	40s
TR	Tremor	800	8	27min48s
EXP	Explosion	154	1396	51s
VT	Volcano-tectonic	800	1958	24s
HIB	Hybrid	466	1647	34s
TOR	Tornillo	105	632	40s

TABLE III

INFLUENCE OF FEATURE SET ON ACCURACY RESULTS ($\alpha = 70\%$)

Features	All features	Time features	Freq. features
Dimension	79	40	39
Accuracy results	$90.4 \pm 0.7\%$	$86.7 \pm 1.0\%$	$73.0 \pm 1.3\%$

on a subset of data as a preliminary study, and are kept constant in all experiments. Results are obtained through a 50-fold cross validation. A preliminary study showed that results are steady if using more trials. Finally, the learning rate α varied during the experiments so as to estimate the variability in each class and the necessary number of examples for the learning stage.

C. Results

In this section, we present results associated to the automatic classification of Ubinas volcano-seismic data over six classes. First, results validating the methodology and features used are presented in Table III. Especially, comparison from features computed from the time representation $x(t)$ (including its energy) and the frequency representation $X(f)$ (including on its Fourier Transform $\mathcal{X}(q)$) used to represent the signal are given. A confusion matrix is also given in Table IV. Secondly, results displaying influence of the learning rate α are shown in Figure 1.

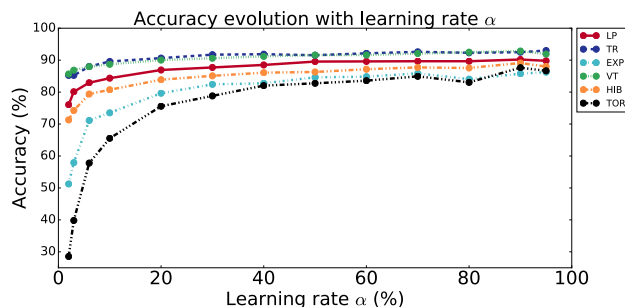


Fig. 1. Accuracy evolution depending on the learning rate α for the 6 classes

TABLE IV

CONFUSION MATRIX BETWEEN TRUE CLASS AND PREDICTED CLASS WITH ALL FEATURES FOR THE 6 CLASSES, LEARNING RATE $\alpha = 70\%$. MEAN ACCURACY FOR EACH CLASSES IS GIVEN. MEAN ACCURACY OVER ALL CLASSES = 90.4%

		True class					
		LP	TR	EXP	VT	HIB	TOR
Predicted Class	LP	215	16	0	2	7	1
	TR	20	222	0	1	2	0
	EXP	0	0	40	6	1	0
	VT	2	0	5	221	8	3
	HIB	2	1	1	10	123	0
	TOR	0	1	0	1	0	27
Mean acc.		89.7%	92.6%	85.9%	92.1%	87.7%	84.9%

D. Results Analysis & Discussion

Table III shows that the proposed architecture reaches an accuracy of 90%, thereby validating the methodology and the features used in the analysis. It is worth noticing the influence of features extracted from several representation of the observations: accuracy does not exceed 73% and 87% if using features computed respectively from the $X(f)$ and $x(t)$. The discriminating information is best underlined when using several representations, each one underlying some properties of a signals. The combined use of temporal, spectral and cepstral representations allows more discriminative feature vectors. Table IV presents the confusion matrix when all features are used to represent the data and $\alpha = 70\%$. The corresponding confusion matrix displays class by class the number or proportion of true predictions, but also the repartition of wrong predictions. It is useful to analyze errors made by a model. In this case, it is worth noticing that errors made by our model come from two different sources. First, tremors and long periods can be mistaken for each other, which is explained by the similarity between the production mechanisms of those signals. Both are associated to the resonance of cavities, filled with magma, gas or water. The similarity of the production mechanisms induces a similarities in the observed signals, hence this first mistakes. The second source of mistake is seen in hybrids being mistaken for long periods or volcano tectonics. Once again this is explains by the physics behind the signals. Hybrids are a class of signals processing both characteristics of long period and volcano tectonics, thereby explaining the possible confusion. Finally, the less frequent error is made between volcano tectonics and explosions, which can occur in the same frequency bandwidth with similar shapes.

The evolution of performances depending on the learning rate α also worths to be considered. In general, classification accuracy increases for increasing α before reaching a plateau. This stability however is reached for different learning rates depending on the considered class. Respectively, tornillo, explosions and hybrids need larger learning rate than long period, tremors and volcano tectonics. Corre-

sponding learning rates are actually ordered depending on N_i , the number of examples by class. Eventually, this plot shows that with a stability in results is reached with about a hundred examples per class.

V. CONCLUSION & PERSPECTIVES

As a conclusion, we would like to underline the need of automatic processing tools for the monitoring of active volcanoes. We propose in this paper to use a supervised machine learning architecture to build a prediction model. The architecture has been tested on 3125 signals of Ubina volcano, split into six different classes. The accuracy reaches 90%, thereby validating (i) the methodology, (ii) the feature extraction from several observation representations and (iii) the extended feature set. It is our belief that this feature set can be used for other and more general applications. As future developments of this work, we are currently working on (i) testing the generalization capacities for the feature set, and (ii) developing prediction models based on a semi-supervised approach. The idea is here to reduce the dependency on a large set of labeled examples, which is currently the main limitation of the proposed model.

ACKNOWLEDGMENT

This work has been supported by a grant from Labex OSUG@2020 (Investissements d'avenir - ANR10 LABX56) and DGA/MRIS. GIPSA-Lab SIGMAPHY is part of Labex OSUG@2020 (ANR10 LABX56). The authors would also like to thank Dr. Orlando Macedo from Instituto Geofísico del Perú for his input in this work.

REFERENCES

- [1] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. John Wiley. Sons Eds, 2001.
- [2] N. Langet, "Détection et caractérisation massives de phénomènes sismologiques pour la surveillance d'événements traditionnels et la recherche systématique de phénomènes rares," Ph.D. dissertation, University of Strasbourg, 2014.
- [3] S. Zaugg, M. Van Der Schaar, L. Houégnigan, C. Gervaise, and M. André, "Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories," *Applied Acoustics*, vol. 71, no. 11, pp. 1011–1019, 2010.
- [4] S. Fagerlund, "Bird Species Recognition Using Support Vector Machines," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, 2007.
- [5] N. C. Han, S. V. Muniandy, and J. Dayou, "Acoustic classification of Australian anurans based on hybrid spectral-entropy approach," *Applied Acoustics*, vol. 72, no. 9, pp. 639–645, 2011.
- [6] S. Esmaili, S. Krishnan, and K. Raahemifar, "Content based audio classification and retrieval using joint time-frequency analysis," in *ICASSP-2004, IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, 2004, pp. 9–12.
- [7] S. Tucker and G. J. Brown, "Classification of transient sonar sounds using perceptually motivated features," *IEEE Journal of Oceanic Engineering*, vol. 30, no. 3, pp. 588–600, 2005.
- [8] S. Gupta, A. D. Dileep, and I. Veena Thenkanidiyoor, "Segment-Level Pyramid Match Kernels For The Classification of Varying Length Patterns of Speech Using SVMs," in *24th European Signal Processing Conference (EUSIPCO'16)*, 2016, pp. 2030–2034.
- [9] K. Mahkonen, A. Hurmalainen, T. Virtanen, and K. JK, "Cascade Processing for Speeding up Sliding Window Sparse Classification," in *24th European Signal Processing Conference (EUSIPCO'16)*, 2016, pp. 2305–2309.
- [10] P. Sharma, V. Abrol, A. Sachdev, A. D. Dileep, A. K. Sao, and I. I. T. Mandi, "Speech emotion recognition using kernel sparse representation based classifier," in *24th European Signal Processing Conference (EUSIPCO'16)*, 2016, pp. 374–377.
- [11] I. Fujinaga and K. MacMillan, "Realtime recognition of orchestral instruments," in *Proceedings of the International Computer Music Conference (ICMC2000)*, vol. 141, 2000, pp. 141–143.
- [12] T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning," *Elements*, vol. 1, pp. 337–387, 2009.
- [13] A. Iosifidis and M. Gabbouj, "Graph-regularized Multi-class Support Vector Machines for Face and Action Recognition," 2016, pp. 96–100.
- [14] O. F. Kilic, N. D. Vanli, H. Ozkan, I. Delibalta, and S. S. Kozat, "Adaptive Hierarchical Space Partitioning for Online Classification," in *24th European Signal Processing Conference (EUSIPCO'16)*, 2016, pp. 2290–2294.
- [15] M. Rahim, "Impact of perceptual learning on resting-state fMRI connectivity : A supervised classification study," in *24th European Signal Processing Conference (EUSIPCO'16)*, 2016, pp. 250–254.
- [16] R. Xie, H. Huttunen, S. Lin, S. S. Bhattacharyya, and J. Takala, "Resource-Constrained Implementation and Optimization of a Deep Neural Network for Vehicle Classification," in *24th European Signal Processing Conference (EUSIPCO'16)*, 2016.
- [17] L. Breiman, "Random forests," *Machine learning*, pp. 5–32, 2001.
- [18] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers," in *Proceedings of the fifth Annual Workshop on Computational Learning*, 1992, pp. 144–152.
- [19] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.
- [20] R. Bellman, "Dynamic Programming and Lagrange Multipliers." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 42, no. 10, pp. 767–769, 1956.