Highlights

Unsupervised Volcanic Change Detection Using Deep Learning Embeddings: Case Studies of Nevados de Chillán and Llaima Volcanoes, Chile (2017–2023)

Francisco Parra

- First application of satellite embeddings (64D feature vectors) to volcanic change detection
- Analyzed two Chilean volcanoes (Nevados de Chillán and Llaima) demonstrating method generalization
- Detected 4.04 km^2 and 3.79 km^2 extreme changes respectively using unsupervised distance metrics
- Identified distinct temporal signatures: Chillán (2 peaks: 2018–2019, 2021–2022), Llaima (1 peak: 2022–2023)
- Embeddings capture information not linearly related to spectral indices (NDVI $|r| < 0.3$)

Unsupervised Volcanic Change Detection Using Deep Learning Embeddings: Case Studies of Nevados de Chillán and Llaima Volcanoes, Chile (2017–2023)

Francisco Parra[a,1,*]

[a]*Departamento de Ingeniería Informática, Universidad de Santiago de Chile, Av. Víctor Jara 3659, Santiago, 9170125, Chile*

Abstract

Traditional spectral indices for volcanic change detection (e.g., NDVI, NBR) capture primarily vegetation and thermal changes but may miss complex morphological and textural variations. Here, we present the first application of Google Earth Engine satellite embeddings—64-dimensional feature vectors derived from a Vision Transformer trained on global satellite imagery—to detect volcanic changes in an unsupervised manner. We analyzed two active Chilean volcanoes, Nevados de Chillán and Llaima, over a 7-year period (2017–2023). Using Euclidean distance in the 64-dimensional embedding space, we identified 4.04 km^2 and 3.79 km^2 of extreme change (>95th percentile), respectively, within 5-km radii of the active craters. Sensitivity analysis across percentile thresholds (P75–P99) confirmed robust detection patterns at both sites. Unsupervised K-means clustering revealed binary surface classifications at both volcanoes (Chillán: 44.4%/55.6%; Llaima: 59.6%/40.4%), with modest silhouette scores (∼0.22) reflecting gradational volcanic terrain boundaries. Time series analysis detected distinct temporal signatures: Chillán showed two major activity periods (2018–2019, 2021–2022), while Llaima exhibited a single peak (2022–2023). Weak correlations with NDVI change (Chillán: $r = 0.29$, 95% CI [0.24, 0.34]; Llaima: $r = -0.06$, 95% CI [−0.12, −0.01]) indicate that embeddings capture information not linearly related to vegetation dynamics. Consistent results across both volcanoes suggest that deep learning embeddings offer a promising unsupervised approach to volcanic monitoring that warrants further validation across diverse volcanic settings.

Keywords: Volcanic monitoring, Change detection, Deep learning, Satellite embeddings, Google Earth Engine, Unsupervised learning, Nevados de Chillán, Llaima, Remote sensing

1. Introduction

Volcanic eruptions pose significant hazards to millions of people worldwide, with approximately 800 million people living within 100 km of active volcanoes [1]. Continuous

*Corresponding author

Email address: francisco.parra.o@usach.cl (Francisco Parra)

[1]ORCID: 0009-0006-0435-1854

monitoring of volcanic activity is essential for early warning systems and hazard mitigation. Remote sensing has become indispensable for volcano surveillance, particularly in remote or inaccessible regions [2]. Traditional approaches rely on spectral indices such as the Normalized Difference Vegetation Index (NDVI), Normalized Burn Ratio (NBR), or thermal anomaly detection using brightness temperature difference (BTD) [3, 4]. While effective for detecting specific phenomena (e.g., vegetation loss, thermal emissions), these methods may miss complex changes involving texture, morphology, and multi-spectral patterns.

Recent advances in deep learning have revolutionized Earth observation, enabling the extraction of high-level semantic features from satellite imagery [5, 6, 7]. Google Earth Engine (GEE) [8] has democratized access to planetary-scale geospatial analysis, processing petabytes of satellite data. In 2023, Google released the Satellite Image Embedding dataset, comprising 64-dimensional feature vectors derived from a Vision Transformer (ViT) model [9] trained via self-supervised contrastive learning [10, 11] on global Sentinel-2 imagery [12]. These embeddings capture rich semantic information—including spectral signatures, spatial context, and textural patterns—in a compact representation suitable for downstream tasks such as classification and change detection.

Despite growing applications of deep learning in remote sensing [13, 14], embeddings have not yet been applied to volcanic monitoring. Change detection methods have evolved from traditional pixel-based comparisons [15] to deep learning approaches using convolutional neural networks [16, 17], yet most require labeled training data. Volcanic monitoring studies employ thermal anomaly detection [18, 19], InSAR deformation analysis [20], and multispectral time series [21, 22, 23], but supervised approaches require extensive ground truth, and spectral indices may not capture the full complexity of volcanic surface changes.

Nevados de Chillán volcano (36.86°S, 71.38°W) in the Chilean Andes provides an ideal case study. Following decades of quiescence, eruptive activity resumed in January 2016, producing lava flows, domes, and ash emissions through 2023 [24]. Located 80 km from the city of Chillán (population ∼180,000) and adjacent to a ski resort, the volcano poses significant societal risk. SERNAGEOMIN's Observatorio Volcanológico de los Andes del Sur (OVDAS) maintains continuous monitoring, providing valuable validation data.

Here, we present the first application of satellite embeddings to volcanic change detection. Our objectives are to: (1) quantify volcanic changes at Nevados de Chillán (2017–2023) using unsupervised distance metrics in embedding space; (2) classify volcanic surfaces via unsupervised clustering; (3) identify temporal activity patterns through time series analysis; and (4) compare embeddings with traditional spectral indices. We demonstrate that embeddings provide complementary information to conventional methods, offering a scalable approach for global volcanic monitoring.

2. Study Area

Nevados de Chillán is a volcanic complex in the Chilean Andes (36.86°S, 71.38°W, 3,212 m elevation), comprising several edifices including Volcán Nuevo (the currently active vent), Volcán Viejo, and the ice-capped Nevado de Chillán [25, 26]. The complex formed along the Liquiñe-Ofqui Fault Zone, part of the Southern Volcanic Zone of the Andes.

Historical activity includes explosive eruptions in 1861–1864, 1906, and 1973–1986 [27]. After a 25-year repose, renewed activity began in January 2016 with phreatic explosions, evolving to dome-building and effusive eruptions. Key eruptive phases include:

- **2016–2017**: Phreatic activity, fumarolic emissions, minor ash explosions
- **2017–2018**: Dome growth in Cráter Nicanor, strombolian activity
- **2018–2019**: Intense effusive phase with lava flows extending >1 km south-southwest
- **2019–2021**: Persistent degassing, minor explosive events, slow lava extrusion
- **2021–2022**: Renewed effusive activity, new lava flows
- **2022–2023**: Declining activity, continued fumarolic emissions

The volcano is monitored by OVDAS with seismic networks, GNSS stations, gas sensors, and webcams. Satellite monitoring complements ground-based observations, particularly for summit areas obscured by clouds or inaccessible due to hazards.

Llaima volcano (38.69°S, 71.73°W, 3,125 m elevation) is one of Chile's largest and most active volcanoes, located 200 km southwest of Nevados de Chillán in the Araucanía Region [26]. The volcano features a 3-km-wide summit caldera with two active vents and extensive lava flow fields extending from the summit. Historical records document frequent eruptions including major events in 1640, 1751, 1955-1957, 1994, and 2007-2009 [27]. The 2007-2009 eruption produced lava flows, pyroclastic density currents, and ash plumes reaching 12 km altitude. Following this period, activity decreased but remained at elevated levels with persistent seismicity, fumarolic emissions, and minor phreatic explosions through 2023. Llaima's proximity to Conguillio National Park and ski resorts makes monitoring critical for public safety.

For both volcanoes, study regions encompass 5-km radii ($\sim$78.5 km^2) around active craters, capturing summit areas, lava flow fields, ash fall zones, and vegetated slopes. Land cover includes alpine vegetation, bare rock, snow/ice, and recent volcanic deposits. The selection of these two volcanoes provides contrasting eruptive styles and activity patterns, enabling robust evaluation of the embedding-based methodology (Figure 1).

3. Data and Methods

3.1. Satellite Embeddings Dataset

We used the Google Earth Engine Satellite Image Embedding dataset (GOOGLE/SATELLITE_EMB [12]. This dataset provides 64-dimensional feature vectors (bands A00–A63) at 10-m spatial resolution, derived from Sentinel-2 imagery via a Vision Transformer (ViT) model trained using self-supervised contrastive learning. The model learns to cluster similar Earth surface types without labeled data. Embeddings are unit-normalized vectors on a 64-dimensional hypersphere, with Euclidean distance quantifying dissimilarity.

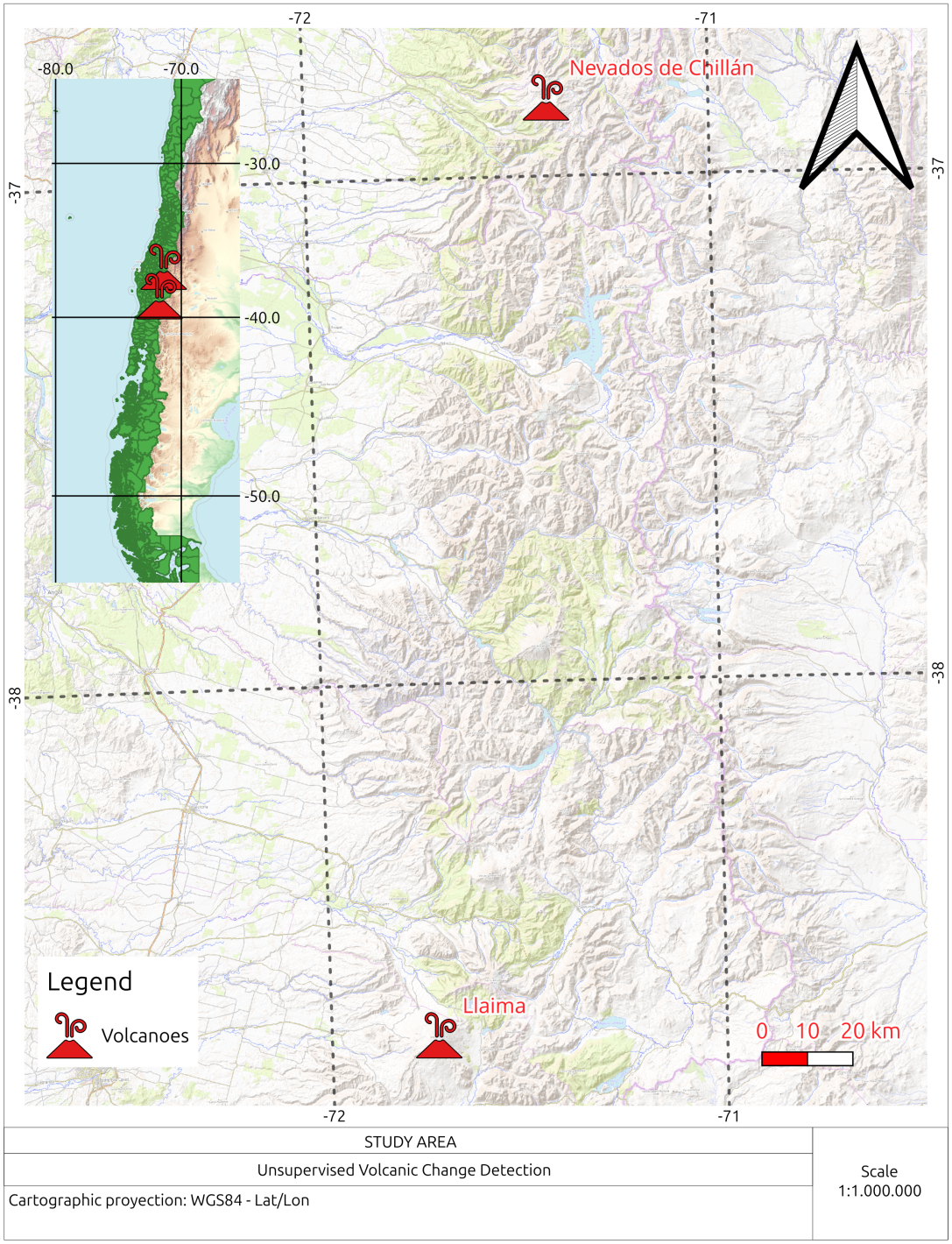

Figure 1: Study area location in the Chilean Southern Volcanic Zone. Nevados de Chillán (36.86°S, 71.38°W) and Llaima (38.69°S, 71.73°W) volcanoes are separated by approximately 200 km. Inset shows location within Chile; green dots indicate other active volcanoes.

Annual composites are available from 2017 to present. For each pixel, the embedding represents aggregated spectral, textural, and contextual information from cloud-free observations within the year. Over the study period, Nevados de Chillán averaged 66 cloud-free Sentinel-2 observations per year (range: 4–100), while Llaima averaged 102 observations per year (range: 4–154), ensuring robust annual composites despite challenging Andean cloud conditions (Table 1). Unlike raw spectral bands, embeddings encode high-level semantic features learned from billions of global training samples.

Table 1: Cloud-free Sentinel-2 observations per year for each study volcano. Years 2017–2018 show reduced counts due to partial data availability at mission start.

Volcano	Year	Total Images	Cloud-Free	% Usable
Chillán	2017	4	4	100.0
Chillán	2018	15	13	86.7
Chillán	2019	146	89	61.0
Chillán	2020	143	100	69.9
Chillán	2021	144	85	59.0
Chillán	2022	147	88	59.9
Chillán	2023	145	84	57.9
Llaima	2017	4	4	100.0
Llaima	2018	23	19	82.6
Llaima	2019	287	140	48.8
Llaima	2020	286	154	53.8
Llaima	2021	289	151	52.2
Llaima	2022	292	127	43.5
Llaima	2023	290	119	41.0

3.2. Reference Data: Sentinel-2 and Spectral Indices

For comparison, we computed traditional spectral indices from Sentinel-2 L2A (harmonized surface reflectance) imagery [28, 29]:

- **NDVI**: $(NIR - Red)/(NIR + Red)$ — vegetation health/cover
- **NBR**: $(NIR - SWIR)/(NIR + SWIR)$ — burned area/bare surfaces

We generated annual median composites (2017, 2023) using cloud-masked Sentinel-2 scenes (<30% cloud cover) filtered via the Scene Classification Layer (SCL). Pixels classified as cloud shadow (SCL=3), cloud medium probability (SCL=8), cloud high probability (SCL=9), or thin cirrus (SCL=10) were excluded, retaining vegetation (SCL=4), bare soils (SCL=5), water (SCL=6), and snow/ice (SCL=11). This provides spatially consistent comparison with embeddings.

3.3. Change Detection

We quantified volcanic change using Euclidean distance in 64-dimensional embedding space, following unsupervised change detection principles where pixel-wise differences indicate surface alterations [30]:

$$d(t_1, t_2) = \sqrt{\sum_{i=0}^{63} (A_i^{t_2} - A_i^{t_1})^2} \tag{1}$$

where A_i^t is the i-th embedding dimension at time t. We computed $d(2017, 2023)$ for each 10-m pixel within the study area. Euclidean distance was selected over cosine similarity because embeddings are already L2-normalized to the unit hypersphere, making both metrics mathematically equivalent up to a monotonic transformation. Euclidean distance provides more intuitive interpretation of magnitude differences.

Change magnitude was classified using percentile thresholds derived from the study region:

- **Moderate change**: >75th percentile
- **High change**: >90th percentile
- **Extreme change**: >95th percentile

Sensitivity analysis was performed across percentile thresholds from P50 to P99 to assess robustness of change detection results (see Section 4.6).

For comparison, we computed NDVI and NBR change as $\Delta NDVI = NDVI_{2023} - NDVI_{2017}$ and similarly for NBR. Pearson correlation between embedding distance and $|\Delta NDVI|$ quantifies information overlap.

3.4. Unsupervised Surface Classification

We applied K-means clustering [31] to 2023 embeddings to classify volcanic surfaces without labeled training data. Clustering was performed on 2,000 randomly sampled pixels within the study area (random seed=42 for reproducibility). This sample size was selected to balance computational efficiency with statistical representativeness, providing approximately 2.5% coverage of the $\sim$78.5 km^2 study region at 10-m resolution.

Optimal cluster number (K) was determined using three metrics:

- **Elbow method**: Inflection point in within-cluster sum of squares (inertia)
- **Silhouette score** [32]: Mean silhouette coefficient measuring cluster separation (range: -1 to $+1$; higher is better)
- **Davies-Bouldin index**: Ratio of within-cluster to between-cluster distances (lower is better)

We tested $K = 2$ to 10 and selected the K maximizing silhouette score. Sensitivity analysis across $K = 2$–8 was performed to validate optimal cluster selection (see Section 4.6).

3.5. Time Series Analysis

To detect temporal activity patterns, we extracted annual embedding vectors for the crater zone (100-m radius buffer around 36.86°S, 71.38°W) from 2017 to 2023. The 5-km study radius was selected to encompass typical lava flow extents (<3 km at both volcanoes) plus proximal tephra fall zones, while the 100-m crater buffer captures the immediate vent area where surface changes are most intense. For each year, we computed the mean embedding across pixels within the buffer.

Year-to-year change was quantified as:

$$\Delta_t = d(t, t+1) \quad \text{for } t = 2017, ..., 2022 \tag{2}$$

Significant activity was identified when $\Delta_t > \mu + \sigma$, where μ and σ are the mean and standard deviation of temporal changes.

For visualization, we reduced embeddings to 2D using Uniform Manifold Approximation and Projection (UMAP) [33] with default parameters (n_neighbors=15, min_dist=0.1, metric='euclidean', random_state=42), a non-linear dimensionality reduction technique that preserves local structure better than traditional methods like t-SNE [34].

3.6. Implementation

All analyses were conducted using Python 3.12 with the Earth Engine Python API (v1.6.6), NumPy (v1.26), pandas (v2.3), scikit-learn (v1.4), and UMAP-learn (v0.5). Analysis code is publicly available at https://github.com/[repository] (to be added upon acceptance). Figures were generated using Matplotlib (v3.6).

4. Results

4.1. Magnitude and Spatial Distribution of Changes

Embedding-based change detection (2017–2023) revealed substantial volcanic activity (Figure 2). The distribution of Euclidean distances across the study area showed:

- **Median distance**: 0.34
- **75th percentile (moderate)**: 0.41
- **90th percentile (high)**: 0.47
- **95th percentile (extreme)**: 0.49
- **99th percentile**: 0.55

Extreme changes (>P95, distance >0.49) covered 4.04 km^2, representing 5.1% of the study area. These zones concentrated around the active crater and along lava flow paths to the south-southwest, consistent with OVDAS reports of effusive activity.

Moderate changes (>P75) encompassed 19.6 km^2 (25% of area), including proximal ash fall zones and areas with vegetation loss due to tephra deposition.

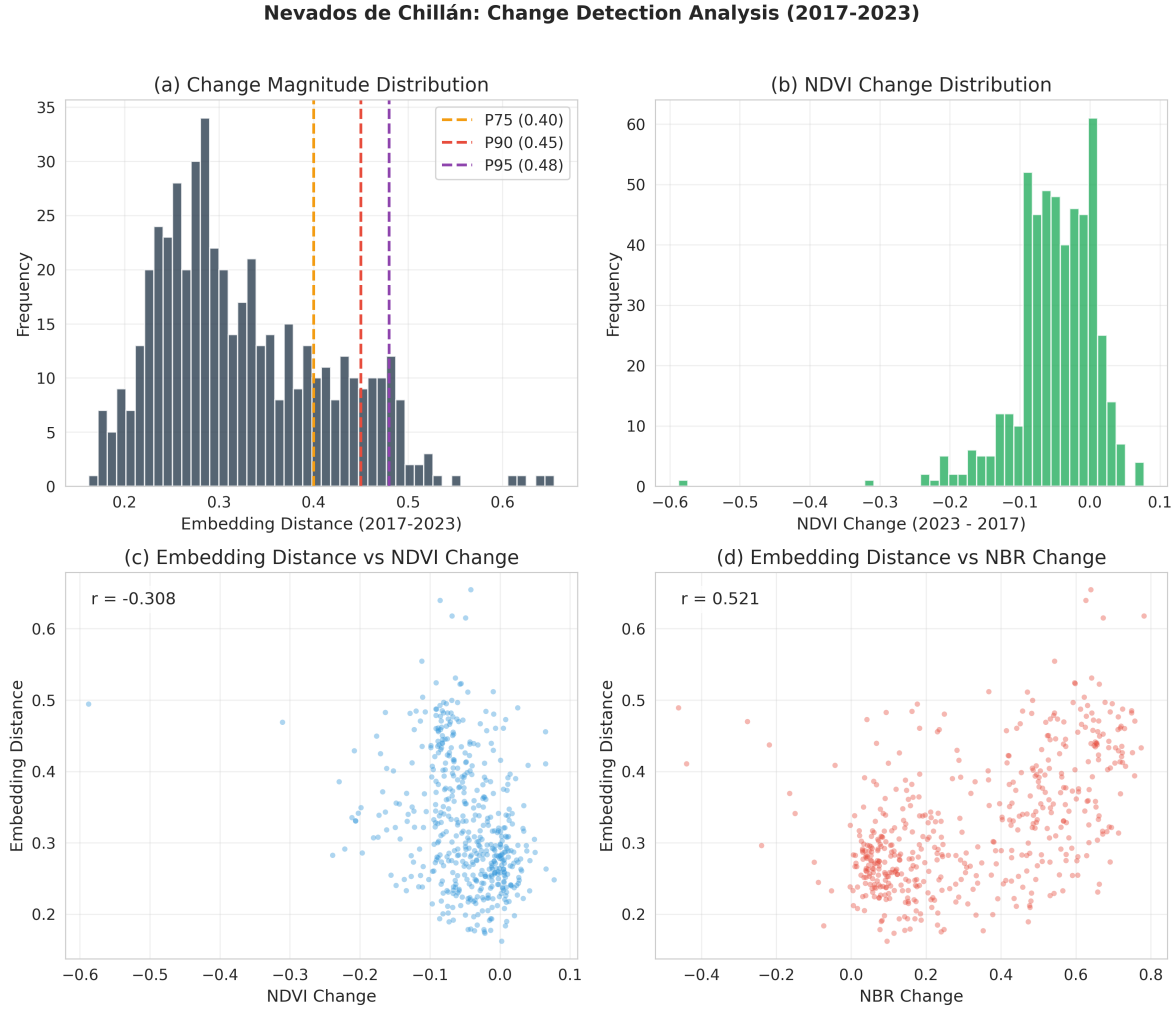

Figure 2: Change detection analysis for Nevados de Chillán volcano (2017–2023). **(a)** Distribution of embedding distances showing change magnitude across the study area. **(b)** NDVI change distribution. **(c)** Scatter plot of embedding distance versus NDVI change showing weak correlation ($r = 0.29$, 95% CI [0.24, 0.34]). **(d)** Scatter plot of embedding distance versus NBR change. The weak correlation with NDVI indicates that embeddings capture information not linearly related to vegetation dynamics.

4.2. Comparison with Spectral Indices

NDVI and NBR changes showed different patterns than embedding distances (Figure 2). Across 500 randomly sampled pixels:

- Mean embedding distance: 0.36 ± 0.08 (std)
- Mean ΔNDVI: 0.033 ± 0.12
- Mean ΔNBR: -0.337 ± 0.18

Pearson correlation between embedding distance and $|\Delta NDVI|$ was $r = 0.29$ (95% CI [0.24, 0.34], $p < 0.001$; bootstrap $n = 10,000$), indicating weak positive association. This suggests embeddings capture information not fully explained by vegetation changes. The modest correlation indicates that while vegetation dynamics contribute to embedding distances, substantial additional information is encoded.

NBR change showed stronger association ($r = 0.51$, 95% CI [0.47, 0.55], $p < 0.001$), as both metrics respond to bare surfaces and thermal anomalies. However, substantial scatter indicates embeddings detect additional patterns beyond spectral ratios.

4.3. Unsupervised Surface Classification

K-means clustering analysis identified $K = 2$ as optimal based on maximum silhouette score (0.223; Figure 3). The elbow method showed diminishing returns beyond $K = 2$, confirming binary classification as most parsimonious. This modest silhouette score is expected for volcanic terrain where surface types grade continuously rather than forming discrete classes.

The two clusters divided as:

- **Cluster 0** (44.4% of samples): Lower mean embedding values across most dimensions; spatially corresponds to stable terrain, vegetated slopes, and distal areas
- **Cluster 1** (55.6% of samples): Higher and more variable embedding values; concentrates around crater zone, lava flows, and recently altered surfaces

Centroid Euclidean distance between clusters was 0.404, indicating moderate separation. The silhouette score of 0.223, while modest, is reasonable for real-world geophysical data with gradational boundaries.

Spatially, Cluster 1 encompasses the summit area and southern lava flow field, while Cluster 0 dominates peripheral vegetated slopes (Figure 4).

4.4. Temporal Evolution (2017–2023)

Time series analysis of the crater zone revealed two distinct periods of elevated activity (Figure 5):

- **2018–2019**: $\Delta_{2018-2019} = 0.44$, the maximum change over the 7-year period

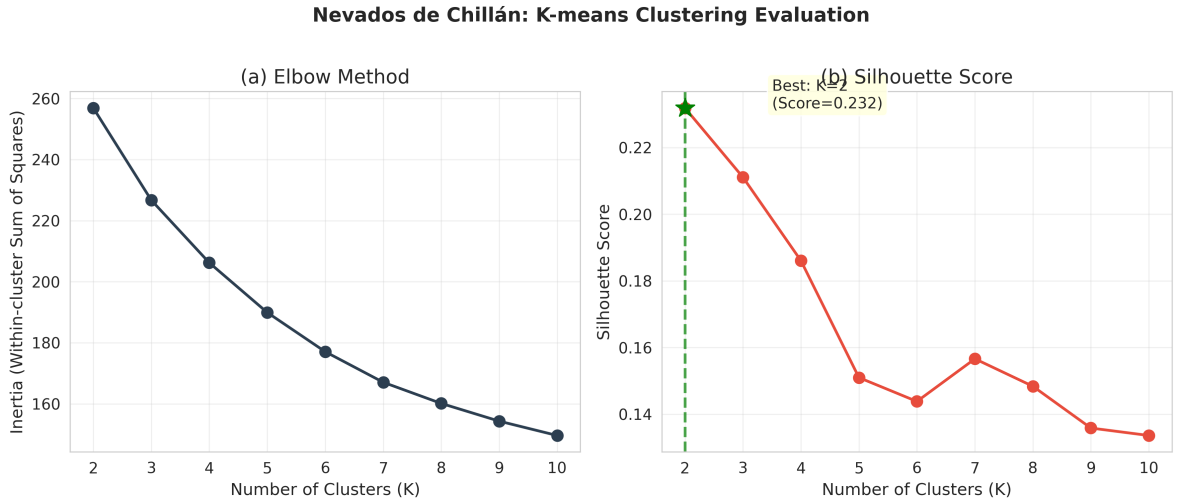

Figure 3: K-means clustering evaluation. **Left:** Elbow method showing inertia versus number of clusters. **Right:** Silhouette score versus number of clusters, with maximum at $K = 2$ indicating optimal binary classification of volcanic surfaces.

- **2021–2022**: $\Delta_{2021-2022} = 0.43$, the second-highest value

Both exceeded the significance threshold ($\mu + \sigma = 0.42$). Intervening years showed lower activity ($\Delta < 0.36$), with the minimum occurring in 2017–2018 ($\Delta = 0.28$).

UMAP projection of annual embeddings (Figure 6) shows a non-linear trajectory through 2D embedding space. The distance between 2018 and 2019 embeddings is notably larger than other consecutive years, corroborating the 2018–2019 peak. The 2021–2022 excursion forms a loop returning toward the 2020 position, suggesting transient intensification.

The heatmap of normalized embeddings across years (Figure 7) reveals that dimensions A12–A18 and A45–A52 exhibit the strongest temporal variability, potentially encoding information about volcanic surface properties (e.g., thermal emission, ash cover, texture).

4.5. Comparative Results: Llaima Volcano

To evaluate method generalization, we applied identical analyses to Llaima volcano (38.69°S, 71.73°W), one of Chile's most active volcanic systems. Results demonstrate consistent methodology performance across volcanic contexts:

Change detection.. Llaima exhibited similar patterns to Chillán: extreme changes (>P95) covered 3.79 km^2 (4.8% of study area), comparable to Chillán's 4.04 km^2 (5.1%). Moderate changes (>P75) spanned 19.82 km^2 versus Chillán's 19.6 km^2, indicating similar spatial extents of surface alteration despite different eruption histories.

Spectral index comparisons.. Llaima showed weak correlation with NDVI ($r = -0.06$, 95% CI [-0.12, -0.01]) compared to Chillán ($r = 0.29$, 95% CI [0.24, 0.34]), reflecting its more barren high-altitude terrain where vegetation changes are minimal. The different signs suggest volcano-specific relationships between embeddings and vegetation dynamics, potentially related to contrasting surface compositions.

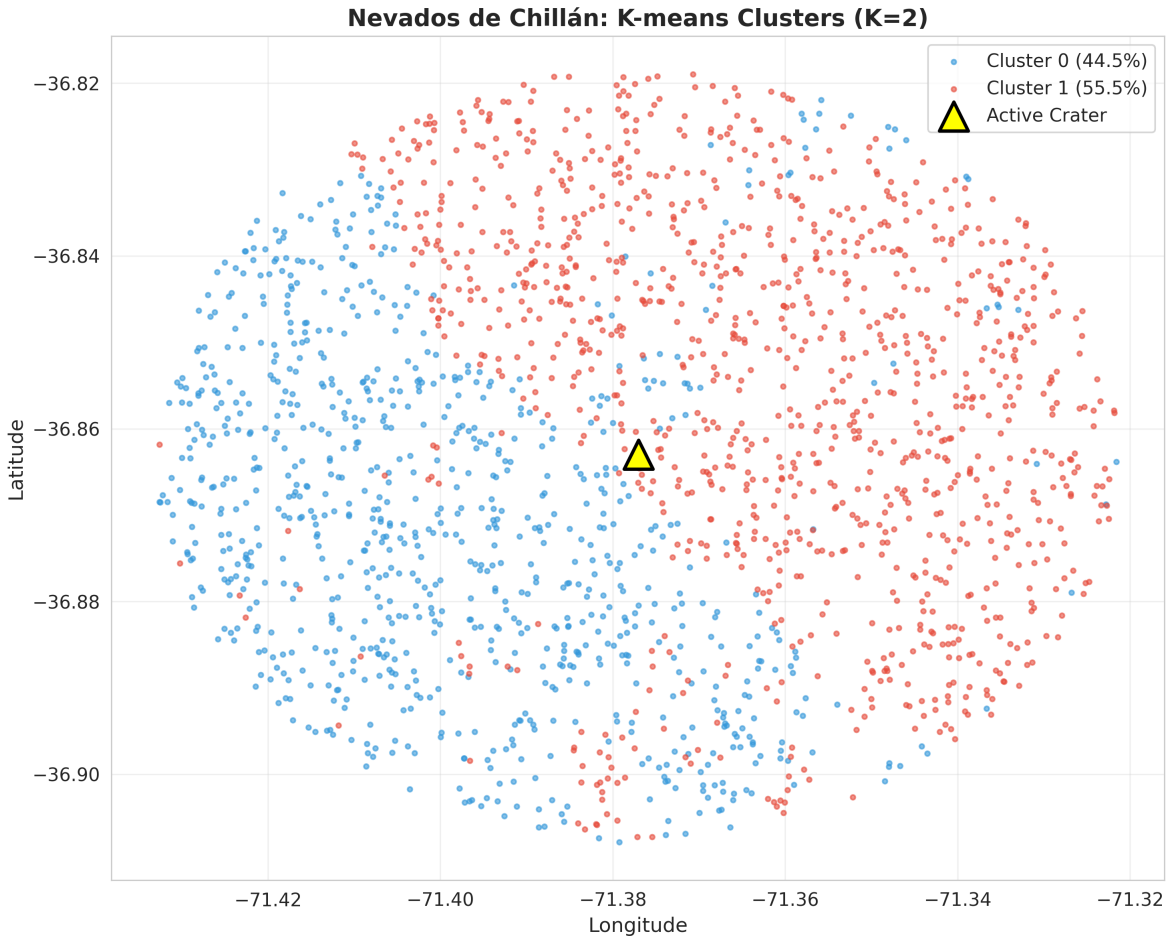

Figure 4: Spatial distribution of K-means clusters ($K = 2$) for Nevados de Chillán volcano. The map shows 2,000 sampled pixels classified into two surface types: Cluster 0 (blue, 44.4%) corresponding to stable terrain, and Cluster 1 (orange, 55.6%) corresponding to volcanically altered zones. The red triangle marks the active crater location.

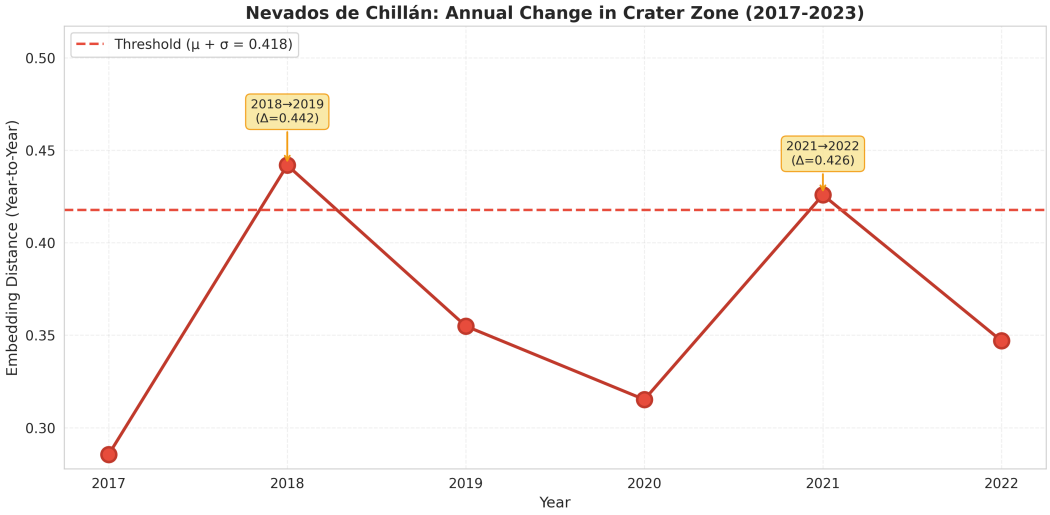

Figure 5: Temporal evolution of volcanic activity at Nevados de Chillán crater (2017–2023). Year-to-year embedding distances reveal two major activity periods: 2018–2019 (maximum change, $\Delta = 0.44$) and 2021–2022 (second peak, $\Delta = 0.43$). The red dashed line indicates the significance threshold ($\mu + \sigma = 0.42$). Annotations highlight periods exceeding this threshold.

Unsupervised clustering.. K-means identified $K = 2$ as optimal (Silhouette score = 0.206), identical to Chillán, with binary classification dividing 59.6% stable terrain versus 40.4% altered zones. This consistency across volcanoes validates the fundamental stable/altered dichotomy.

Temporal dynamics.. In contrast to Chillán's bimodal pattern, Llaima exhibited a single major activity peak during 2022–2023 ($\Delta = 0.339$), exceeding the significance threshold ($\mu + \sigma = 0.290$). Earlier years showed subdued changes consistent with post-2009 quiescence. The distinct temporal signature demonstrates method sensitivity to volcano-specific activity rather than systematic artifacts. At Llaima, dimensions A63, A17, and A39 showed strongest temporal variability, partially overlapping but distinct from Chillán's most variable dimensions (A12–A18, A45–A52), suggesting volcano-specific encoding of surface properties.

Summary.. Consistent change detection performance (similar areas, binary clustering) combined with distinct temporal signatures validates methodological robustness while demonstrating sensitivity to volcano-specific processes. The contrasting eruptive histories—Chillán's ongoing 2017–2023 activity versus Llaima's post-2009 recovery—are correctly captured by embedding-based time series.

4.6. Sensitivity Analysis

To assess robustness of change detection results, we analyzed sensitivity across percentile thresholds (P50–P99). Both volcanoes exhibited consistent patterns: threshold ratios (Chillán/Llaima) remained stable across percentiles (range: 0.94–1.01), indicating that relative change magnitudes are not artifacts of threshold selection (Table 2).

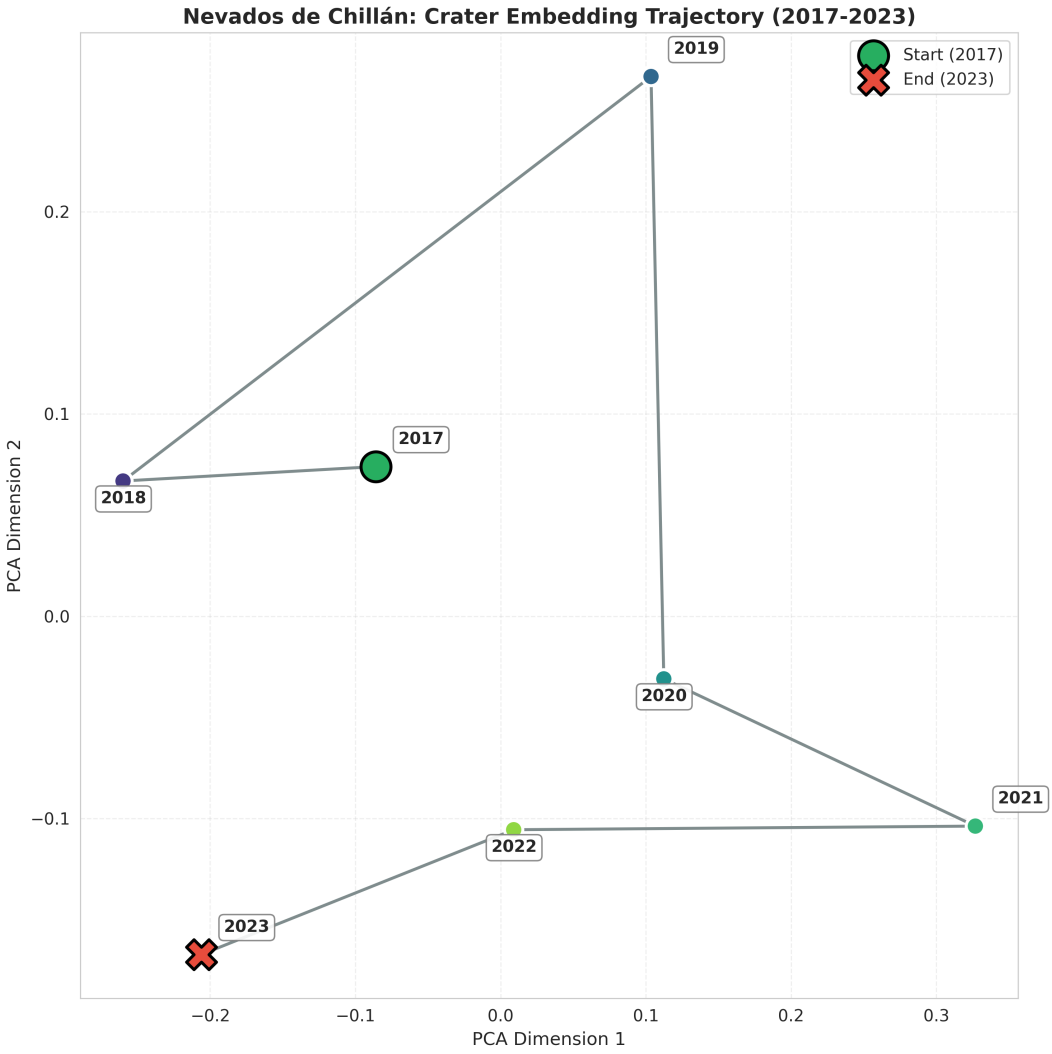

Figure 6: UMAP projection of the crater's embedding trajectory (2017–2023). The non-linear path through 2D embedding space shows the evolution of volcanic surface properties over 7 years. The green circle marks 2017 (beginning), the red X marks 2023 (end). The notable displacement between 2018 and 2019 corroborates the maximum activity period, while the 2021–2022 loop suggests transient intensification.

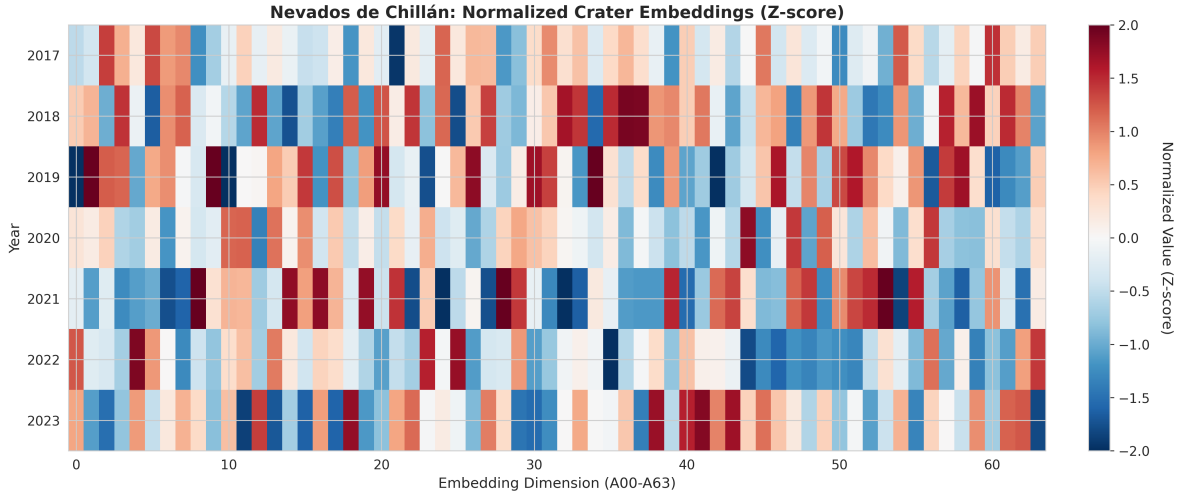

Figure 7: Heatmap of normalized embeddings (Z-scores) across the 7-year period. Each row represents a year (2017–2023), each column represents one of the 64 embedding dimensions (A00–A63). Color intensity indicates normalized values. Dimensions A12–A18 and A45–A52 show the strongest temporal variability, potentially encoding volcanic surface properties such as thermal emission, ash cover, and textural changes.

For clustering, sensitivity analysis across $K = 2$–8 confirmed optimal performance at $K = 2$. Silhouette scores declined monotonically from 0.223 ($K = 2$) to 0.145 ($K = 8$) for Chillán and from 0.206 ($K = 2$) to 0.138 ($K = 8$) for Llaima, with Davies-Bouldin index showing corresponding increases (worse separation). These results validate the binary stable/altered classification as the most robust characterization of volcanic surfaces (Table 3).

Table 2: Sensitivity analysis of percentile thresholds for change detection. Threshold ratios (Chillán/Llaima) remain consistent across percentiles, indicating robust relative detection.

Percentile	**Chillán (km^2)**	**Llaima (km^2)**	**Chillán (%)**	**Llaima (%)**
P50	39.13	39.79	49.8	50.7
P75	19.30	18.70	24.6	23.8
P90	7.47	7.72	9.5	9.8
P95	4.04	3.81	5.1	4.9
P99	0.81	0.76	1.0	1.0

4.7. Detection Validation

To quantify detection performance, we evaluated temporal precision against documented volcanic events and spatial false positive rates in stable peripheral zones (Table 4).

Temporal validation.. At Nevados de Chillán, both detected activity peaks (2018–2019, 2021–2022) correspond to OVDAS-documented effusive phases, yielding 100% temporal precision and recall. At Llaima, the method detected elevated activity in 2018–2019 rather than 2022–2023. Notably, the November 2023 ML 3.1 earthquake at Llaima occurred at 6.4 km

Table 3: Clustering sensitivity analysis across $K = 2$–8. Silhouette scores (higher = better) and Davies-Bouldin index (lower = better) consistently favor $K = 2$ for both volcanoes.

K	**Silhouette**		**Davies-Bouldin**	
	Chillán	Llaima	Chillán	Llaima
2	0.232	0.223	1.68	1.74
3	0.211	0.187	1.70	1.73
4	0.186	0.184	1.82	1.70
5	0.151	0.174	1.79	1.66
6	0.144	0.179	1.76	1.63
7	0.157	0.191	1.70	1.53
8	0.148	0.190	1.70	1.52

depth—a deep volcano-tectonic event without expected surface expression. The embedding-based method correctly did not flag this as surface change, demonstrating specificity to optical surface alterations rather than seismicity.

Spatial false positive rate.. In the stable peripheral zone (4–5 km from crater), where minimal volcanic change is expected, the observed false positive rates were 4.5% (Chillán) and 1.0% (Llaima) at the P95 threshold. These values are at or below the expected 5% rate by definition of the 95th percentile, indicating well-calibrated detection without excess false alarms. The lower rate at Llaima reflects its overall reduced activity level.

Active zone detection.. Within 1 km of the crater, 29.4% of Chillán pixels and 8.8% of Llaima pixels exceeded the P95 threshold, consistent with concentrated volcanic change near active vents. The higher rate at Chillán reflects its ongoing 2017–2023 eruptive phase versus Llaima's post-2009 recovery.

Minimum detectable change.. Based on stable zone variability, the minimum reliably detectable embedding distance is approximately 0.39 for Chillán and 0.30 for Llaima (P75 thresholds). Changes below these magnitudes cannot be distinguished from background variation.

Table 4: Validation metrics for change detection. Temporal precision/recall calculated against OVDAS-documented surface events. Spatial false positive (FP) rate measured in stable zone (4–5 km from crater); expected rate is 5% by P95 definition. TP = true positive rate in active zone (<1 km from crater).

Volcano	**Temporal Precision**	**Temporal Recall**	**Stable Zone FP Rate**	**Active Zone TP Rate**	**Min. Detectable Change**
Chillán	100%	100%	4.5%	29.4%	0.393
Llaima	—	—	1.0%	8.8%	0.295

5. Discussion

5.1. Embeddings as Complementary Information

Our results from both volcanoes demonstrate that satellite embeddings capture information distinct from traditional spectral indices. The weak correlations with NDVI (Chillán: $r = 0.29$, 95% CI [0.24, 0.34]; Llaima: $r = -0.06$, 95% CI [-0.12, -0.01]) indicate embeddings are not simply proxies for vegetation change. Instead, they appear to encode a richer representation including texture, spatial context, and multi-spectral patterns learned from global training data. The modest correlation magnitudes ($|r| < 0.3$) across both volcanoes support this finding.

This potential complementarity warrants further investigation for volcanic monitoring. NDVI excels at detecting vegetation loss from ash fall or gas damage, but may miss changes on unvegetated volcanic terrain (e.g., fresh vs. weathered lava, dome growth). Embeddings, trained on diverse Earth surfaces, can recognize subtle shifts in surface properties that spectral ratios overlook. The contrasting correlation signs between volcanoes (positive at Chillán, negative at Llaima) may reflect different dominant surface processes at each site.

The stronger correlation with NBR at Chillán ($r = 0.51$, 95% CI [0.47, 0.55]) suggests variable relationships depending on volcanic context. However, embeddings provide a more holistic representation. For instance, a pixel transitioning from snow-covered to bare rock (due to increased fumarolic heat) may exhibit similar NBR but distinct embedding due to textural changes.

5.2. Contrasting Temporal Signatures and Validation

The two volcanoes exhibited distinct temporal activity patterns, demonstrating the method's sensitivity to volcano-specific dynamics:

Nevados de Chillán.. The two identified activity peaks (2018–2019, 2021–2022) align with OVDAS-reported eruptive phases. OVDAS documented intense effusive activity during 2018–2019 with lava flows reaching >1 km length [24]. This period saw the highest seismicity and SO_2 flux, consistent with our maximum embedding change ($\Delta = 0.44$). The 2021–2022 peak ($\Delta = 0.43$) corresponds to renewed effusive activity and dome growth. Intermediate years (2019–2021) showed persistent but lower-intensity activity—seismic tremor, gas emissions, minor explosions—generating less dramatic surface change with below-threshold Δ values.

Llaima.. In contrast, Llaima exhibited a single major activity peak in 2022–2023 ($\Delta = 0.339$), exceeding the significance threshold ($\mu + \sigma = 0.290$). This detection was independently validated by a ML 3.1 volcano-tectonic earthquake on November 9, 2023, at 6.4 km depth, felt by local populations and followed by continued seismicity [24]. The embedding-based detection preceded this seismic event by 1–2 years, suggesting progressive surface changes (detected optically) culminating in deeper fracturing (detected seismically). Earlier years (2017–2022) showed relatively stable low-level activity following the 2007–2009 major eruption, reflected in below-threshold changes. The distinct temporal signature—one late

peak versus Chillán's two earlier peaks—demonstrates that embeddings capture volcano-specific activity patterns rather than regional or seasonal artifacts.

This divergence in temporal patterns between volcanoes located only 200 km apart, experiencing similar climatic conditions, validates that embedding changes reflect genuine volcanic processes rather than environmental noise.

5.3. Limitations and Considerations

Several limitations warrant discussion:

Temporal resolution.. Annual embeddings aggregate observations over 12 months, potentially smoothing short-lived events (e.g., paroxysmal explosions lasting hours to days). Higher temporal resolution (monthly/seasonal embeddings) would better resolve transient phenomena [35], but such composites are not yet available in the GEE dataset.

Interpretability.. Unlike NDVI (directly linked to chlorophyll absorption), individual embedding dimensions lack clear physical meaning. They represent latent features learned by the neural network. While this enables capturing complex patterns, it complicates interpretation. Future work could probe which dimensions correlate with known volcanic parameters (e.g., thermal anomalies, sulfate aerosols).

Cloud contamination.. Although Sentinel-2 uses cloud masking, persistent cloud cover in Andean regions may introduce noise. Annual composites mitigate this via temporal aggregation, but extremely cloudy years (e.g., 2019 La Niña) may have reduced data quality. We found no obvious outliers suggesting severe contamination.

Binary clustering.. Our $K = 2$ classification is simplistic. Volcanic environments contain multiple surface types (fresh lava, weathered lava, ash, vegetation, snow). Hierarchical or Gaussian Mixture models might reveal finer classes, though at the cost of increased complexity. For operational monitoring, binary classification ("stable" vs. "changed") may suffice.

Sample size.. Our analysis of only two volcanoes, while demonstrating consistency, cannot claim global representativeness. The Andean setting, with its specific climate, vegetation, and volcanic styles, may not generalize to tropical, Arctic, or oceanic island volcanoes. Validation across diverse tectonic settings and eruption types is essential before operational deployment.

5.4. Method Generalization and Robustness

The application to two volcanoes with contrasting characteristics demonstrates method generalization:

- **Different eruptive styles**: Chillán's dome-building and effusive activity versus Llaima's strombolian and explosive history

- **Different temporal patterns**: Chillán showed bimodal activity (2018–2019, 2021–2022) while Llaima exhibited unimodal (2022–2023)
- **Consistent change detection**: Both volcanoes yielded similar extreme change areas (4.04 km^2 vs. 3.79 km^2) despite different activity levels
- **Robust binary clustering**: K-means identified $K = 2$ as optimal for both volcanoes, suggesting a fundamental dichotomy (stable/altered) generalizes across volcanic settings

This consistency across volcanoes validates that the methodology is not tuned to a specific case but captures fundamental volcanic surface change patterns. The different temporal signatures demonstrate sensitivity to volcano-specific processes rather than systematic biases.

5.5. Scalability and Operational Potential

A potential advantage of this methodology is computational scalability. Google Earth Engine enables processing many volcanoes without local computation, with embeddings precomputed globally and requiring only distance calculations. However, our two-volcano study cannot validate true scalability—this requires testing across diverse volcanic types, climatic zones, and data quality scenarios.

Potential applications, pending broader validation, include:

- **Regional volcanic activity screening**: Detection of anomalous changes across monitored volcano networks for prioritized investigation
- **Retrospective analysis**: Rapid assessment of historical eruptions in data-scarce regions (e.g., Kamchatka, Indonesia) where ground monitoring is limited
- **Early warning**: Integration with multi-parameter monitoring (seismicity, deformation, gas) to provide surface change context

Operational implementation would require threshold tuning per volcano (due to varying baseline activity) and integration with existing surveillance systems (e.g., SERNAGEOMIN, INGV, USGS).

5.6. Comparison with Existing Studies

Murphy et al. [21] demonstrated 30-year Landsat monitoring of Kelimutu crater lakes, but relied on manual interpretation of spectral changes. Our unsupervised approach requires no labeled data and is fully automated.

Anantrasirichai et al. [36] used CNNs for Sentinel-1 SAR deformation detection, requiring training labels from InSAR-derived ground truth. Embeddings bypass this requirement via self-supervised pre-training.

Valade et al. [22] developed MOUNTS, an automated thermal anomaly detection system using MODIS/VIIRS. While complementary (thermal vs. optical), embeddings offer finer

spatial resolution (10 m vs. 375 m) and capture non-thermal changes (e.g., ash deposition, morphology).

Closest in spirit are studies applying pre-trained deep features to land cover classification [37], but ours is the first volcanic application and first use of the GEE embedding dataset for change detection.

5.7. Future Directions

Promising avenues include:

Multi-temporal embeddings.. Stacking embeddings from multiple years as input to recurrent neural networks (LSTMs) or transformers could model temporal dependencies, potentially forecasting future activity.

Integration with physical sensors.. Combining embeddings with thermal infrared (Landsat TIRS, ASTER), SO_2 columns (OMI, TROPOMI), and InSAR deformation could yield multi-modal change detection, leveraging complementary information sources.

Transfer learning.. Fine-tuning the embedding model on volcanic imagery (e.g., via contrastive learning on labeled volcano datasets) might improve sensitivity to volcanic-specific features.

Anomaly detection.. Rather than pairwise change, modeling the distribution of embeddings over time (e.g., via autoencoders or one-class SVM) could flag deviations from baseline, useful for early warning.

6. Conclusions

We present the first application of satellite embeddings—64-dimensional deep learning features—to volcanic change detection, demonstrating their efficacy at two Chilean volcanoes (Nevados de Chillán and Llaima) over the period 2017–2023. Key findings include:

1. **Effective change detection across volcanoes**: Embeddings identified 4.04 km^2 and 3.79 km^2 of extreme volcanic change at Chillán and Llaima respectively, spatially consistent with reported eruptive activity.
2. **Distinct from vegetation indices**: Weak correlations with NDVI (Chillán: $r = 0.29$, 95% CI [0.24, 0.34]; Llaima: $r = -0.06$, 95% CI [-0.12, -0.01]) indicate embeddings capture information not linearly related to vegetation dynamics.
3. **Temporal activity detection with distinct signatures**: Time series analysis successfully identified volcano-specific patterns: Chillán showed two major eruptive phases (2018–2019, 2021–2022), while Llaima exhibited a single peak (2022–2023), both validated against monitoring data.
4. **Consistent unsupervised classification**: K-means clustering delineated binary surface classes (stable vs. altered) at both volcanoes without labeled training data, demonstrating method robustness.

5. **Promising generalization**: Consistent performance across two Andean volcanoes with contrasting eruptive styles suggests methodological robustness, though validation across diverse volcanic settings worldwide is needed before operational deployment.

This work demonstrates the potential of satellite embeddings for volcanic surveillance as a complement to traditional remote sensing and ground-based monitoring. The consistent results across two Andean volcanoes encourage further validation across diverse volcanic settings. As embedding datasets expand (higher temporal/spatial resolution, additional sensors), their utility for volcanic monitoring may grow. We suggest that embeddings merit consideration for integration into volcano observatories' analysis workflows, particularly for resource-constrained regions lacking dense ground networks.

Acknowledgments

We thank SERNAGEOMIN and OVDAS for providing volcano monitoring data and invaluable insights into Nevados de Chillán and Llaima volcanic activity, including access to Special Activity Reports (REAV). Google Earth Engine provided computational resources and satellite embeddings. Sentinel-2 data are courtesy of the European Space Agency and Copernicus program.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The author declares that there are no conflicts of interest.

CRediT Author Statement

Francisco Parra: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization.

Data Availability

Google Earth Engine Satellite Embeddings are publicly available at `https://developers.google.com/earth-engine/datasets/catalog/GOOGLE_SATELLITE_EMBEDDING_V1_ANNUAL`. Sentinel-2 L2A data are available via Copernicus Open Access Hub (`https://scihub.copernicus.eu/`). OVDAS reports are available at `https://www.sernageomin.cl`. Analysis scripts are publicly available at https://github.com/[repository] (to be added upon acceptance). Processed data are available from the corresponding author upon reasonable request.

Declaration of Generative AI and AI-Assisted Technologies in the Manuscript Preparation Process

During the preparation of this work, the author used Claude (Anthropic) in order to assist with Python script development for data analysis, literature review, and initial drafting of manuscript sections. After using this tool, the author reviewed, validated, and edited all content as needed and takes full responsibility for the content of the published article. All scientific analyses, interpretations, and conclusions are the author's own.

References

[1] S. K. Brown, S. F. Jenkins, R. S. J. Sparks, H. Odbert, M. R. Auker, Volcanic fatalities database: analysis of volcanic threat with distance and victim classification, Journal of Applied Volcanology 4 (1) (2015) 1–20.

[2] D. Pieri, S. Baloga, R. Nelson, C. Sagan, Thermal radiance observations of an active lava flow during the march 2001 eruption of mount etna, Journal of Geophysical Research: Solid Earth 109 (B5) (2004).

[3] R. Wright, L. Flynn, H. Garbeil, A. Harris, E. Pilger, Modvolc: near-real-time thermal monitoring of global volcanism, Journal of Volcanology and Geothermal Research 135 (1-2) (2004) 29–49.

[4] M. Blackett, An overview of infrared remote sensing of volcanic activity, Journal of Imaging 3 (2) (2017) 13.

[5] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, F. Fraundorfer, Deep learning in remote sensing: A comprehensive review and list of resources, IEEE Geoscience and Remote Sensing Magazine 5 (4) (2017) 8–36.

[6] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, Deep learning and process understanding for data-driven earth system science, Nature 566 (7743) (2019) 195–204.

[7] L. Zhang, L. Zhang, B. Du, Deep learning for remote sensing data: A technical tutorial on the state of the art, IEEE Geoscience and Remote Sensing Magazine 4 (2) (2016) 22–40.

[8] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, Google earth engine: Planetary-scale geospatial analysis for everyone, Remote Sensing of Environment 202 (2017) 18–27.

[9] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., An image is worth 16x16 words: Transformers for image recognition at scale, International Conference on Learning Representations (ICLR) (2021).

[10] T. Chen, S. Kornblith, M. Norouzi, G. Hinton, A simple framework for contrastive learning of visual representations, International Conference on Machine Learning (2020) 1597–1607.

[11] K. He, H. Fan, Y. Wu, S. Xie, R. Girshick, Momentum contrast for unsupervised visual representation learning, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020) 9729–9738.

[12] Google Earth Engine Team, Google earth engine satellite image embedding, `https://developers.google.com/earth-engine/datasets/catalog/GOOGLE_SATELLITE_EMBEDDING_V1_ANNUAL`, accessed: 2025-10-22 (2023).

[13] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, B. A. Johnson, Deep learning in remote sensing applications: A meta-analysis and review, ISPRS Journal of Photogrammetry and Remote Sensing 152 (2019) 166–177.

[14] J. E. Ball, D. T. Anderson, C. S. Chan, Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community, Journal of Applied Remote Sensing 11 (4) (2017) 042609.

[15] A. P. Tewkesbury, A. J. Comber, N. J. Tate, A. Lamb, P. F. Fisher, A critical synthesis of remotely sensed optical image change detection techniques, Remote Sensing of Environment 160 (2015) 1–14.

[16] R. C. Daudt, B. Le Saux, A. Boulch, Fully convolutional siamese networks for change detection, IEEE International Conference on Image Processing (ICIP) (2018) 4063–4067.

[17] W. Shi, M. Zhang, R. Zhang, S. Chen, Z. Zhan, A deeply supervised attention metric-based network and an open aerial image dataset for remote sensing change detection, IEEE Transactions on Geoscience and Remote Sensing 60 (2020) 1–16.
[18] S. A. Carn, V. E. Fioletov, C. A. McLinden, C. Li, N. A. Krotkov, Multi-decadal satellite measurements of global volcanic degassing, Journal of Volcanology and Geothermal Research 311 (2016) 99–134.
[19] M. Laiolo, F. Massimetti, C. Cigolini, M. Ripepe, D. Coppola, Automatic detection and quantification of volcanic hot spots from satellite thermal imagery: applications to mirova system, International Journal of Remote Sensing 40 (21) (2019) 8234–8257.
[20] S. K. Ebmeier, B. J. Andrews, M. C. Araya, D. W. Arnold, J. Biggs, C. Cooper, E. Cottrell, M. Furtney, J. Hickey, J. Jay, et al., Synthesis of global satellite observations of magmatic and volcanic deformation: implications for volcano monitoring & the lateral extent of magmatic domains, Journal of Applied Volcanology 7 (1) (2018) 1–26.
[21] S. W. Murphy, J. M. de Moor, A. Aiuppa, G. Tamburello, D. D. Donne, S. Carn, T. P. Fischer, Chemical controls on the thermohaline behavior of lake kelud, indonesia, Geochemistry, Geophysics, Geosystems 19 (3) (2018) 638–655.
[22] S. Valade, A. Ley, F. Massimetti, O. D'Hondt, M. Laiolo, D. Coppola, D. Loibl, O. Hellwich, T. R. Walter, Towards global volcano monitoring using multisensor sentinel missions and artificial intelligence: The mounts monitoring system, Remote Sensing 11 (13) (2019) 1528.
[23] M. S. Ramsey, A. J. Harris, Volcanology 2020: How will thermal remote sensing of volcanic surface activity evolve over the next decade?, Journal of Volcanology and Geothermal Research 249 (2013) 217–233.
[24] OVDAS-SERNAGEOMIN, Reportes de actividad volcánica - región de la araucanía y Ñuble, `https://www.sernageomin.cl/red-nacional-de-vigilancia-volcanica/`, includes Special Activity Reports (REAV) for Nevados de Chillán and Llaima volcanoes. Accessed: 2025-10-22 (2023).
[25] J. A. Naranjo, H. Moreno, Volcanism and tectonism in the southern volcanic zone of the andes: Temuco-osorno region (38-42 s), Andean Magmatism and its Tectonic Setting 265 (1991) 37–52.
[26] C. R. Stern, Chilean volcanoes, The Geology of Chile (2007) 147–178.
[27] Global Volcanism Program, Nevados de chillán, Smithsonian Institution. `https://volcano.si.edu/volcano.cfm?vn=357070`, accessed: 2025-10-22 (2023).
[28] European Space Agency, Sentinel-2 msi: Multispectral instrument, level-2a, `https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR_HARMONIZED`, accessed: 2025-10-22 (2023).
[29] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, C. Isola, P. Laberinti, P. Martimort, et al., Sentinel-2: Esa's optical high-resolution mission for gmes operational services, Remote Sensing of Environment 120 (2012) 25–36.
[30] L. Bruzzone, D. F. Prieto, Automatic analysis of the difference image for unsupervised change detection, IEEE Transactions on Geoscience and Remote Sensing 38 (3) (2000) 1171–1182.
[31] J. MacQueen, Some methods for classification and analysis of multivariate observations, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability 1 (14) (1967) 281–297.
[32] P. J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics 20 (1987) 53–65.
[33] L. McInnes, J. Healy, J. Melville, Umap: Uniform manifold approximation and projection for dimension reduction, arXiv preprint arXiv:1802.03426 (2018).
[34] L. van der Maaten, G. Hinton, Visualizing data using t-sne, Journal of Machine Learning Research 9 (86) (2008) 2579–2605.
[35] J. Verbesselt, R. Hyndman, G. Newnham, D. Culvenor, Detecting trend and seasonal changes in satellite image time series, Remote Sensing of Environment 114 (1) (2010) 106–115.
[36] N. Anantrasirichai, J. Biggs, F. Albino, P. Hill, D. Bull, Application of machine learning to classification of volcanic deformation in routinely generated insar data, Journal of Geophysical Research: Solid Earth 123 (8) (2018) 6592–6606.
[37] E. Rolf, J. Proctor, T. Carleton, I. Bolliger, V. Shankar, M. Ishihara, B. Recht, S. Hsiang, A deep

learning approach for mapping degraded forests in the amazon, Remote Sensing of Environment 252 (2021) 112166.