IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE GEOSCIENCE AND REMOTE SENSING SOCIETY SECTION

Received 10 December 2023, accepted 21 December 2023, date of publication 15 January 2024,
date of current version 12 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3354717

RESEARCH ARTICLE

Volcanic Micro-Earthquake Classification With Spectral Manifolds in Low-Dimensional Latent Spaces

CARLOS-PAUL BERNAL-OÑATE[1,2], (Member, IEEE), ENRIQUE V. CARRERA[1], (Senior Member, IEEE), FRANCISCO-MANUEL MELGAREJO-MESEGUER[2], RODOLFO GORDILLO-ORQUERA[1], JOSÉ LUIS ROJO-ÁLVAREZ[2,3], (Senior Member, IEEE), AND ROMÁN A. LARA-CUEVA[1], (Senior Member, IEEE)

[1]Departamento de Eléctrica, Electrónica y Telecomunicaciones, Universidad de las Fuerzas Armadas—ESPE, Sangolquí 171103, Ecuador
[2]Department of Signal Theory and Communications, Telematics, and Computing Systems, Universidad Rey Juan Carlos, 28933 Madrid, Spain
[3]D!lemmaLab Ldt., Fuenlabrada, 28943 Madrid, Spain

Corresponding author: Carlos-Paul Bernal-Oñate (cpbernal@espe.edu.ec)

This work was partly supported by the Ministry of Economy and Competitiveness, research grants IPT-2012-1126-300000, PID2019-106623RB/AEI/10.13039/5011000110033, PID2019-104356RB-C42/AEI/10.13039/5011000110033, PID2022-140786NB-C31/AEI/10.13039/5011000110033, 2022-REGING-95982, and 2022-REGING-92049, and partly by the Universidad de las Fuerzas Armadas-ESPE under research grants 2013-PIT-014 and 2016-EXT-038. Research funded by the Autonomous Community of Madrid (ELLIS Madrid Unit).

ABSTRACT Micro-earthquakes are frequently associated with volcanic activity and are vital indicators of volcanic processes. These minor seismic events occur within or near volcanic systems, yielding valuable insights into subsurface activities. Geologists meticulously record and analyze these events to monitor volcanoes and forecast eruptions. While recent years have seen several studies proposing automated detection and classification systems of seismic events, approaches based on Manifold Learning techniques could be beneficial in terms of information interpretability and transfer learning to other Machine Learning tasks. This paper presents a novel approach employing audio features and psychoacoustic scales to represent micro-earthquakes at Cotopaxi and Llaima Volcanoes, and these representations are then transformed into low-dimensional latent spaces. We implemented a multi-class classification system for events generated by these volcanoes, incorporating feature selection techniques based on audio-inspired features. This approach enhances the detection of volcanic phenomena triggering eruptions and improves interpretability. Our results indicated high accuracy, with rates of 94.44% for Llaima Volcano and 95.45% for Cotopaxi Volcano when utilizing mutual information to select the most relevant features. Spectral Roll-off Point and Spectral Flux dominate in classifying events for both volcanoes. These findings suggest that low-dimensional latent spaces, particularly when utilizing spectral features, can be a promising foundation for developing transfer learning schemes in general, and new multi-class classification systems in particular, for detecting volcanic micro-earthquakes.

INDEX TERMS Audio features, latent space separability, manifold learning, micro-earthquake classification, mutual information.

I. INTRODUCTION

Around the world, institutions in charge of monitoring active volcanoes, such as Cotopaxi in Ecuador and Llaima in Chile, have installed several seismic monitoring stations equipped with high-precision instruments capable of recording seismic signals that travel through the ground. The purpose of these seismic networks is to continuously monitor volcanic activity to determine possible upheaval behavior that may force authorities to generate early warning strategies for the

The associate editor coordinating the review of this manuscript and approving it for publication was Gerardo Di Martino.

population at risk of an imminent eruption [1], [2]. In this way, those monitoring networks are intended to minimize the catastrophic consequences that could occur due to a volcanic eruption.

Numerous studies have been conducted over the years in seismic event detection and classification. These studies vary primarily in terms of methodology and tools employed to identify the type of seismic event. For instance, in [3], a classifier was developed by extracting 79 features related to time, frequency, and scale domains. The classifier considered metrics such as the Gini index, the drag rule, and the standard deviation. This system achieved an approximate classification accuracy of 95% for Long-Period (LP), Volcano-Tectonic (VT), and Hybrid (HB) events from Cotopaxi volcano (Ecuador). Similarly, in [4], a study of signals from Nevado del Ruiz (Colombia) demonstrated distinctions among various seismic events based on a stochastic-type measure applied to a wide range of time-varying features. The classification accuracy for LP, VT, HB, and Tremors (TR) events achieved by this system was approximately 88%. On the other hand, in [5], the authors utilized a deep-learning framework to detect and classify LP and VT events from the Cotopaxi volcano, achieving a classification accuracy of 97%. Another noteworthy example can be found in [6], where the authors compared several neural networks and deep-learning architectures to classify LP, VT, HB, TR, and Tectonic (TEC) events. Their system achieved an impressive classification accuracy of 98%. Another approach for the analysis of seismic signals from volcanoes is the proposal of [7] that uses the databases of the Popocatépetl and Volcán de Fuego in Colima (Mexico), and through the parameterization of seismic data using Mel Frequency Cepstral Coefficients as features used in Hidden Markov Models to classify the 7 events contained in the databases obtaining an accuracy over 78%. Following this line of research in [8] proposes using Linear Prediction Coefficients (LPC) and statistical properties as features to feed deep neural networks obtaining varied results for the 7 types of events of the Volcán de Fuego. In [9] they propose the use of Bayesian Neural Networks together with the extraction of features proposed in [8] to improve the Volcán de Fuego classifier, for which it uses the Volcán de Fuego features for the training and testing set and for the blind tests with Mount St. Helens (USA) and Peteroa (Chile) volcanoes. Finally, in [10] the use of recurrent neural networks to detect and classify the events of the Deception Island Volcano in Antarctica using LPC and Logarithmic Filter Banks (LFB), this database contains only 5 types of events and its results are over 90%. These studies collectively illustrate the diversity of approaches and the effectiveness of various methodologies in classifying seismic events, providing valuable insights into the field of vulcanology.

Despite this variety of existing works, and to the best of our knowledge, a few efforts have been made to scrutinize the possibilities of Manifold Learning (MnL) techniques in volcanic event classification. MnL is a set of techniques for finding low-dimensional representations of high-dimensional data by assuming that said data have some intrinsic, low-dimensional, geometrical structure. They were extremely popular in machine learning in the 2000s, mostly for data visualization. MnL seeks to find a method to generalize to all data structures by applying differential geometry to machine learning. The fundamental assumptions [11] are: (1) There exist nonlinear relationships in the data that can be modeled with manifolds, this is, surfaces that are smooth, not too complex, and continuous, and span multiple dimensions; (2) It is not important to maintain the shape of the original data, but; (3) If each data point preserves the distance with the close ones, the geometric relationships can be maintained. Clustering, dimensionality reduction, and many semi-supervised and supervised machine learning algorithms can be used with great advantage in applications whose data are well represented by manifolds [12]. Several methods have been proposed for MnL, which initially focused on visualization, but their practical relevance was then reinforced. Latent-variable models aim to explain the surface structure of the data with some underlying hidden variables [13], and their advantages include dimensionality reduction (and subsequently, increased robustness to overfitting), generalization improvement, and especially, suitability for transfer learning for other machine-learning related tasks, as well as interpretability and visualization in their analysis. However, it is important to evaluate empirically how these techniques can affect the performance of a specific data model in a given practical problem. Additionally, these models provide interpretability and visualization capabilities crucial for the in-depth analysis of volcanic datasets. However, it is paramount to empirically evaluate how these MnL techniques can impact the performance of specific data models in the context of volcanic event detection. By bridging the gap between traditional methodologies and the innovative applications offered by MnL, our study aims to shed light on the potential of these techniques for advancing the field of vulcanology.

In this work, we characterize and classify the different types of micro-earthquakes generated by the eruptive process of volcanoes, for which we use the databases of Cotopaxi and Llaima Volcanoes, which as a common denominator are strato volcanoes. In contrast to previous work in which frequency, voice, LFB, statistical properties or time characteristics were extracted to feed a machine-learning classifier, we start from the assumption that using audio and voice analysis with psychoacoustic scales can better represent the evolution in frequency and time of micro-earthquakes. In this sense, a total of 14 features are extracted with linear and nonlinear scales divided into 32, 64, 128, and 256 bands. The use of frequency features and psychoacoustic scales are associated with the term instantaneous feature, short-term feature, low-level feature, or audio descriptor is generally used for measurements that generate one value per (short)

block of audio samples. An instantaneous feature is not necessarily significant from a musical, musicological, or perceptual point of view, which would be the case for a signal of volcanic origin, so it is often referred to as a low-level feature. A low-level feature can be a basis for constructing higher-level features describing more meaningful properties of the (musical) signal [14]. In this context, we use several of these features over time to increase the amount of information delivered to the classifier through a windowing process that somehow preserves certain features in time, helping with the classification process. With all these features, manifold algorithms visualize low-dimensional latent spaces where the relationships between micro-earthquake classes for different volcanoes can be observed and analyzed. Autoencoders are also used to generate latent spaces, and classifiers of several micro-earthquakes are built by supervised learning to obtain objective metrics such as accuracy, precision, and sensitivity. From the obtained results, the features, scales, and bands that better represent the evolution in time and frequency of the different types of micro-earthquakes are chosen. For a systematic selection of features, a technique based on mutual information (MI) is also used to determine the relevance of each feature in the performance of the classifier and its contribution to the dynamics of the eruptive process of the volcanoes [15]. Based on all these analyses, we were able to establish similarities between the micro-earthquake events at Cotopaxi and Llaima Volcanoes.

The remainder of the manuscript is organized as follows. Section II reviews the main concepts of manifold algorithms, psychoacoustic audio scales, and feature selection methods. Section III describes the databases and the features used in this work. Section IV presents the main results obtained using supervised learning autoencoders, together with the analysis of the most relevant features. Finally, Section V concludes this work and foresees relevant future lines of research.

II. BACKGROUND

A. MANIFOLD LEARNING

High-dimensional datasets are difficult to analyze due to different issues known as curse of dimensionality [16]. Therefore, dimensionality reduction techniques are used to transform data to lower-dimensional subspaces for better representation [17]. Usually, representing high-dimensional structured data within a low-dimensional latent space allows us to find a compressed state containing a small set of essential and informative variables. Traditional and machine-learning manifold algorithms are presented below.

1) TRADITIONAL MANIFOLD ALGORITHMS

Many linear dimensionality reduction techniques, such as principal component analysis (PCA), have been developed to simplify the representation of high-dimensional datasets. These algorithms choose a linear projection of the data according to specific rules. Although these techniques are powerful, they often overlook essential nonlinear structures in the data. On the other hand, t-distributed stochastic neighbor embedding (t-SNE) converts affinities of data points into probabilities [18]. Joint Gaussian probabilities represent affinities in the original space, while students' t distributions represent affinities in the embedded space. This feature makes t-SNE quite sensitive to local structures, revealing structures at many scales.

Unlike t-SNE, uniform manifold approximation and projection (UMAP) is a general nonlinear dimensional reduction technique. This algorithm is based on three assumptions about the data [19]: *(i)* data are uniformly distributed on a Riemannian manifold, *(ii)* the Riemannian metric is locally constant (or it can be approximated as such), and *(iii)* the manifold is locally connected.

2) MACHINE-LEARNING MANIFOLD ALGORITHMS

Compact representations of data are essential to machine and deep learning since lower dimensional data representations permit automatic feature extraction [20]. Thus, manifold learning is a machine-learning scheme that enables the map of observed high-dimensional data into a low-dimensional embedded space. Despite the existence of specific techniques to carry out MnL in machine learning, autoencoders have emerged as flexible and valuable manifold learners. An autoencoder is an artificial neural network based on unsupervised learning with a particular architecture to learn a lower-dimensional encoding from higher-dimensional input data. In this way, the neural network architecture can be trained to capture the intrinsic structure of the input data within each encoding and use this information to perform quantitative and qualitative analyses on the learned associations.

The autoencoder architecture consists of an encoder and a decoder. The encoder compresses the input data $\boldsymbol{x}_i$ into an encoding symbol $\boldsymbol{h}_i$, and the decoder decompresses the encoding $\boldsymbol{h}_i$, reconstructing the input data as an approximated output (or estimation) $\hat{\boldsymbol{x}}_i$ [20]. The encoding $\boldsymbol{h}_i$ represents the autoencoder input data in the latent space (or bottleneck).

Encoder and decoder architectures can be extended to include multiple hidden layers, providing increased representational capacity. These layers are trained similarly to other artificial neural networks, with the objective of minimizing the difference between the input $\boldsymbol{x}_i$ and the output $\hat{\boldsymbol{x}}_i$ of the network [21]. Mathematically, this can be expressed as follows for an autoencoder with L multiple hidden layers:

$$\boldsymbol{h}_i^{(1)} = \phi^{(1)}(\boldsymbol{W}_e^{(1)}\,\boldsymbol{x}_i + \boldsymbol{b}_e^{(1)}) \tag{1}$$

$$\boldsymbol{h}_i^{(2)} = \phi^{(2)}(\boldsymbol{W}_e^{(2)}\,\boldsymbol{h}_i^{(1)} + \boldsymbol{b}_e^{(2)}) \tag{2}$$

$$\cdots \tag{3}$$

$$\boldsymbol{h}_i^{(L)} = \phi^{(L)}(\boldsymbol{W}_e^{(L)}\,\boldsymbol{h}_i^{(L-1)} + \boldsymbol{b}_e^{(L)}) \tag{4}$$

$$\hat{\boldsymbol{x}}_i = \varphi(\boldsymbol{W}_d\,\boldsymbol{h}_i^{(L)} + \boldsymbol{b}_d) \tag{5}$$

where $\boldsymbol{h}_i^{(l)}$ denotes the encoding at the l-th layer in the latent space corresponding to the input $\boldsymbol{x}_i$, $\phi^{(l)}(\cdot)$ and $\varphi(\cdot)$ are nonlinear transformations, $\boldsymbol{W}_e^{(l)}$ is the weight matrix of the

TABLE 1. Selected audio-based features to be extracted from windowed volcanic signals.

No.	Audio Feature
1	Spectral Centroid
2	Spectral Crest
3	Spectral Decrease
4	Spectral Entropy
5	Spectral Flatness
6	Spectral Flux
7	Harmonic Ratio
8	Spectral Kurtosis
9	Spectral Rolloff Point
10	Spectral Skewness
11	Spectral Slope
12	Spectral Spread
13	Short-time Energy
14	Zero-crossing Rate

encoder at layer l, $\boldsymbol{b}_e^{(l)}$ is the bias vector of the encoder at layer l, $\boldsymbol{W}_d$ is the decoder weight matrix, and $\boldsymbol{b}_d$ is the decoder bias vector. During training, weights and biases (i.e., $\boldsymbol{W}_e^{(l)}$, $\boldsymbol{W}_d$, $\boldsymbol{b}_e^{(l)}$, and $\boldsymbol{b}_d$) are adjusted using a set of samples such that $\hat{\boldsymbol{x}}_i \approx \boldsymbol{x}_i$. Common loss functions include mean squared error and cross-entropy.

B. AUDIO FEATURES, SCALES, AND SELECTION

Psychoacoustic scales have been used to extract features for voice and audio analysis and for characterizing, analyzing, and modeling many other types of signals [22], [23], [24]. The psychoacoustic scales aim to extract information from audio signals regarding the perceptual characteristics of the environment and the listener [25]. Considering the physical nature of audio signals, the best results are usually obtained by analyzing them in the frequency domain and considering that the response of the auditory nerve is logarithmic. This has led to the development of a series of scales based on critical bands (*i.e.*, filter banks that attempt to represent the response of the auditory nerve) and psychoacoustic tests. Thus, the most relevant scales in this field are the Mel, Bark, and Equivalent Rectangular Bandwidth (ERB) scales [14], [26]. Many masking phenomena can be explained regarding sound frequency ranges known as critical bands. A critical band can be related to a bandpass filter whose frequency response roughly corresponds to the tuning curves of auditory neurons. In psychoacoustic experiments, a critical band defines a frequency interval in which perception changes abruptly when a narrow-band sound stimulus is modified to have frequency components beyond the band. When two competing sound signals pass energy through a critical band filter, the sound with higher energy within the critical band dominates perception and masks the other sound [27]. The shapes of critical band filters have been determined in experiments in which broadband low-pass or high-pass noise is used to mask a tone. As the cutoff frequency of a low-pass noise is increased beyond the tone frequency f. An analytical expression that maps the frequency f to the critical band rate m is:

$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{6}$$

called Mel scale [27], the mapping is approximately linear in frequency up to 1kHz and logarithmic at higher frequencies, such tonal frequency can be adjusted to the sampling frequency ranges of micro-earthquakes. The analytical representation of the Bark scale can be found [28], and that for the ERB scale in [29].

The number of frequency bands into which each scale can be divided is variable and is typically a power of 2. This processing in the frequency domain and based on psychoacoustic scales can help us extract features such as centroids, kurtosis, and many others from these frequency bands, resulting in a more extensive representation of the speech and audio signals. In this work, 14 spectral features were used, as detailed in Table 1. These features can be extracted linearly or logarithmically using scales divided into 32, 64, 128, and 256 bands.

A widely used feature selection method is Statistical Dependence (SD). Its goal is to measure whether the values of a feature depend on the associated class labels or whether the two coincide by chance. Each feature value is first quantized at one of the Q_S levels, where the feature-specific quantization scale is adaptively determined so that each bin will contain approximately the same number of samples across the entire data set [30]. SD is typically measured as MI, which is given by:

$$\mathrm{MI} = \sum_{y \in Y} \sum_{z \in Z} p(y, z) \frac{p(y, z)}{p(y)p(z)} \tag{7}$$

where the probability $p(\cdot)$ of each bin y and each class z is computed by frequency of occurrence. The MI score allows the user to rank the features, selecting the ones with the highest values.

III. PROPOSED METHOD AND DATASET

The proposed system evaluates the separability of micro-earthquakes of Cotopaxi and Llaima Volcanoes in a three-dimensional latent space using machine-learning algorithms and mutual information for feature selection. The aim is to visualize and analyze these latent spaces corresponding to several volcanic events and classify them among LP, VT, and REG for Cotopaxi Volcano and among LP, VT, TR, and TEC for Llaima Volcano. The main steps of the proposed methodology are shown in Figure 1 and can be summarized as follows:

1) Preliminary processing, including signal conditioning and noise removal, generates 60-second sequences for both volcanoes.
2) Fourteen audio features are extracted from the sequences, which serve as inputs to the different network architectures

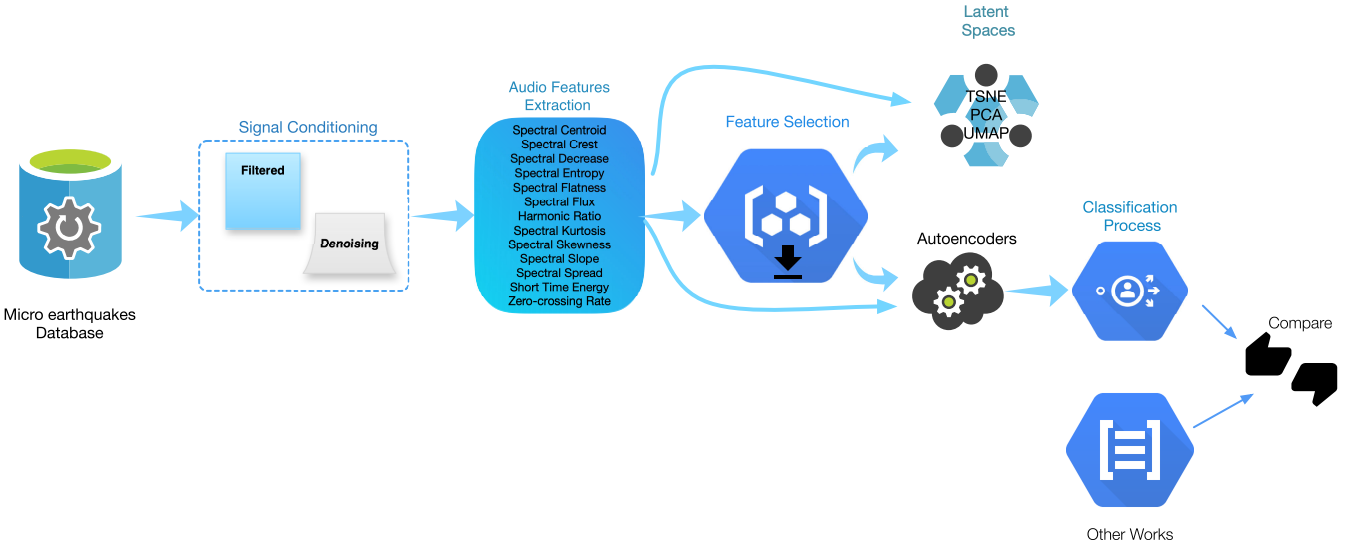

FIGURE 1. Pipeline of the proposed approach for feature reduction and classification of micro-earthquakes in Cotopaxi and Llaima Volcanoes.

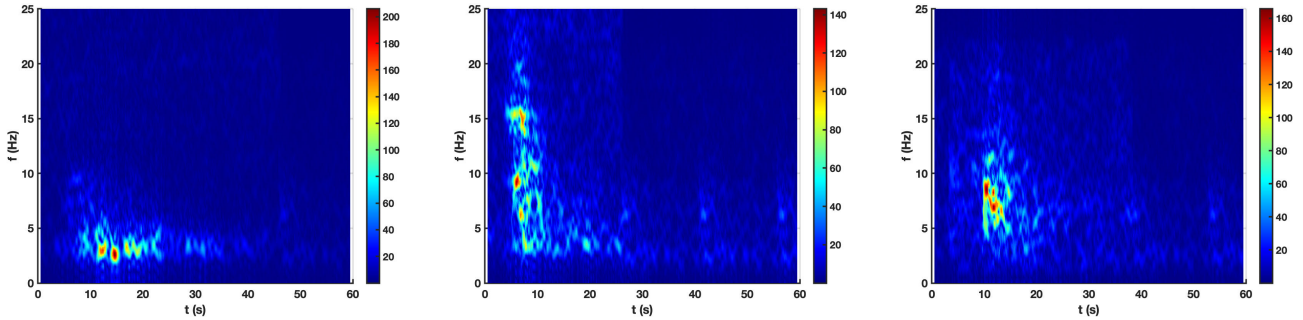

FIGURE 2. Two-dimensional cut of the spectrograms of Cotopaxi Volcano events using a 100-Hz sampling frequency: (a) LP, (b) VT, and (c) REG events.

3) Mutual information and automatic feature selection are used to reduce the number of features and improve network performance.
4) In parallel, latent spaces are observed using t-SNE, PCA, and UMAP techniques to assess existing separability. Also, autoencoders are used to generate latent spaces. Original features with and without feature selection are analyzed.
5) Classifiers based on artificial neural networks are used to detect the different types of micro-earthquakes of the two volcanoes.

All this analysis is focused on two datasets of micro-earthquakes. The first one is available from the ESeismic[1] repository [31], which is the first annotated Ecuadorian volcano seismic public repository, and the second one is available from Mendeley Data repository,[2] which is the first annotated Chilean volcano seismic public dataset [6].

The first repository contains a dataset named *MicSigV1*, which has been obtained from Cotopaxi Volcano, located in Ecuador (latitude 0°41'05'' S and longitude 78°25'54.8'' W). This dataset contains a total of 1187 independent micro-earthquakes recorded at two different seismic stations installed on flanks of the volcano and sampled at 50 Hz and 100 Hz, respectively. These particular recordings have been extracted, identified, and labeled by experts at the IGEPN from continuous monitoring seismogram recordings. The dataset contains samples distributed in five classes corresponding to LP, VT, REG, HYB, and Icequakes (ICE), and distributed as follows: $N_{LP} = 1044$, $N_{VT} = 101$, $N_{REG} = 27$, $N_{HYB} = 8$, and $N_{ICE} = 7$. Since the last types of events are scarce, all of them have been labeled as

[1]ESeismic Repository was provided by courtesy of the IGEPN and collaborators, and it is available at http://www.igepn.edu.ec/eseismic_web_site/index.php

[2]Llaima dataset was provided by courtesy of the Observatorio Vulcanológico de los Andes Sur (OVDAS). It is available at http://dx.doi.org/10.17632/dv8nwdd36k.1

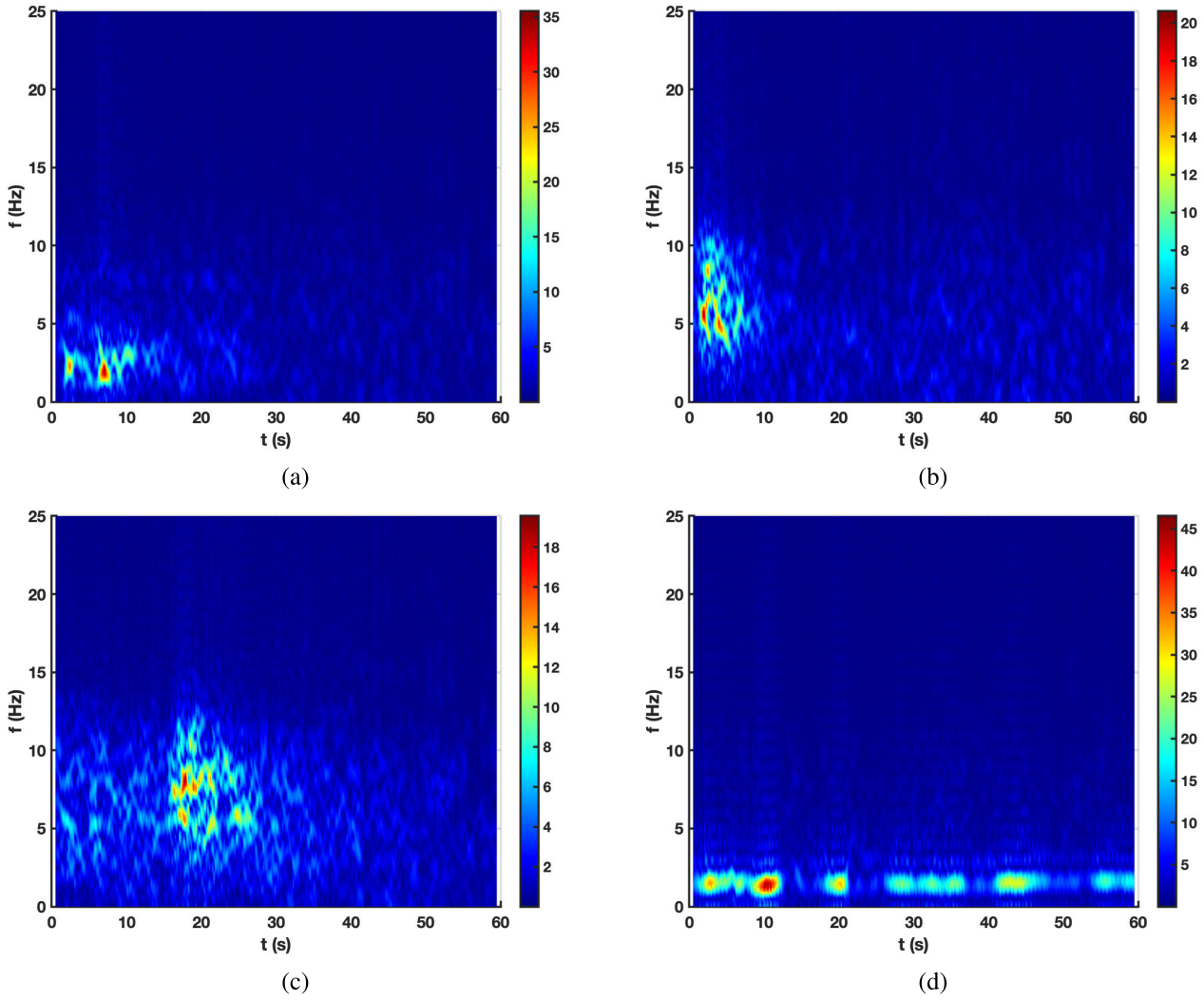

FIGURE 3. Two-dimensional cut of the spectrograms of the typical Llaima Volcano events using a 100-Hz sampling frequency: (a) LP, (b) VT, (c) TEC, and (d) TR events.

others. Micro-earthquakes at Cotopaxi Volcano have a typical duration below 60 seconds.

The second dataset was obtained from Llaima Volcano in Chile (latitude 38°41' S and longitude 71°44' W). This dataset contains 3592 independent micro-earthquakes recorded at one of the seven seismic stations deployed at Llaima Volcano and sampled at 100 Hz. These particular recordings have been normalized using their maximal values and labeled by the specialists from OVDAS. The dataset contains samples distributed in four classes corresponding to LP, VT, TR, and TEC, and distributed as follows: $N_{LP} = 1310$, $N_{VT} = 304$, $N_{TR} = 490$, and $N_{TEC} = 1488$. Micro-earthquakes at Llaima Volcano also have a typical duration below 60 seconds.

IV. EXPERIMENTS AND RESULTS

This section presents the outcomes of our experiments along with relevant details regarding the various steps involved in data processing, including preprocessing and feature extraction, feature selection, low-dimensional representation, classification, and sensitivity analysis. The methodology employed in this study is illustrated and summarized in Figure 1.

A. PREPROCESSING AND FEATURE EXTRACTION

The Cotopaxi Volcano database encompasses five types of events, but the hybrids and icequakes together total only 15 instances. Initially, we worked with the entire database and obtained promising results. However, due to the limited number of signals from these events, we found that the results lacked generalization. Therefore, we focused on the three event types with more signals (*i.e.*, LP, VT, and REG), restricting the duration to a maximum of 60 seconds and upsampling all the signals to 100 Hz. Figure 2 shows the spectrograms of typical examples of LP, VT, and REG events.

TABLE 2. Classification results for different scales and the number of bands in Cotopaxi Volcano.

Type of Characteristics	Number of Features	Supervised Precision %	Supervised Recall %	Supervised Accuracy %
Linear	3848	91.42	84.16	94.60
Linear	3000	33.69	35.49	74.72
Linear	2000	30.85	30.37	70.17
Linear	1000	34.49	35.72	72.73
Linear	400	34.06	39.80	50.85
Linear	296	34.06	39.69	50.85
Linear	115	27.65	18.42	31.82
Linear	59	26.34	19.34	50.28
ERB-32	3848	88.52	87.12	95.45
ERB-32	3000	32.28	32.30	73.01
ERB-32	2000	29.53	26.26	55.11
ERB-32	1000	35.89	35.89	29.83
ERB-32	400	34.23	38.53	23.86
ERB-32	296	34.24	38.97	25.00
ERB-32	115	36.23	40.17	32.10
ERB-32	59	33.47	39.19	25.57

TABLE 3. Centroids and standard deviation for UMAP, t-SNE, PCA and autoencoders with 3848 Cotopaxi volcano features.

Type	Centroids	σ
LP-TSNE	$(4.29, -12.69, -1.11)$	5.78
VT-TSNE	$(8.37, 14.80, -0.09)$	6.96
REG-TSNE	$(0.46, 5.92, -2.92)$	8.94
LP-PCA	$(-1.53, 0.29, -0.91)$	46.49
VT-PCA	$(-2.38, -23.39, 6.44)$	58.12
REG-PCA	$(-7.25, 2.76, -2.38)$	35.20
LP-UMAP	$(-3.44, -2.76, 0.76)$	7.01
VT-UMAP	$(-2.30, -2.62, 0.96)$	5.29
REG-UMAP	$(-4.83, -2.67, 0.27)$	6.68
LP-Autoencoder	$(-6.02, -1.88, -2.51)$	10.56
VT-Autoencoder	$(1.01, 1.26, -0.62)$	3,80
REG-Autoencoder	$(3.77, -0.26, 4.31)$	4.49

TABLE 4. Classification results for different scales and number of bands in Llaima Volcano.

Type of Characteristics	Number of Features	Supervised Precision %	Supervised Recall %	Supervised Accuracy %
ERB-32	3848	92.50	91.62	94.07
ERB-32	400	93.02	93.69	94.44
ERB-32	300	79.00	81.16	75.81
ERB-32	200	80.65	83.45	79.24
ERB-32	121	80.65	83.45	79.24

Conversely, the Llaima Volcano database comprises four event types (*i.e.*, LP, TEC, TR, and VT) with a maximum duration of 60 seconds. Figure 3 depicts the spectrograms for these typical Llaima events. The evolution of the different events was analyzed within a 60-second time frame, enabling the observation of micro-earthquake development.

From the signals of Cotopaxi and Llaima Volcanoes, we extracted 14 features using psychoacoustic scales and a sliding window size of 100 samples. A rectangular window with 80% overlap was employed, resulting in 296 values for each of the 14 features. The Harmonic Ratio feature, containing many zeros, was excluded. Consequently, the final dataset included 3848 features for each event. The MI method was then utilized for feature reduction.

Then, a supervised classifier based on autoencoders was employed to assess the effectiveness of feature

TABLE 5. Predominant features according to Mutual Information for Llaima Volcano.

Type of characteristic	Number of features	Number of features with mutual information	Percentage selected
Spectral Centroid	296	7	2.36
Spectral Crest	296	11	3.72
Spectral Decrease	296	7	2.36
Spectral Entropy	296	5	1.69
Spectral Flatness	296	5	1.69
Spectral Flux	296	9	3.04
Spectral Kurtosis	296	5	1.69
Spectral Rolloff Point	296	295	99.66
Spectral Skewness	296	10	3.38
Spectral Slope	296	4	1.35
Spectral Spread	296	8	2.70
Short Time Energy	296	10	3.38
Zero-crossing Rate	296	24	8.11

reduction. Our proposed autoencoder architecture included a low-dimensional projection encoder generating a 3-dimensional latent space with four layers (32, 16, 8, and 3 neurons) using the leaky-ReLU layer activation function. This encoder was linked to a classifier based on neural networks composed of three hidden layers (10, 20, and 30 neurons) with the hyperbolic tangent activation function and a final softmax layer. The complete network was trained with 70% of the database for training and 30% for testing, aiming to create low-dimensional embeddings that minimize classification errors and allow for comparisons of relative separations among different classes.

The network training process involved extracting features from all signal segments, conducting the training process in batches, and calculating results in terms of accuracy (A), precision (P), and recall (R), given by the following equations:

$$A = \frac{\sum_{i=1}^{N} TP(i)}{N_s} \tag{8}$$

$$P = \frac{1}{N}\left[\frac{TP(1)}{\sum_{j=1}^{N} E(j,1)} + \ldots + \frac{TP(N)}{\sum_{j=1}^{N} E(j,N)}\right] \tag{9}$$

$$R = \frac{1}{N}\left[\frac{TP(1)}{\sum_{j=1}^{N} E(1,j)} + \ldots + \frac{TP(N)}{\sum_{j=1}^{N} E(N,j)}\right] \tag{10}$$

where $TP(i)$ is the true positives for class i, $E(i,j)$ is the errors made when class i is classified as class j, N is the total number of samples, and N_s is the number of classes.

B. COTOPAXI VOLCANO

Cotopaxi Volcano's database was initially unbalanced, and linear features were extracted using the MEL, BARK, and ERB psychoacoustic scales in 32, 64, 128, and 256 bands.

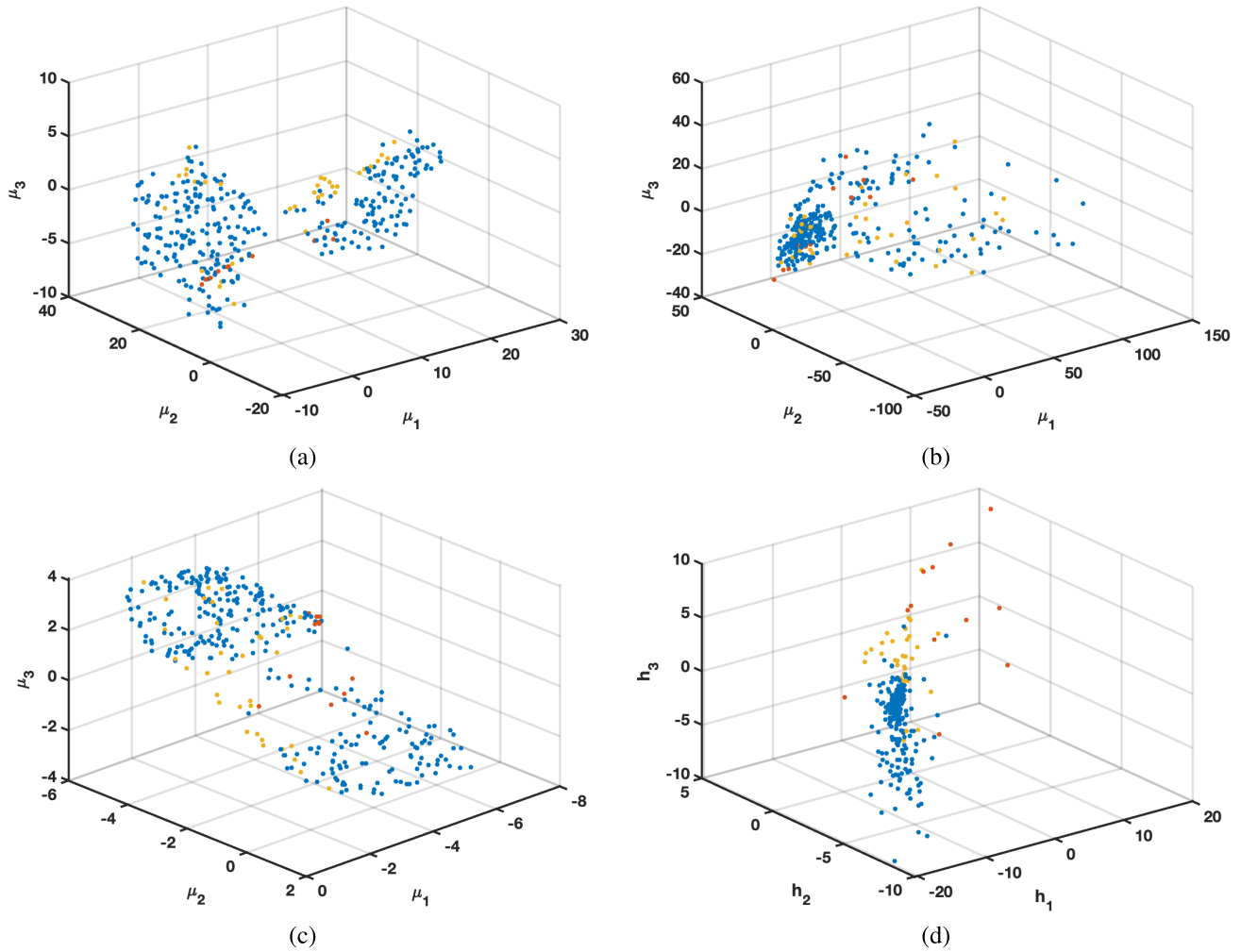

FIGURE 4. Embedding spaces represented as three latent variables for (a) t-SNE, (b) PCA, (c) UMAP, and (d) supervised autoencoder, using 3848 ERB-32 scale features from Cotopaxi Volcano. In blue are LP events, in orange are VT events, and in red are REG events.

However, results were inconsistent for the MEL and BARK scales. Consequently, we focused on linear psychoacoustic scales and ERB with 32 bands. Figures 4(a), 4(b), and 4(c) illustrate the latent spaces obtained for t-SNE, PCA, and UMAP, respectively, with 3848 features. While separability among the three types of events is observable, the abundance of LP events limits the comparison to the other types. Feature reduction using Mutual Information (MI) was attempted due to the deterministic nature of t-SNE, PCA, and UMAP mappings. Table 2 demonstrates that the ERB psychoacoustic scale outperforms the linear one, reaching 95.45% accuracy with 3848 features and 270 training epochs. Feature reduction did not enhance results, but relevant features identified by the MI method included the Spectral Rolloff Point and Spectral Flux. The imbalance among Cotopaxi different events primarily influences classification behavior. Additionally, Figure 2(d) displays the latent space using autoencoders with the same 3848 features, revealing clustering of volcanic events similar to t-SNE, PCA, and UMAP.

The latent spaces presented in this study can be a subjective form of analysis as the events are represented in three dimensions, so we calculate the centroids in Table 3, together with the standard deviation of the Euclidean distance of each point with respect to the centroid. If we observe the σ values, the highest values correspond to the PCA technique, which indicates that the points are more dispersed with respect to the centroids; in the case of t-SNE, σ is low with respect to the centroids, but two point clusters can confuse the classifier. In the case of UMAP and the autoencoders, the σ value is low and the points with respect to the centroid have reasonable values, so we considered the most suitable techniques to work with the classifier.

C. LLAIMA VOLCANO

Following the Cotopaxi results, Llaima Volcano analysis used only the ERB psychoacoustic scale with 32 bands and all 3848 features. MI feature reduction was employed, and results are tabulated in Table 4. Ac can be seen, the classifier with the best performance used the top 400 features.

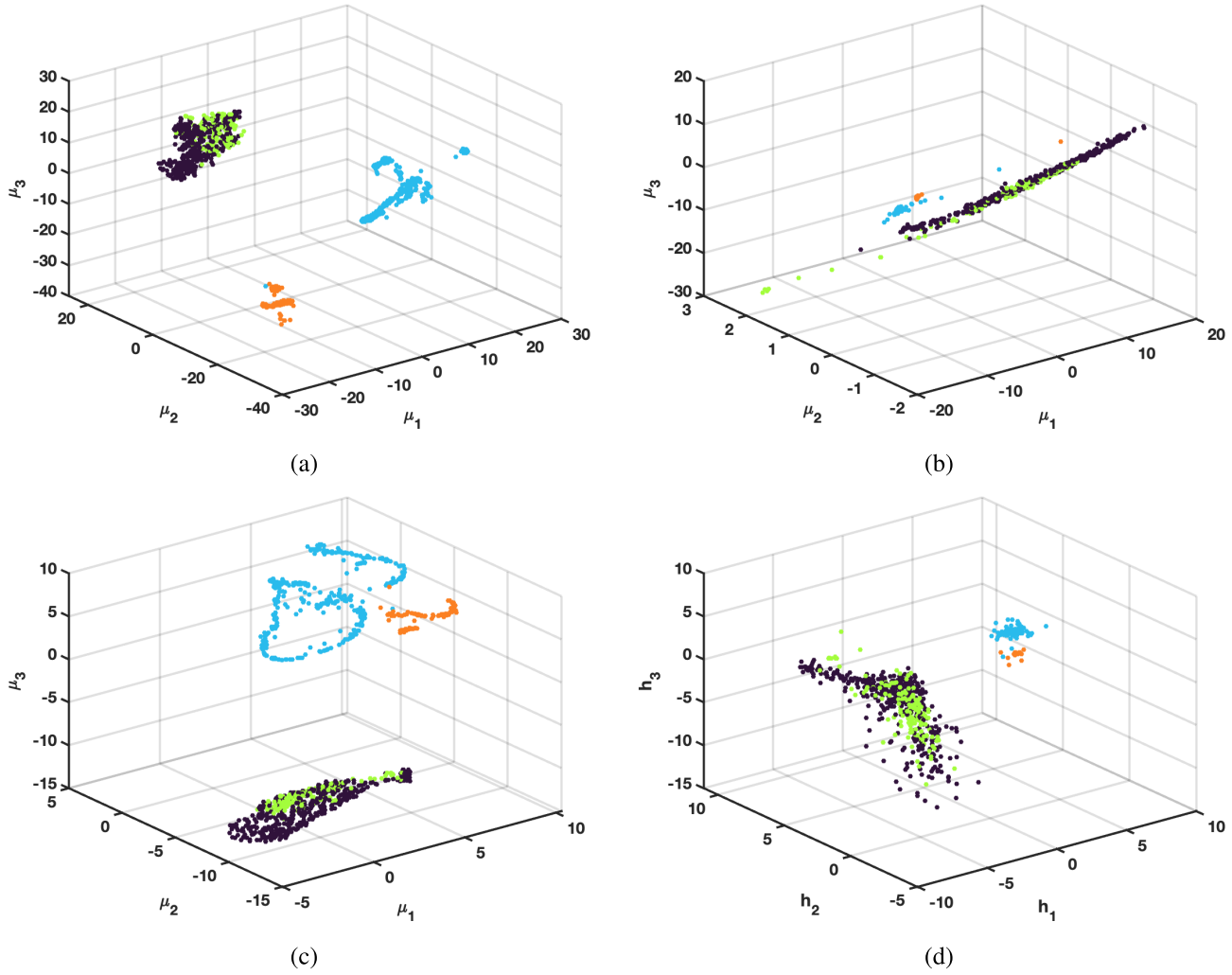

FIGURE 5. Embedding spaces represented as three latent variables for (a) t-SNE, (b) PCA, (c) UMAP, and (d) supervised autoencoder, using 400 ERB-32 scale features from Llaima Volcano. In black are TEC events, in green LP events, in light blue TR events, and in orange VT events.

Table 5 shows that, unlike Cotopaxi, Llaima required only 400 features, with 99.66% being the Spectral Rolloff Point. The Spectral Flux was less critical than in Cotopaxi. Latent spaces using t-SNE, PCA, UMAP, and autoencoders (for 150 training epochs) are shown in Figures 5(a), 5(b), 5(c), and 5(d), respectively, indicating separability among the four types of events.

For the Llaima volcano, we also calculated the centroids and the standard deviation of the Euclidean distance, and the results are presented in Table 6. If we observe the results and the σ values as in the Cotopaxi volcano, we could conclude that the autoencoders and the PCA have the best performance, but here as the database is better balanced, we can see that for the PCA case, the centroids of the 4 events are very close to each other, which could confuse the events, so as in the case of Cotopaxi, the autoencoders are repeated considered as the best results for classification.

D. OTHER TECHNIQUES

To validate the MI results and rule out overfitting, we contrasted our feature selection with the random forest technique [32]. Notable results are summarized in Table 7, showing a close agreement with MI outcomes. The only relevant improvement for the linear features of the Cotopaxi volcano, but it is corroborated that the ERB-32 features have a better performance. As for which are the most relevant features selected by the random forest technique for the Cotopaxi volcano is the Spectral RolloffPoint with 122 out of 296 corresponding to 41.22%, followed by Spectral Flux with 31 out of 296 corresponding to 10.47%. For the Llaima Volcano, it is Spectral RolloffPoint with 198 out of 296, which corresponds to 66.89%. Continuing with the analysis, the Cotopaxi database was merged with the Llaima database to observe how general the characteristics were found for this. The REG events of Cotopaxi were matched with the TC of Llaima, obtaining results close to those obtained by

TABLE 6. Centroids and standard deviation for UMAP, t-SNE, PCA and autoencoders with 400 LLaima volcano features.

Type	Centroids	σ
TEC-TSNE	$(-9.22, 16.13, 5.38)$	18.17
LP-TSNE	$(11.51, -14.81, -2.99)$	29.79
TR-TSNE	$(-6.39, 15.46, 8.62)$	15.57
VT-TSNE	$(-13.29, -14.62, -29.10)$	39.84
TEC-PCA	$(8.02, 0.08, -2.73)$	9.44
LP-PCA	$(-11.55, -0.21, 1.49)$	12.14
TR-PCA	$(6.39, 0.23, -5.14)$	11.97
VT-PCA	$(-9.80, -0.38, 4.28)$	11.57
TEC-UMAP	$(-12.69, -9.67, -7.72)$	20.59
LP-UMAP	$(7.51, 0.67, 0.50)$	6.22
TR-UMAP	$(-12.08, -8.92, -7.22)$	18.60
VT-UMAP	$(7.39, -4.81, -2.09)$	9.70
TEC-Autoencoder	$(-3.85, 3.56, -2.38)$	6.99
LP-Autoencoder	$(2.73, 0.94, 5.03)$	5.15
TR-Autoencoder	$(-5.49, 0.72, -0.45)$	7.38
VT-Autoencoder	$(5.09, 3.35, -0.51)$	5,67

TABLE 7. Classification results using random forest for feature selection for Cotopaxi and Llaima volcanoes.

Type of Characte	Volcano	Number of Features	Supervised Precision %	Supervised Recall %	Supervised Accuracy %
Linear	Cotopaxi	765	86.83	80.21	93.47
ERB-32	Cotopaxi	259	90.20	87.92	94.87
ERB-32	Llaima	225	87.63	91.40	88.97

TABLE 8. Results of the classification of merging the bases of Cotopaxi and Llaima volcanoes.

Type of Characteristics	Number of Features	Supervised Precision %	Supervised Recall %	Supervised Accuracy %
ERB-32	3848	90.80	89.83	93.71
ERB-32 (RF)	900	92.06	90.19	94.38
ERB-32 (MI)	400	86.43	83.32	90.78

separate databases with an increase in the number of features, as shown in Table 8.

In order to compare with the best option obtained with the Llaima volcano, we have decided to set it at 400 to with a reasonable result. On the other hand, it is maintained that the main features are Spectral RolloffPoint and Spectral Flux.

E. THE MOST RELEVANT FEATURES

Based on both volcanoes, it is crucial to utilize the Spectral Rolloff Point for event classification. Furthermore, we can also observe the presence of the Spectral Flux in both cases. The Spectral Rolloff Point represents the 95th percentile of the power spectral distribution, and it serves to differentiate between voiced and unvoiced speech. Unvoiced speech tends to have a significant amount of energy in the high-frequency range of the spectrum, while voiced speech and music predominantly concentrate their energy in lower-frequency bands. This measurement quantifies the skewness of the spectral shape, with higher values indicating a right-skewed distribution [33].

On the other hand, Spectral Flux (Delta Spectrum Magnitude) is the 2-norm of the frame-to-frame spectral amplitude difference vector. Generally, music has a higher rate of change and goes through more drastic frame-to-frame changes than speech does. Thus, this value is higher for music than for speech. Note that speech alternates periods of transition (consonant – vowel boundaries) and periods of relative stasis (vowels), where music typically has a more constant rate of change. This method is similar to Hawley's, which attempts to detect harmonic continuity in music [33].

If we make an analogy with the voice and music from which these two characteristics come, they seek to find the voiced and unvoiced sounds and the evolution from an unvoiced sound to a voiced one. Thus, these features are indispensable for the classification of the events of Cotopaxi and Llaima Volcanoes, with the hypothesis that the LP events are considered as voiced, the VT as unvoiced, and the other events they are considered as an evolution to go from voiced to unvoiced.

TABLE 9. Comparison of techniques for classifying Cotopaxi and Llaima Volcano events.

Volcano	Technique	Precision %	Recall %	Accuracy %
Cotopaxi	TFFCV [3]	–	–	95.80
Cotopaxi	ERB-32/3848	88.52	87.12	95.45
Llaima	SeismicNet [6]	–	93.87	98.04
Llaima	ERB-32/400	93.02	93.69	94.44

Finally, comparisons with other techniques are presented in Table 9. Firstly, in [3], the classification of three events (*i.e.*, LP, VT, and HYB) for Cotopaxi Volcano was performed, achieving an accuracy percentage of 95.80%, compared to the 95.45% we have just obtained. It is important to note that the method we use exhibits lower dimensionality. Secondly, in [6], the SeismicNet network, a variant of the SoundNet network [34], is proposed for classifying the four types of events at Llaima Volcano. This system obtained a recall of 93.87% and an accuracy of 98.04%, in contrast to our results of 93.69% and 94.44%, respectively. Whereas the performance in classification can be considered as comparable, these results show that they can be obtained with a highly simplistic network scheme in comparison with these deep models, assuming that the input space is expressed in terms of spectral representations.

V. DISCUSSION AND CONCLUSION

The micro-earthquakes produced by the eruptive process of volcanoes are subject to significantly changing dynamics, challenging to model, and almost impossible to predict. In this work, we used the databases of two volcanoes whose main similarity is strato volcanos. The rest of the characteristics differ in their sizes, the amount of snow, and many others. However, their location concerning populated sites for a possible eruption is essential. We used a variety of spectral characteristics in both cases based on features that are well-known in audio digital processing.

Aiming to gain interpretability of the results in the multi-class classifiers, we scrutinized three different manifold algorithms for identifying unsupervised and intrinsic separability among the different classes and autoencoder-based architectures that were subsequently used

to yield supervised classifiers. Latent spaces were checked to be low-dimensional, representing evidence for using these and other machine-learning schemes with advantage in the problem of multi-class micro-earthquake classification. Data from other volcanoes could exhibit this property by following our methodology.

Other ways of parameterization could be used, for example, features in the time domain. In [3], we already established a mixture of time and frequency characteristics, obtaining reasonable results, but it was observed that the most relevant were those extracted in the frequency domain, in addition to the fact that noise affected them less. At the beginning of the research, certain features were tested in the time domain, obtaining varied results, and when introducing infrequent events such as those of the Cotopaxi volcano, they confused the classifier, obtaining results between 50% and 60% accuracy. In the end, the linear features were discarded, and only the ERB-32 was used. These obtained features conserve certain information in time since they are the product of a signal windowing over time.

From our results, the extraction of psychoacoustic characteristics, specifically the ERB scale, showed great versatility in classifying LP, VT, and other events regardless of the volcano. Using MI to reduce the number of features improved the classifier results considerably. Latent spaces were a helpful tool for visualization and checking for overfitting in this case. Moreover, the suitability of low-dimensional latent spaces for supporting multiclass data-driven schemes shows that transfer learning could be readily done in this kind of signal, and other machine learning paradigms will be exploited with advantage in this setting. Finally, interpretability can also be attained, for instance, in terms of the inter-class separation among classes that can be observed in the 3-dimensional maps.

We could determine by using MI that two features are predominant for the two volcanoes: Spectral Rolloff Point and Spectral Flux. In audio-digital processing, these measurements allow for characterizing the voiced and unvoiced sounds and their evolution of passing from one to the other as used in voice and audio applications. There is the presumption that the LP events are considered to be voiced and the VT events are unvoiced. As for the others, they remain as an evolution towards passing from voiced to unvoiced and vice versa. In a broader context, when referring to voiced signals, it is considered that said signals concentrate more narrow-band energy and tend to be more predictable. On the other hand, unvoiced signals are more similar to noise, so their analysis is more complicated when classifying them.

The results obtained demonstrate that various volcanoes share similarities in their micro-earthquake activities. This finding paves the way for a new paradigm in volcano analysis, where a generalized approach can be employed. Such an approach would facilitate the development of volcano-independent algorithms capable of improving the current generation of detection algorithms by leveraging insights from various volcanic sources. In essence, this article underscores the audio feature similarities in micro-earthquake signals between different volcanoes, pointing toward the potential for a more advanced generation of algorithms that surpass current methods.

The classification accuracy obtained was high for both volcanoes. However, improvements need to be addressed shortly to yield online surveillance systems based on multi-class classifiers of this kind. The imbalance in the number of events and their very different duration has to be considered. Some other events are scarce and were not considered here, as documented in the experiments. Nevertheless, the low-dimensional latent spaces exhibited separability among classes using adequate spectral features, thus opening new doors to their consideration and adaptation in studying and monitoring these and other volcanoes.

REFERENCES

[1] B. A. Chouet, "Long-period volcano seismicity: Its source and use in eruption forecasting," *Nature*, vol. 380, no. 6572, pp. 309–316, Mar. 1996.

[2] B. A. Chouet and R. S. Matoza, "A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption," *J. Volcanol. Geothermal Res.*, vol. 252, pp. 108–175, Feb. 2013.

[3] R. Lara-Cueva, P. Bernal, M. G. Saltos, D. S. Benítez, and J. L. Rojo-Alvarez, "Time and frequency feature selection for seismic events from Cotopaxi Volcano," in *Proc. Asia–Pacific Conf. Comput. Aided Syst. Eng.*, 2015, pp. 129–134.

[4] D. Cárdenas-Peña, M. Orozco-Alzate, and G. Castellanos-Dominguez, "Selection of time-variant features for earthquake classification at the Nevado–del–Ruiz Volcano," *Comput. Geosci.*, vol. 51, pp. 293–304, Feb. 2013.

[5] F. Lara, R. Lara-Cueva, J. C. Larco, E. V. Carrera, and R. León, "A deep learning approach for automatic recognition of seismo-volcanic events at the cotopaxi volcano," *J. Volcanology Geothermal Res.*, vol. 409, Jan. 2021, Art. no. 107142.

[6] J. P. Canário, R. Mello, M. Curilem, F. Huenupan, and R. Rios, "In-depth comparison of deep artificial neural network architectures on seismic events classification," *J. Volcanol. Geothermal Res.*, vol. 401, Sep. 2020, Art. no. 106881.

[7] G. Cortés, R. Arámbula, L. A. Gutiérrez, C. Benítez, J. Ibánez, P. Lesage, I. Alvarez, and L. Garcia, "Evaluating robustness of a HMM-based classification system of volcano-seismic events at Colima and popocatepetl volcanoes," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, vol. 2, 2009, pp. 1012–1015.

[8] M. Titos, A. Bueno, L. García, and C. Benítez, "A deep neural networks approach to automatic recognition systems for volcano-seismic events," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 5, pp. 1533–1544, May 2018.

[9] A. Bueno, M. Titos, L. García, I. Álvarez, J. Ibañez, and C. Benítez, "Classification of volcano-seismic signals with Bayesian neural networks," in *Proc. 26th Eur. Signal Process. Conf. (EUSIPCO)*, 2018, pp. 2295–2299.

[10] M. Titos, A. Bueno, L. García, M. C. Benítez, and J. Iba nez, "Detection and classification of continuous volcano-seismic signals with recurrent neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 1936–1948, Apr. 2019.

[11] L. Ruff, J. R. Kauffmann, R. A. Vandermeulen, G. Montavon, W. Samek, M. Kloft, T. G. Dietterich, and K.-R. Müller, "A unifying review of deep and shallow anomaly detection," *Proc. IEEE*, vol. 109, no. 5, pp. 756–795, May 2021.

[12] P. Baldi, "Autoencoders, unsupervised learning, and deep architectures," in *Proc. ICML Workshop Unsupervised Transf. Learn.*, 2012, pp. 37–49.

[13] X. Kong, X. Jiang, B. Zhang, J. Yuan, and Z. Ge, "Latent variable models in the era of industrial big data: Extension and beyond," *Annu. Rev. Control*, vol. 54, pp. 167–199, 2022.

[14] A. Lerch, *An Introduction to Audio Content Analysis Applications in Signal Processing and Music Informatics*. Piscataway, NJ, USA: IEEE Press, 2012.

[15] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Trans. Neural Netw.*, vol. 5, no. 4, pp. 537–550, Jul. 1994.
[16] R. Bellman and R. Kalaba, "On adaptive control processes," *IRE Trans. Autom. Control*, vol. 4, no. 2, pp. 1–9, Nov. 1959.
[17] B. Ghojogh, M. Crowley, F. Karray, and A. Ghodsi, *Elements of Dimensionality Reduction and Manifold Learning*. Berlin, Germany: Springer, 2023.
[18] A. J. Izenman, "Introduction to manifold learning," *Wiley Interdiscipl. Rev., Comput. Statist.*, vol. 4, no. 5, pp. 439–446, Sep. 2012.
[19] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform manifold approximation and projection for dimension reduction," 2018, *arXiv:1802.03426*.
[20] W. H. López-Pinaya, S. Vieira, R. Garcia-Dias, and A. Mechelli, "Autoencoders," in *Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 193–208.
[21] S. E. Chazan, S. Gannot, and J. Goldberger, "Deep clustering based on a mixture of autoencoders," in *Proc. IEEE 29th Int. Workshop Mach. Learn. Signal Process. (MLSP)*, Oct. 2019, pp. 1–6.
[22] C. P. Bernal Oñate, F.-M. Melgarejo Meseguer, E. V. Carrera, J. J. Sánchez Muñoz, A. García Alberola, and J. L. Rojo Álvarez, "Different ventricular fibrillation types in low-dimensional latent spaces," *Sensors*, vol. 23, no. 5, p. 2527, Feb. 2023.
[23] C.-P. Bernal-Oñate, E. V. Carrera, F.-M. Melgarejo-Meseguer, R. Gordillo-Orquera, A. Garcí-A-Alberola, and J. L. Rojo-álvarez, "Atrial fibrillation detection with spectral manifolds in low-dimensional latent spaces," *IEEE Access*, vol. 11, pp. 103364–103376, 2023.
[24] K. Kalgaonkar, R. Hu, and B. Raj, "Ultrasonic Doppler sensor for voice activity detection," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 754–757, Oct. 2007.
[25] L. R. Rabiner and R. W. Schafer, *Theory and Applications of Digital Speech Processing*, 1st ed. London, U.K.: Pearson, 2011.
[26] V. Pulkki and M. Karjalainen, *Communication Acoustics an Introduction to Speech, Audio and Psychoacoustics*. Hoboken, NJ, USA: Wiley, 2015.
[27] D. O'Shaughnessy, *Speech Communications Human and Machine*, 2nd ed. Hoboken, NJ, USA: Wiley, 1999.
[28] H. Traunmüller, "Analytical expressions for the tonotopic sensory scale," *J. Acoust. Soc. Amer.*, vol. 88, no. 1, pp. 97–100, Jul. 1990.
[29] B. R. Glasberg and B. C. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.*, vol. 47, nos. 1–2, pp. 103–138, 1990.
[30] J. Pohjalainen, O. Räsänen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits," *Comput. Speech Lang.*, vol. 29, no. 1, pp. 145–171, Jan. 2015.
[31] N. Pérez, D. Benítez, F. Grijalva, R. Lara-Cueva, M. Ruiz, and J. Aguilar, "ESeismic: Towards an ecuadorian volcano seismic repository," *J. Volcanol. Geothermal Res.*, vol. 396, May 2020, Art. no. 106855.
[32] W.-Y. Loh, "Regression trees with unbiased variable selection and interaction detection," *Statistica Sinica*, vol. 12, pp. 361–386, Apr. 2002.
[33] E. Scheirer and M. Slaney, "Construction and evaluation of a robust multifeature speech/music discriminator," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 1997, pp. 1331–1334.
[34] Y. Aytar, C. Vondrick, and A. Torralba, "SoundNet: Learning sound representations from unlabeled video," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 892–900.

CARLOS-PAUL BERNAL-OÑATE (Member, IEEE) received the Graduate degree in electronics and telecommunications engineering from Escuela Politécnica del Ejército (ESPE), Ecuador, in 2000, and the master's degree from the Military Institute of Engineering (IME), Rio de Janeiro, Brazil, in 2005. He is currently pursuing the Ph.D. degree in information and communication technologies with King Juan Carlos University, Spain. He is also a Principal Professor with the Electric, Electronics and Telecommunications Department, Universidad de las Fuerzas Armadas—ESPE. His research interests include digital speech processing, adaptive filtering, deep learning, and machine learning.

ENRIQUE V. CARRERA (Senior Member, IEEE) received the bachelor's degree in electronic engineering from Armed Forces University—ESPE, Ecuador, in 1992, the master's degree in electrical engineering from the Pontifical Catholic University of Rio de Janeiro, Brazil, in 1996, and the Ph.D. degree in computer engineering from the Federal University of Rio de Janeiro, Brazil, in 1999. He was a Visiting Scholar with the University of Rochester, USA, in 1999. From 2000 to 2004, he was a Postdoctoral Associate with the Department of Computer Science, Rutgers University, USA. He was also an Associate Professor with the University of San Francisco of Quito, Ecuador, from 2004 to 2011. Since 2011, he has been a Professor with Universidad de las Fuerzas Armadas—ESPE. He has also been collaborated as an external Professor with King Juan Carlos University, Spain, since 2015. He has participated in approximately 20 research projects, and holds more than 100 scientific publications in recognized journals and conferences with more than 3400 citations. His main research interests include signal processing and artificial intelligence.

FRANCISCO-MANUEL MELGAREJO-MESEGUER received the bachelor's degree in telecommunication engineering and the master's degree from Miguel Hernández University, Elche, Spain, in 2014 and 2016, respectively, and the Ph.D. degree in multimedia and communications from King Juan Carlos University, Madrid, in 2019. From 2015 to 2019, he was a Predoctoral Fellow with Hospital Clínico Universitario Virgen de la Arrixaca, Murcia. In 2019, he was an Assistant Professor with the University of Murcia. Since 2020, he has been an Associate Professor with King Juan Carlos University, specializing in bio-signal processing, deep learning, and machine learning, with over 15 publications and 100 citations. His research interests include statistical learning methods for signal and image processing, the study of arrhythmia mechanisms, the development of robust signal processing techniques, addressing inverse problems in cardiology sensors, and applying data science to seismic signal processing.

RODOLFO GORDILLO-ORQUERA received the B.Sc. degree in electronic and telecommunications engineering and the M.Sc. degree from the Faculty of Electronic Engineering, Escuela Politécnica del Ejército, Ecuador, in 1996 and 2002, respectively, and the Ph.D. degree in information and communication technologies from Universidad Rey Juan Carlos, Spain, in 2019. He is currently a Principal Professor with the Electric, Electronics, and Telecommunications Department, Universidad de las Fuerzas Armadas (ESPE), Sangolquí, Ecuador. His research interests include machine learning, dynamical systems, adaptive systems, and intelligent control systems.

JOSÉ LUIS ROJO-ÁLVAREZ (Senior Member, IEEE) received the B.Sc. degree in telecommunication engineering from Universidade de Vigo, in 1996, and the Ph.D. degree in telecommunication engineering from Universidad Politécnica de Madrid, in 2000. He is currently a Full Professor with the Department of Signal Theory and Communications, Universidad Rey Juan Carlos, Spain. He has coauthored over 160 international articles and contributed to over 180 conference proceedings. His research interests include statistical learning methods for signal and image processing, arrhythmia mechanisms, robust signal processing methods, inverse problems in cardiology sensors, and data science for seismic signal processing.

ROMÁN A. LARA-CUEVA (Senior Member, IEEE) received the B.Eng. degree in electronic and telecommunications engineering from Escuela Politécnica Nacional, Quito, Ecuador, in 2001, the M.Sc. degree in wireless systems and related technologies from Politecnico di Torino, Turin, Italy, in 2005, and the M.Sc. and Ph.D. degrees in telecommunication networks for developing countries from Rey Juan Carlos University, Fuenlabrada, Spain, in 2010 and 2015, respectively. In 2002, he joined the Departamento de Eléctrica, Electrónica y Telecomunicaciones, Universidad de las Fuerzas Armadas—ESPE, Sangolquí, Ecuador, where he has been an Associate Professor, since 2005, and since 2016, he has been a Full Professor, where he founded and heads the Ad Hoc Network Research Center (CIRAD) and the Smart Systems Research Group (WiCOM-Energy), Universidad de las Fuerzas Armadas—ESPE, and also he has been collaborated as an external Professor with King Juan Carlos University, Spain, since 2017. He received the Prize to the Best Junior Researcher from Universidad de las Fuerzas Armadas—ESPE, in 2014, and a Best Researcher from IEEE Ecuador Section, in 2017. He has authored more than 50 refereed conference papers on topics related to wireless communications, signal processing, and machine learning. He is the author or coauthor of 13 publicly funded research projects and directed eight. His main research interests include digital signal processing and machine learning theory applied to wireless communications systems and volcano seismology, also in the scope of the Internet of Things by developing smart gadgets for smart cities. He has served as the Chair of IEEE Communications Society Ecuador, from 2020 to 2021.