# Deep Learning for Subtle Volcanic Deformation Detection With InSAR Data in Central Volcanic Zone

Teo Beker, *Member, IEEE*, Homa Ansari, Sina Montazeri, Qian Song, *Member, IEEE*, and Xiao Xiang Zhu, *Fellow, IEEE*

***Abstract*— Subtle volcanic deformations point to volcanic activities, and monitoring them helps predict eruptions. Today, it is possible to remotely detect volcanic deformation in mm/year scale thanks to advances in interferometric synthetic aperture radar (InSAR). This article proposes a framework based on a deep learning model to automatically discriminate subtle volcanic deformations from other deformation types in five-year-long InSAR stacks. Models are trained on a synthetic training set. To better understand and improve the models, explainable artificial intelligence (AI) analyses are performed. In initial models, Gradient-weighted Class Activation Mapping (Grad-CAM) linked new-found patterns of slope processes and salt lake deformations to false-positive detections. The models are then improved by fine-tuning (FT) with a hybrid synthetic-real data, and additional performance is extracted by low-pass spatial filtering (LSF) of the real test set. The t-distributed stochastic neighbor embedding (t-SNE) latent feature visualization confirmed the similarity and shortcomings of the FT set, highlighting the problem of elevation components in residual tropospheric noise. After fine-tuning, all the volcanic deformations are detected, including the smallest one, Lazufre, deforming 5 mm/year. The first time confirmed deformation of Cerro El Condor is observed, deforming 9.9–17.5 mm/year. Finally, sensitivity analysis uncovered the model's minimal detectable deformation of 2 mm/year.**

***Index Terms*— Deep learning (DL), interferometric synthetic aperture radar (InSAR), minimal deformation analysis, volcanic deformation simulation, volcanic deformation.**

## I. Introduction

AROUND 1350 volcanoes are active today; there are 50–90 eruptions per year, and 40–50 volcanoes continue eruptions at any given time [1]. Volcanic deformation is statistically linked to eruptions [2] and is a dominantly preeruptive sign of volcanic activity [3]. Hence, timely detection of volcanic preeruptive activities is crucial for understanding and characterizing the tectonic processes [4], [5] and predicting the approximate time of an eruption [6]. Volcanoes have been monitored by ground monitoring stations, which measure seismicity patterns, surface strain, and microgravity changes. Traditionally, the areal change has been measured using multiple measuring stations or wireless sensor networks [7], [8]. Nowadays, the areal deformation can be measured accurately by a remote sensing technique, interferometric synthetic aperture radar (InSAR) [9], [10], [11], substituting the in situ measurements.

The deformations in a satellite's line of sight (LOS) during an observed period can be derived from interferograms. An interferogram is a phase difference map of two SAR images acquired on the same track at two time points [12]. Accordingly, interferogram time series [13], consisting of multiple interferograms, is further processed to estimate the long-term deformation of the land surfaces through multipass SAR Interferometry techniques, such as persistent scatterer interferometry (PSI) [14], [15], [16], distributed scatterer interferometry (DSI) [15], [17], [18], or differential SAR tomography [19], [20]. However, analyzing big InSAR data is time-consuming and demands expertise. Therefore, automatic information extraction from big InSAR data is critical for large-scale volcano monitoring and timely decision-making.

Deep learning (DL), which usually uses convolutional layers to extract spatial features from input data, has been widely applied to SAR and InSAR in general [21] and specifically to detect volcanic deformations in interferograms [22], [23], [24] and filter out the noise from long-term time series [25], [26], [27]. However, automatic subtle volcanic deformations' detection at mm/year level using DL is still underexplored. Besides, previous research generally focused on improving the DL structure and training process, which is the model-centric method [22], [24], [25], [26], [28], [29]. However, our discovery reveals that the main challenges remain related more to the data than the model: the volcanic deformations are swallowed by residual atmospheric noise; real data (RD) are scarce and imbalanced with significantly fewer volcanic deformation samples; and nonvolcanic deformations that accumulated over the observed period are prone to be confused with volcanic deformations by the DL models. In this regard, model-centric

Manuscript received 26 March 2023; revised 9 July 2023; accepted 18 August 2023. Date of publication 24 October 2023; date of current version 26 October 2023. The work of Qian Song and Xiao Xiang Zhu was supported in part by the German Federal Ministry for Economic Affairs and Climate Action in the framework of the National Center of Excellence ML4Earth under Grant 50EE2201C and in part by the Munich Center for Machine Learning (MCML), Munich, Germany. *(Corresponding author: Xiao Xiang Zhu.)*

Teo Beker is with the Chair of Data Science in Earth Observation (SiPEO), Technical University of Munich (TUM), 80333 Munich, Germany, and also with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Weßling, Germany (e-mail: Teo.Beker@dlr.de).

Homa Ansari and Sina Montazeri are with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), 82234 Weßling, Germany.

Qian Song is with the Chair of Data Science in Earth Observation (SiPEO), Technical University of Munich (TUM), 80333 Munich, Germany (e-mail: Qian.Song@tum.de).

Xiao Xiang Zhu is with the Chair of Data Science in Earth Observation (SiPEO), Technical University of Munich (TUM), 80333 Munich, Germany, and also with the Munich Center for Machine Learning, Munich, Germany (e-mail: XiaoXiang.Zhu@tum.de).

Digital Object Identifier 10.1109/TGRS.2023.3318469

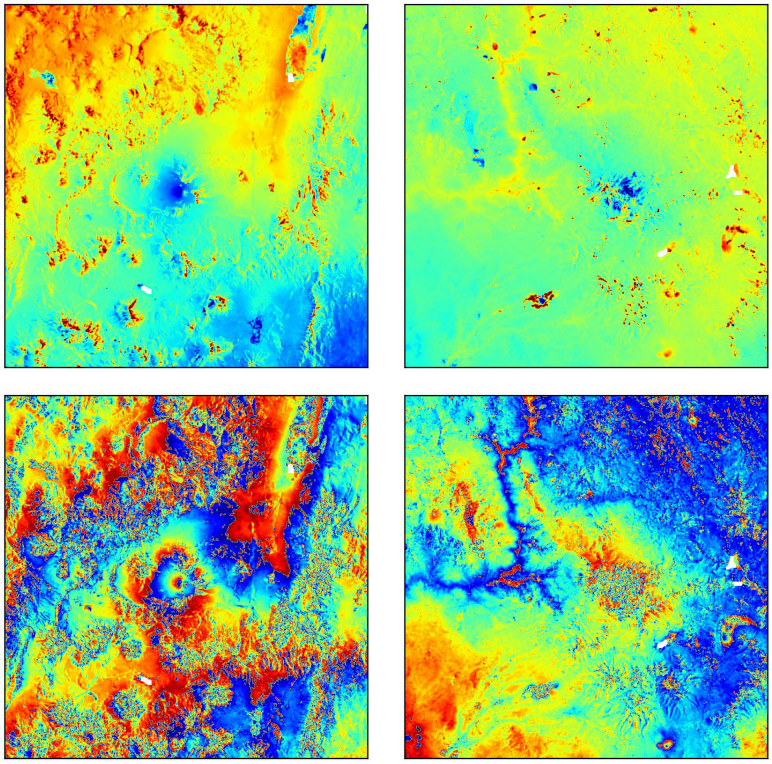


Fig. 1. Example of (left) volcanic and (right) nonvolcanic deformations in (top) deformation maps and (bottom) rewrapped data. The volcanic deformation is shown as a round conical or convex hill in the middle of the image, which is swallowed by other nonvolcanic deformation patterns.

methods bring limited gains, according to our experiment results. Therefore, we apply a data-centric approach, which improves the data used to train the model to increase detection accuracy.

In this article, we proposed a CNN framework for subtle long-term volcanic deformations' detection in stacked InSAR deformation maps. Deformation maps' derivation removes most of the noise in the time series, increases the signal-to-noise ratio (SNR), and compresses the interferogram time series into a single map [30], [31], [32], [33], thus significantly reducing the amount of data. A simulated dataset is used as the training set to train the model, and all RDs are reserved as the test set. Subsequently, analysis using Gradient-weighted Class Activation Mapping (Grad-CAM) [34] uncovered the influence of nonvolcanic deformations. To rectify the model, a data-centric approach is employed. We fine-tune the initial model with a hybrid set of simulated volcanic deformations and the real nonvolcanic deformations patterns. Besides, to make test data more similar to the synthetic data, we used a low-pass spatial filtering (LSF) to remove small patterns in the real deformation maps. Furthermore, the similarity of fine-tuning (FT) and test sets is presented with t-distributed stochastic neighbor embedding (t-SNE) [35] by visualizing the extracted features [36], [37]. Finally, an analysis is performed to determine the model's sensitivity to the volcanic deformation scale.

In this work, we present the following contributions.

1) We generated a large-scale, long-term subtle volcanic deformation detection dataset covering 46 volcanoes in the Atacama desert. The collected Sentinel-1 SAR image time series are preprocessed to allow millimeter-per-year scale deformation detection. Besides, a synthetic set that consists of 330 848 deformation maps is simulated.
2) A subtle volcanic deformation detection pipeline based on DL models is proposed in this article. We compared several DL models and devised a data-centric approach to improve detection accuracy and reduce false positives (FPs).
3) We utilized two explainable artificial intelligence (AI) tools for analyzing the results. First, Grad-CAM is used to visualize the model's activation, where we found that the models confuse salt lakes and slope-induced signals with volcanic deformation signals. Besides, the t-SNE dimension reduction tool is applied to compare features extracted from FT and test sets.
4) We propose a sensitivity analysis to expose the model to an extended range of volcanic deformations and determine the most subtle detectable volcanic deformation by the model.

## II. Related Work

### A. Challenges in Volcanic Deformations' Detection

Compared with common target detection problems, volcanic deformation detection is a challenging task. Usually, volcanic deformations are not observable from high-resolution space-borne optical data. In InSAR data, the volcanic deformation patterns often appear as protruding and/or subsiding round or eight-shaped objects in the unwrapped deformation map, as seen in Fig. 1. However, the shape of these patterns is often modified by the topography and varies from volcano to volcano.

The volcanic deformation detection task also suffers from data shortage. On the one hand, volcanic activities are severely rare, unlike other types of changes. There are only less than

TABLE I

COMPARISON OF DIFFERENT DL DEFORMING VOLCANO DETECTION METHODS AND THE USED DATASETS

| Paper | Data | | | | | | | | | Performance | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Type | Wrap | Defo scale | Time frame | Geographic scale | Spatial Resolution | Objects [Defo/All] | Task | Training Set [Synth/Real] | TP/P | FP/N | AUC |
| [24] | IFGS | Wrapped | / | 2016-2017 | Global | 110m | 4/$\geq$ 900 | Classification | 0/15,125[1] | 38/42[1] | 242/30,207[1] | 99.5%[1] |
| [23] | IFGS | Wrapped | $\geq 1800mm/year$ | 2016-2017 | Global | 110m | 4/$\geq$ 900 | Classification | 0/20,000[2] | 42/42 | 1327/30,207 | 88.2% |
| [39] | IFGS | Wrapped | 150mm | 2016-2017 | Global | 110m | 4/$\geq$ 900 | Classification | 910,000/0 | 38/42 | 295/30,207 | 93.12% |
| [40] | TS IFGS | Wrapped | $\geq 35mm/year$ | 2015-2019 | Campi Flegrai, Dallol | 100m | 2/2 | Classification | 910,000/0 | ?/?[3] | ?/?[3] | 98.94%[3] |
| [28] | TS | Unwrapped | 20-200mm | 2014-2018 | Wolf, Etna | | 2/2 | Denoising | ?/0 | / | / | / |
| [41] | TS | Unwrapped | -2.4m | 2014-2018 | Sierra Negra | ?m | 1/1 | Denoising | ?/0 | / | / | / |
| [25] | IFGS | Unwrapped | ?mm(50mm) | 2014-2018 | Campi Flegrei, Agung, Wolf, Sierra Negra, Alcedo | 90m | 5/5 | Classification, Detection | 20,000/0<br>20,000/20,000[4] | 8/20<br>16/20 | 6/32<br>5/32 | /<br>/ |
| [29] | TS | Unwrapped | ?mm | 2016-2019 | Erta Ale, Etna | 14m | 2/2 | Phase Gradient Gen. | / | / | / | / |
| [27] | TS | Unwrapped | 30-100mm | 2015-2016 | Masaya | 180m | 1/1 | Denoising | 400,000/0[6] | / | / | / |
| [30] | IFGS | Wrapped | 10-100mm | 2014–2020 | Turkey | 100m | 3/10 | Classification | Synth | 0/0 | 4/ 5000 | / |
| [26] | PSI TS | Unwrapped | 2-20mm | 2013, 2016 | Anatolia, Coso | 70m | 2/2[7] | Denoising | 225,000,000/0[7] | / | / | / |
| [42] | IFGS | Both, Wrapped | >28mm[5] | 2019-2021 | 29 Earthquake Zones | 100m[8] | ?/29 | Classification, Segm | 0/15,478 | /252 | /3,826 | 86.4% |
| [43] | IFGS | Wrapped | >28mm[5] | 2015-2018 | 32 Earthquakes | 480m[9] | 32/32 | Classification | 4,000,000/5,184 | Acc = | 982/1,152[10] | = 85.22%[10] |
| [44] | IFGS | Wrapped | 100-250mm | 2014-2020 | Taal, Etna, Erta Ale... | 100m | 10/14 | Classification | 0/7536 | 31/32 | 4/32 | / |
| | IFGS | Wrapped | 100-250mm | 2016-2021 | | 100m | 6/7 | Classification | 0/7536 | 347/404 | 10/365 | / |
| [45] | IFGS | Wrapped | 100-250mm | 2016-2021 | Taal, Cerro Azul, Sierra Negra, Fagradalsfjall, Etna, La Cumber, Erta Ale | 100m | 6/7 | Classification | 25,000/0 | 367/404<br>398/404 | 57/365<br>16/365 | /<br>/ |
| [46] | IFGS | Wrapped | > 28mm[5] | 2014-2021 | 44 Most Active | 100m | ?/44 | Classification | 0/>1,200,000 | 699/826 | 176/1648 | / |
| Ours | PSI IFGS | Wrapped | 5-30mm/year | 2015-2020 | Atacama | 200m | 4/46 | Classification | 330,848/836[11] | 134/142 | 201/633 | 86.44% |

[1] Performed 2-fold cross-validation on the training set, TPR = 89.9%, FPR = 0.8%

[2] 300 positive EnviSat samples are augmented to 10,000

[3] Time series' over two volcanoes are tested, Campi Flegrai 57, and Dallol 73 cumulative interferograms (IFGS).

[4] 173 Interferograms are augmented to create a 20,000-sample real data set

[5] Unreported deformation scale. The models rely on the fringes to detect deformations. Therefore theoretical limit is 28mm in between two time points (not accounting for noise)

[6] 20,000 time series of 20 interferograms each

[7] Denoising tested on areas of an earthquake and subsidence. For training, they used 25,000,000 time series of 9 timesteps each.

[8] Estimated from the paper and plots. If Sentinel-1 80-100km wide strip map is used, the angle of interferograms is 8-10 degrees (8.18 degrees of Sun-synchronous Orbit), divided by 1024 pixels. This results in 90-110m spatial resolution.

[9] Estimated from synthetic interferogram size of 108km and model input of 224x224 pix.

[10] Only accuracy is reported

[11] 836 non-volcanic deformation samples, 418 as is, and 418 added synthetic volcanic deformation

1500 active volcanoes in the world. However, nonvolcanic deformations such as slope processes and salt lake activities are common. As a result, the data are imbalanced. For example, 633 out of 775 samples contain only nonvolcanic deformations in our case. On the other hand, SAR interferometry processing needs expert knowledge and usually takes time. Therefore, the data that can be used to train the models are very limited.

Besides, volcanic deformations are usually swallowed by noise. The collected data suffer from atmospheric noise, which makes it not possible to detect subtle deformations. Although, after PSI processing, atmospheric noise can be reduced, the residual atmospheric noise remains; although the subtle long-term volcanic deformations are accumulated, there are more nonvolcanic deformation patterns. As shown in Fig. 1, the deforming volcano signal in the middle of the image is hidden by atmospheric noise of various shapes and intensities.

In summary, the volcanic deformation detection task is mainly restricted by data availability and data quality. The detection performance varies with the used datasets. In Table I, we compared the different datasets used and achieved performance by existing models. In Section II-B, we elaborate on how previous works deal with these challenges and improve their models.

### *B. Existing Volcanic Deformations' Detection Methods*

Fig. 2 and Table I show the overview of the current state of volcanic deformation detection from InSAR data using DL. The approaches can be distinguished by the used data type and selected modeling approaches.

First, the works differ in what type of input data is used and how the data are preprocessed. Short-term interferograms [I1 in Fig. 2] are used in many larger scale volcanic deformation detection studies [22], [23], [24], [28], [29], [38], [43], while long-term deformation detection using time series [II1] has been utilized in [25], [26], [27], [39], and [40]. These can be further split by how the atmospheric correction is handled: separately by removing atmospheric phase screen (APS) [I1a and II1a] [46] or by leaving this to be handled by DL [I1b and II1b] [22], [23], [24], [25], [26], [28], [29], [38], [41], [43]. Albino et al. [46] stated that atmospheric correction is of great importance for classification models, especially in steep tropical volcanoes. To the best of our knowledge, the velocity maps [II2] were utilized only for denoising task [25] and were not utilized before this study in the classification task.

Second, the preprocessing steps generate products of different levels of quality and, thus, impact the models' performances and detection sensitivity. Previous models use either wrapped [III1b] [22], [23], [24], [28], [29], [38], [39], [41] or unwrapped data [III1a] [24], [25], [26], [41], [46]. Unwrapping data need additional steps consuming time and resources. In some cases where the unwrapped data are rewrapped, depending on which wavelength is used, the rewrap methods are divided into overwrapping [III1c] and regular rewrap (using the original satellite wavelength) [III1b]. Studies [24], [41] show that fringes present in wrapped data facilitate convolutional neural networks (CNNs). Therefore, overwrapping is anticipated to increase models' sensitivity [39]. On the other hand, additional data, digital elevation model (DEM), the inclusion of amplitude, and so on can potentially benefit the models [III2]. Nevertheless, Gaddes et al. [24] did not show the benefits of additional channels for the transfer learning model. However, adding DEM data in the middle of a denoising model did help remove the residual noise from the data [25].

Third, the gap between training and test sets impacts the results. Synthetic training sets [III3] were proved to be helpful in training models when insufficient deformation data are available [24], [25], [26], [28], [29], [38], [39]. The parameters for simulations can be extracted empirically [III3b] or statistically [III3a] by variogram modeling [47]. To the best of our knowledge, the statistical approach explored in this work [47] has not yet been employed.

The data augmentation methods [III4] are helpful for increasing the number of RD training samples and the robustness of the model [22], [23], [43], [45]. However, augmentations were avoided in synthetic training datasets for more controlled and realistic sample distribution.

While test data should be taken from a separate distribution from training data, the domain should be kept the same. In cases where that is completely impossible (e.g., a model trained on imperfect synthetic data, which often keeps differences from RD), spatial filtering [IV1] could improve the similarity. That is tested in this study. Edge detection [IV2] was used [22], [23], [38] to exclude the cases of interferograms that do not contain strong deformations, therefore reducing the number of samples tested by the model (decreasing running time) and increasing the performance by removing the possibility of FP classifications by the model.

Most work can be split into two basic modeling tasks: denoising and detection, with which other techniques can be applied. Denoising models clear the images of atmospheric noise or unwanted deformations so that volcanic deformation and its scale can be easily identified. Detection models have focused on handling the noise and finding the volcanic deformations in large amounts of data. The modeling process of the deformation detection can be split into model-centric [V] and data-centric [VI] approaches. The model-centric approach focuses on model architecture improvements to increase performance. It can be divided by the implemented DL techniques: self-supervised learning [V1] [43], [44], classification [V2] [23], [24], [28], [38], [41], [42], [43], [43], object detection [V3] [24], semantic segmentation [V4] [41], denoising [V5] [25], [26], [27], [28], [40], and, a subcategory of these, the multioutput approach [V6] using single architecture and double-headed method to tackle different tasks simultaneously [24]. While most papers focused on binary classification, Gaddes et al. [24] use a multiclass approach to classify the types of volcanic deformations into point/sill and opening dyke.

The data-centric approach improves the model by better understanding and adapting the data. To understand the data and model's weak points, explainable AI (XAI) techniques are used [VI1]. So far, CAMs [42], [43] and case studies have been

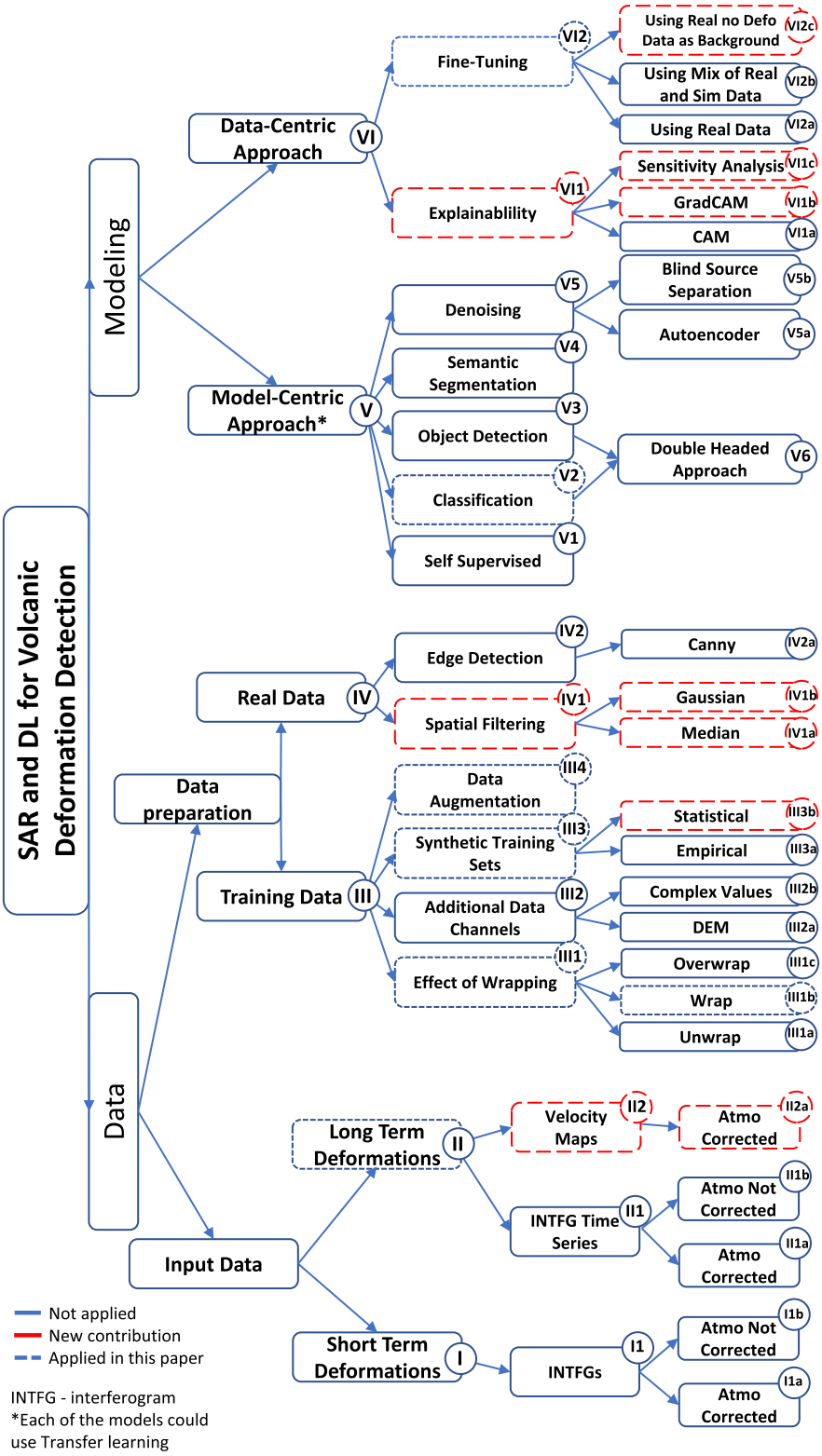



Fig. 2. Overview of the current state of research in the field of SAR and DL for volcanic deformation detection. The development is grouped in data- and modeling-related experiments. The fields explored in this article are shown in dashed lines, in red if it is a new contribution or blue if it is previously explored in another paper.

utilized to understand the models better. The models can be improved through different data processing adjustments and FT [VI2].

Due to the differences in the used datasets, comparing all those models fairly is not easy. Most comprehensive metrics, such as F1 score and precision, often are sensitive to skew data distribution, which is often the case in real volcanic deformation datasets. In these cases, more elementary metrics, true-positive rate (recall, TPR) and false-positive rate (FPR), are more informative. Also, the area under the curve receiver

operating characteristic (AUC ROC) is an informative metric as it combines TPR and FPR and is unaffected by the skew [48]. The datasets and achieved model performances are summed up in Table I.

Using edge detection and DL model trained on synthetic data, [38] reports 90.4% recall [38/42 true positive (TP)], 0.98% FPR [295 FP], and 93.12% AUC ROC using their best model on their test set of 30 249 interferograms and 42 deforming cases. In [39], they report the minimal deformation detection of 3.5 cm/year after overwrapping is applied.

Another approach training a multioutput DL model trained on 20 000 synthetic data intended for three-class classification and localization of the volcanic deformations [24] achieved 40% TPR (8/20 TP), 19% FPR (6/32 FP), and accuracy of 65% (34/52 correct class) on 52 interferogram classification task, and after inclusion of 173 real samples to the training set, 80% TPR (16/20 TP), 16% FPR (5/32 FP), and 83% accuracy (43/52 correct class) are achieved.

It has been shown that self-supervised learning significantly increases robustness, and the models perform well in a broader range of data [43]. In contrast, the transfer learning model of Inception v4-ImageNet fine-tuned on 7536 samples of real interferograms outperformed the compared models, giving the 96.87% TPR (31/32 TP) and 12.5% FPR (4/32 FP) out of balanced 64-sample validation set. The same model did not perform as well on the out-of-distribution test set. ResNet50-SimCLR achieved 85.89% TPR (347/404) and 2.74% (10/365 FP) out of a slightly imbalanced 769 samples set, containing ascending and descending, and filtered and unfiltered data.

To sum up, the classification approach is the most explored and the most reliable method for volcano detection in InSAR. More available and easily approachable short-term interferograms are widely used, while few studies focus on long-term velocity maps. RDs are often limited. Having a larger quantity of it for training is beneficial. When sufficient RDs are unavailable, synthetic data can help, but it introduces a problem of bridging a domain gap between the synthetic data and RD. The reported model performance varies and depends on the datasets used. While the majority of research has been focused on model-centric approaches, few works point out that data-centric improvements could make a larger impact. High-performing models have been published, but most approaches do not report the scale of detected volcanic deformations. The data SNR and the model's sensitivity to the scale of the volcanic deformation play a significant role when estimating models and models' metrics. The smallest reported detection is of 35 mm/year. In addition, the explainability of the models is underexplored. Often, single-case studies are shown, which are very limiting and cannot present the model capabilities with statistical significance.

## III. Study Area and Datasets

### A. Study Area

The study area covers the Andean Central Volcanic Zone (see Fig. 3), the Atacama Desert with 1113 volcanic edifices [49], of which a fraction is still active today. It contains high elevation differences of high planes and mountains of more than 6000 m, mainly covered by sand, stone, and lava, and is dotted with salt lakes. Slope processes, such as weathering, erosion, transport, and deposition of the material [50], [51], are common in the steep mountainous regions, which are prevalent in our InSAR stack data. Furthermore, the area has a high density of volcanoes and a sparse population distribution [9]. Therefore, remote sensing is the optimal way to monitor this area.

### B. InSAR Data

The gathered Sentinel-1 data cover 46 potentially active volcanoes in this region from October 2014 to June 2020; 35 interferometric stacks, with track numbers ascending 047, 076, and 149 and descending 010, 054, 083, 127, and 156, consist of 48–144 interferograms each (generally an acquisition every 6–12 days), spanning 4.5–5.7 years.

Different InSAR error sources must be considered in the processing chain to achieve high precision and accuracy in the deformation estimation. These error sources include atmospheric errors [52], [53], solid Earth tides [54], and subtle temporally short-lived signals, which may be confused with surface deformation [55]. If not corrected, the delays in a single interferogram could amount to tens of centimeters error [56]. We use IERS 2010 convention [57] to correct solid Earth tides, ECMWF ERA5 reanalysis data to correct tropospheric delay, and Tec maps CODE data to correct ionospheric delay using SAR Geodesy Processor (SGP) of DLR [58]. In addition, to overcome the sparsity of identified points for estimation, combined PSI [14], [15], [16] and DSI [32], [59] are performed to create stacked deformation maps. Processed this way, InSAR deformation maps achieve accuracy in the mm/year level [33].

The geocoded deformation maps are converted to wrapped interferograms in two steps. First, the map is scaled by the temporal baseline to achieve cumulative surface displacement. Second, the displacement is wrapped between $-\pi$ and $\pi$, given the Sentinel-1 wavelength. The interferograms generated as such are expected to be free from erroneous atmospheric and short-lived bias signals and have a high SNR. The accuracy of such interferograms in showing the surface displacement is related to the accuracy of the techniques for correcting the error sources. The resulting dataset is referred to hereafter as RD to distinguish it from the simulated data. RD samples containing volcanic deformations are used only to validate our method. RD samples over known volcanoes are used for validation and the rest for fine-tuning the model.

The velocity maps created this way are cropped around the potentially active volcanoes and filtered out if they contain missing values. They are resampled using an averaging window of 200 m. This gave 32 frames of 512 $\times$ 512 pixels and 102.4 $\times$ 102.4 km$^2$ (achieving pixels of 200 m), centered over 24 volcanoes. For feeding the data to the model, the whole of the volcanic deformation must be included in a patch while keeping the patch not excessive in size. To this end, the frames are split into patches of 256 $\times$ 256 pixels, with a stride of 64 pixels, producing 25 patches for each frame.

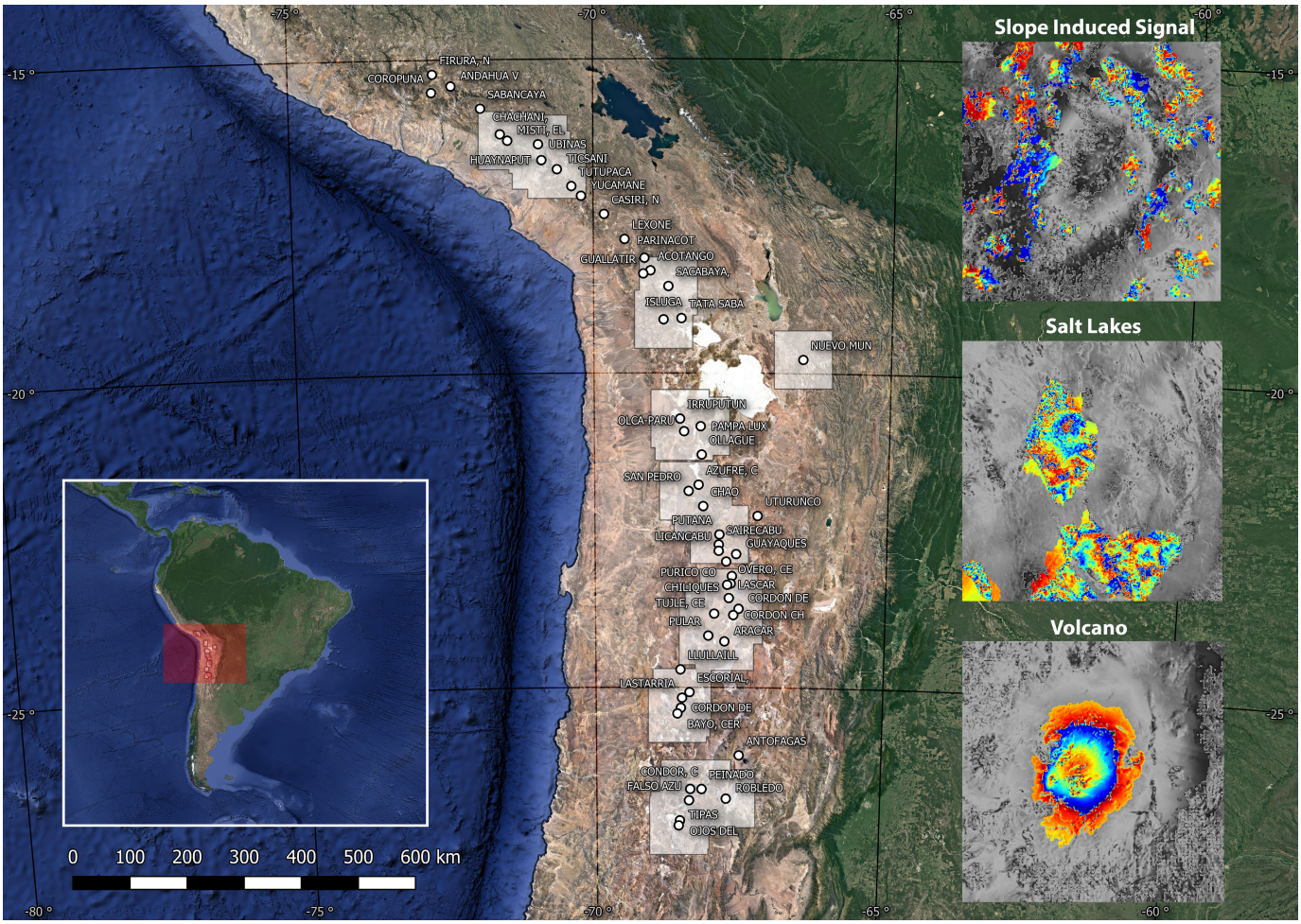



Fig. 3. Examined Andean Central Volcanic Zone. The area is densely packed with volcanoes but also with slopes and salt lakes. The InSAR deformation patterns caused by slope-induced signals and salt lakes are widely distributed in the deformation maps, which can confuse the detection models from detecting volcanic deformation patterns.

### C. Simulated Data

The training set comprises synthetic volcanic deformations and residual atmospheric noise patches. After the atmospheric correction, the latter remains in the interferogram and is considered when creating a simulated dataset with a similar distribution of the RD set. Thus, half of the simulated residual atmospheric noise is added to the volcanic deformations, and half is used as the nondeformation class. For the residual atmospheric noise simulation, an assumption is made that most of the stratified atmosphere is eliminated by corrections algorithms and PSI [60], [61]. Therefore, residual noise patterns are statistically extracted from RD using variogram modeling [47] and then simulated using a Monte Carlo approach [27]. A difference in the noise scale between ascending and descending tracks ($1.6598 \pm 0.3873$ and $0.9205 \pm 0.03438$ mm/year, respectively, Fig. 4) is observed due to different acquisition time in a day [47].

Volcanic deformations are simulated with a volumetric source using the compound dislocation model (CDM) [62]. The simulations correspond to 23 631 diverse volcanic patterns of the sill, sill-like, Mogi-like, near-equidimensional volumetric source, oblate ellipsoidal source, and penny-shape-like crack. Thus, the synthetic volcanic deformations depict global variance in volcanic deformations.

Originally, the scale of volcanic deformations was set to be 0.2–20 mm/year. The distribution of volcanic deformations is exponential and skewed toward more minor deformations. This in hand meant that a significant fraction of the training set might not have an easily distinguishable volcanic uplift when covered by atmospheric noise, leading to the model being trained to detect spurious correlations. Therefore, an additional refined set is created.

The synthetic set is refined by rescaling the volcanic deformations to be the same or larger than $\mu \pm 3\sigma$ of atmospheric noise. For ascending tracks, the corrected range is from 3–20 mm/year and, for descending tracks, 1.2–20 mm/year. The volcanic deformations below the threshold are redistributed to normal distribution centered on 8-mm/year deformations, which is found to be the most common deformation scale in the data. The original distribution and the new distribution can be seen in Fig. 5.

The simulated data are then projected to similar LOSs as RD. The LOS parameters are picked from the RD if they significantly differ from others. Otherwise, they are grouped and averaged to represent morefold data. This grouping results

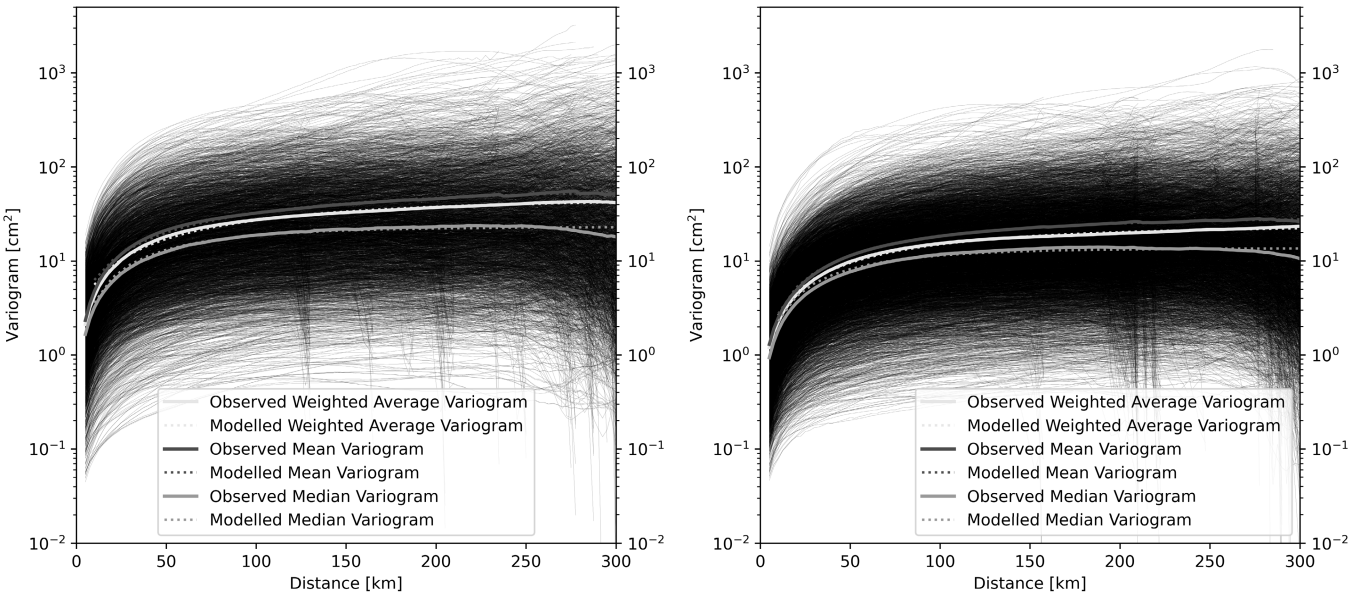



Fig. 4. Separately modeled variograms for ascending and descending tracks. Different statistical functions are explored to determine the values for modeling the variograms. The ascending track maintains a higher noise level and variance. (a) Ascending. (b) Descending.

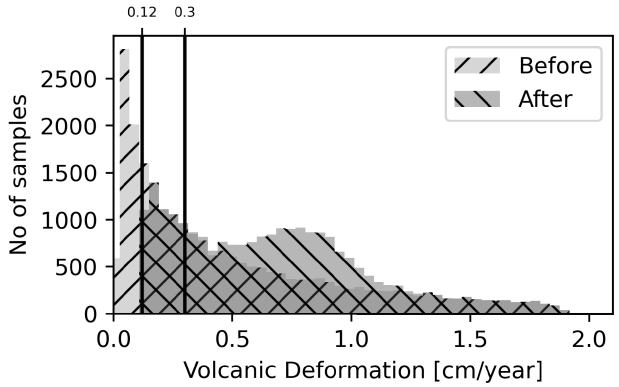



Fig. 5. Histogram of the maximum volcanic deformation present in training set before and after refinement. The noise thresholds of 1.2 mm/year for descending and 3 mm/year for ascending geometry are marked. Refined data redistribute the samples below the threshold to Gaussian distribution centered around 8 mm/year.

in five descending and two ascending sets of LOS parameters. To create an equal amount of data for each LOS, we created seven subsets of simulated data. The volcanic simulations are reused, and the different residual atmosphere data are simulated for each set, amounting to 330 848 samples. Of all the simulated samples, 90% is used for training, 5% for validation, and the rest for synthetic testing.

When using simulation data to solve the problem of the shortage of training data, there is a domain gap between the synthetic data and RD. The velocity map data span six years, over which terrain suffers many changes. Slope processes, erosion, salt lake mining, and level changes contribute to the final product. None of these patterns are easy to simulate and, therefore, have not been considered in the training data. As a result, they constitute a majority of the domain gap.

## IV. Method

To automate the detection of such deformations and timely monitor the activity of volcanoes, we use a DL model to detect the presence of slow, sustained volcanic deformations in velocity maps. We use the detection approach formulated as a binary classification problem.

Initially, the models are trained on synthetic data with volcanic deformation and nondeformation classes. Explainability analysis via Grad-CAM helps guide the improvement process and identifies the model's difficulties. The initial synthetic set is refined accordingly, and the model is fine-tuned. Spatial filtering and fine-tuning using a hybrid synthetic-real set help improve detection accuracy. Finally, the sensitivity analysis tests the model on an extended range of data and uncovers the model's minimal detectable deformation.

### A. CNN Models

In [43], it is demonstrated that transfer learning from models trained in optical data, such as ImageNet, provides less meaningful features for the volcanic detection tasks. Therefore, we train and validate the models entirely using simulated data and test on the real dataset.

Due to their popularity in classification tasks and the advantage of extracting and squeezing spatial features from input images, CNNs are also used in this article. The models use a one-channel image of 256 × 256 pixels, corresponding to an area of 51.2 × 51.2 km$^2$, having a spatial resolution of 200 m as input. Different hidden layer configurations are tested. Final binary classification probabilities are produced by the last layer using one node with sigmoid activation.

To select the best model suited for the task, AlexNet, VGG16-inspired architectures [63], and recent InceptionResNet v2 [64] are compared in this article. Our experiment results show that InceptionResNet v2 outperforms other models by 14% in the AUC ROC on the synthetic test set; hence, it is selected for further experiments.

The models are trained from random initialization. The labeled synthetic training dataset is used for training, and

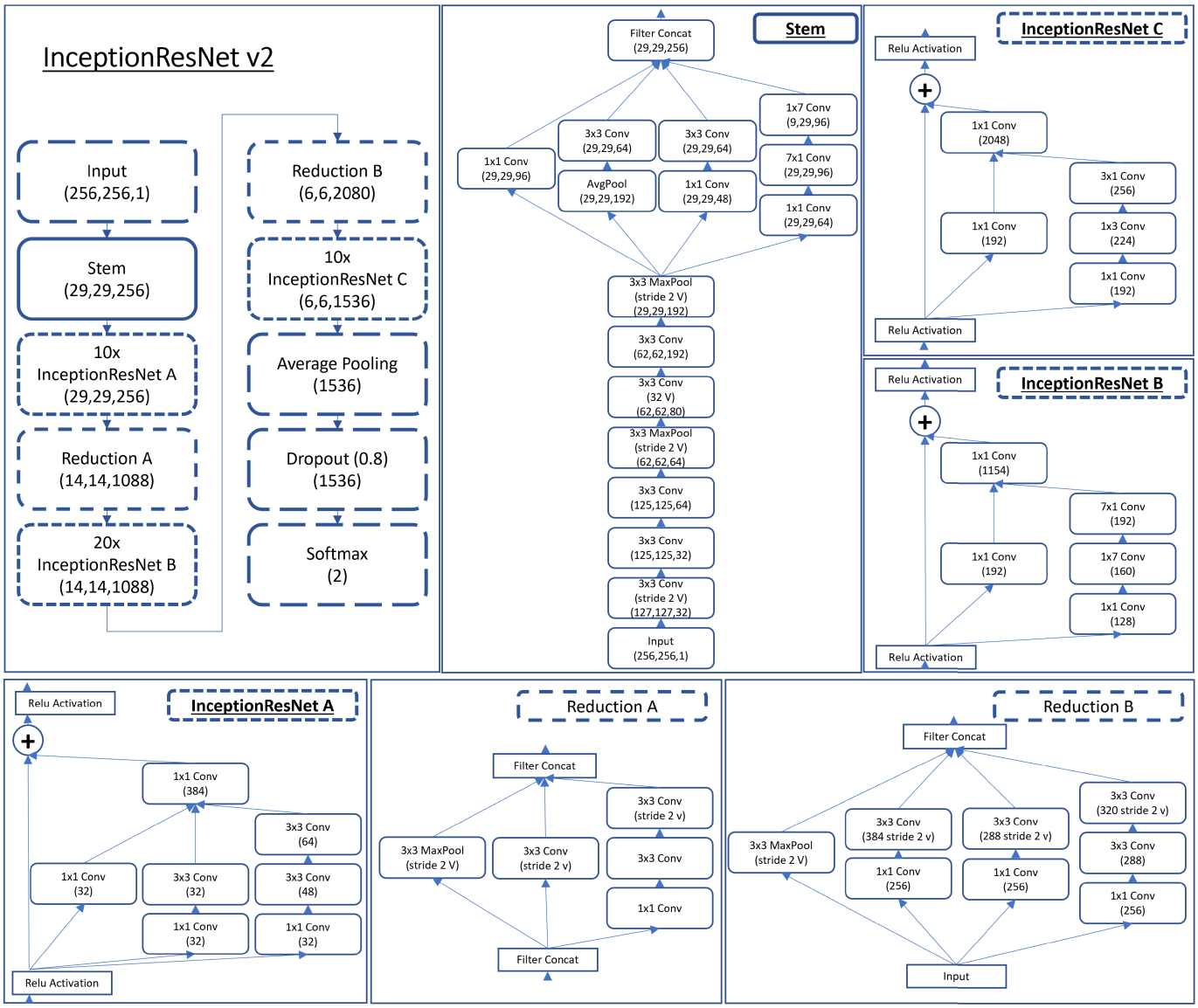



Fig. 6. Network architecture of used InceptionResNet v2 showing the layer dimensions using an image size of 256 × 256 pixels. The network consists of stem and multiple modules of InceptionResNet A, B, and C, and reductions after them Reductions A and B, illustrated in subplots.

the models are trained using Adam optimizer [65] and cross-entropy loss. To select the best model, we employed the early stopping mechanism. The training process is stopped when no further loss improvement is discovered after ten consecutive epochs with a maximum of 100 epochs. With an extensive training set, most models need no more than 30 training epochs.

*1) CNN Architectures:* The InceptionResNet v2 model contains residual and inception blocks, which allows it high flexibility and increased capacity when learning to distinguish features. Residual blocks introduce skip connection, which improves the stability of a network allowing deeper networks and merging low-level features with high-level features [66]. Inception blocks introduce different optimizations, such as asymmetric convolutions, which reduces the number of parameters while maintaining the expressivity and extracts features at different scales, small details, and mid and larger sized features [67]. The Inception ResNet v2 architecture can be seen in Fig. 6.

Alexnet is chosen and used as a best-performing architecture, among the tested, for volcanic deformation detection in [22], [23], and [38]. It has fewer layers compared to more recent models, and it uses large convolutional matrices, which allows it to handle larger and more complex features.

VGG16 [63] has also been successfully applied for volcanic deformation detection and classification [24]. It uses convolutional layers with small kernels (3 × 3), reducing the number of parameters while allowing depth to handle the more complex and larger features. Two architectures, ConvNet8 and ConvNet16 (the number representing the number of layers), inspired by it are tested.

*2) Detection Metrics:* We use five standard detection metrics for evaluating the model's performance

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

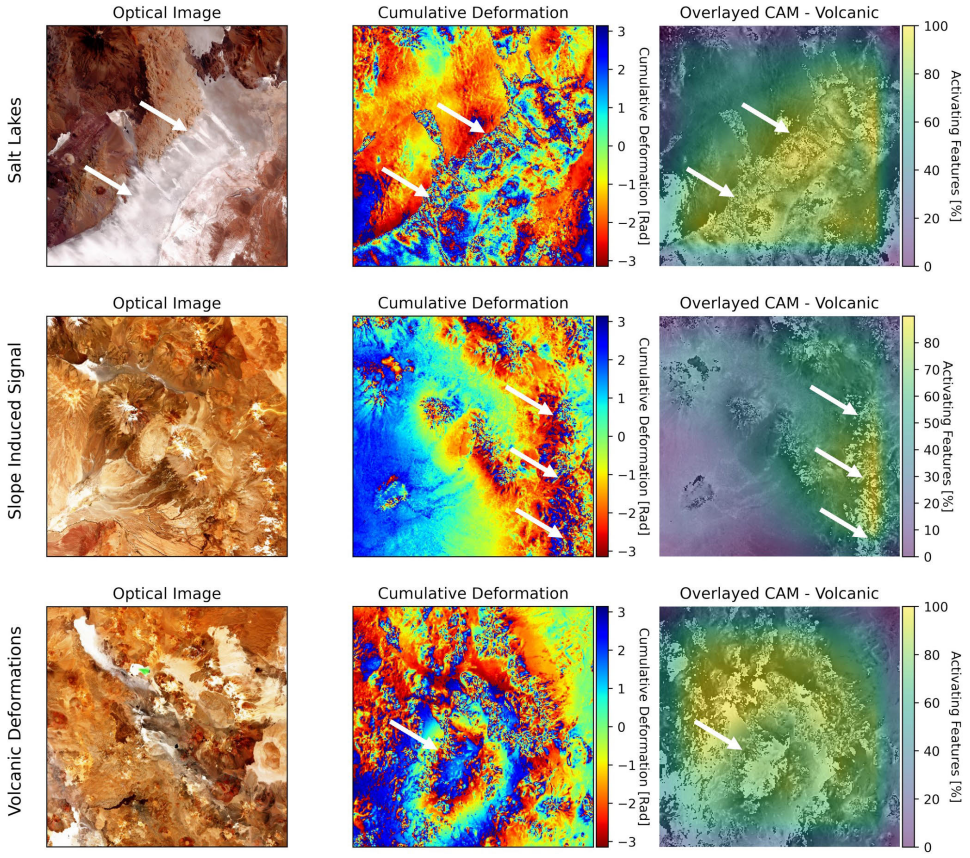



Fig. 7. Three typical regional deformation patterns and the corresponding saliency maps of an initial model generated by Grad-CAM. The model mistakes the patterns of the salt lake and slope-induced signal for volcanic deformations.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{TPR} = \frac{\text{TP}}{\text{FN} + \text{TP}} \tag{3}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{4}$$

$$\text{F1} = 2 * \frac{1}{\frac{1}{\text{Precision}} + \frac{1}{\text{TPR}}} \tag{5}$$

where TP, FP, TN, and FN are numbers of true-positive, false-positive, true-negative, and false-negative samples. TPR and FPR are true-positive and false-positive rates. Since the test set is highly imbalanced, most metrics are skewed. Detecting all volcanic deformations is paramount for our task (best quantified by TPR), followed by reducing the number of false detections (shown by the FPR).

The ROC is a curve showing TPR plotted against FPR when the threshold is varied and is not affected by the data skew [48]. The area under the curve is derived to evaluate the models' ability to balance the high detection rate and the low false-positive rate, and is used to represent overall model performance in this work. To trade off between TPR and FPR, classification threshold selection is essential.

### B. Fine-Tuning Strategy

Since the geographic and temporal patterns were not accounted for in simulations, all-present slope-induced signals and salt lakes confuse the models. To bridge this gap, the FT step is performed.

The RD needed to be used for fine-tuning, but, since volcanic data were scarce, all the data on known volcanoes were kept for testing. For fine-tuning, we use the region around the central Andes, which does not contain volcanic deformations or test data. The volcanic simulation data were added to half of the extracted patches of slope-induced signal and sparse salt lakes. This way, the balanced hybrid synthetic-real set is created, having a deformation and nondeformation class.

Since the RDs are very scarce, the FT set is limited. For fine-tuning the more significant parts of the model, it is necessary to have a larger dataset. Therefore, we tested freezing the model and fine-tuning only the last layer, the last two layers, and the last two layers fine-tuned individually (last, second last, and then last again). Our previous experiment demonstrated that retraining only the last layer worked the best [68], which is also adopted here.

### C. Explainability Analysis

DL models are often called black box models, but, in the past decade, there has been extensive work to improve the explainability and interpretability of the models. While the explainability is helpful for different purposes [69], we use it to verify and improve the system. We apply two explainability tools, i.e., Grad-CAM for analyzing the model's focus and t-SNE to visualize extracted features.

*1) Models' Activation Maps Analysis:* We needed feature attribution methods of explainability for the computer vision tasks and classification problems. Many feature attribution

TABLE II

DETECTION METRICS OF INCEPTIONRESNET V2 WITH DIFFERENT TRAINING STRATEGIES. USING A REFINED SYNTHETIC TRAINING SET AND FINE-TUNING THE AFOREMENTIONED MODEL HAVE IMPROVED THE AUC ROC BY 12.01% AND 28.52%, RESPECTIVELY. THE HIGHEST AUC ROC OF 90.29% IS ACHIEVED VIA FINE-TUNING THE REFINED MODEL AND TESTING IT ON THE FILTERED DATA. BESIDES, FINE-TUNING STRATEGY AND ADJUSTING A TEMPORAL BASELINE CAN HELP IN REDUCING THE FPR

| Model | TP | FP | FN | TPR ↑ | FPR ↓ | Accuracy ↑ | Precision ↑ | F1 ↑ | AUC ROC ↑ |
|---|---|---|---|---|---|---|---|---|---|
| Wrapped data | 139 | 452 | 3 | 97.89 | 71.41 | 41.29 | 23.52 | 37.93 | 57.92 |
| Refined | 142 | 627 | 0 | 100.00 | 99.05 | 19.10 | 18.47 | 31.17 | 69.93 |
| Spatial Filtering (LSF) | 142 | 537 | 0 | 100.00 | 84.83 | 30.71 | 20.91 | 34.59 | 76.42 |
| Fine-tuning (FT) | 134 | 201 | 8 | 94.37 | 31.75 | 73.03 | 40.00 | 56.18 | 86.44 |
| FT + LSF | 62 | 37 | 80 | 43.66 | 5.85 | 84.90 | 62.63 | 51.45 | 90.29 |
| 4y Temporal Baseline | 125 | 102 | 17 | 88.03 | 16.11 | 84.64 | 55.07 | 67.75 | 93.20 |

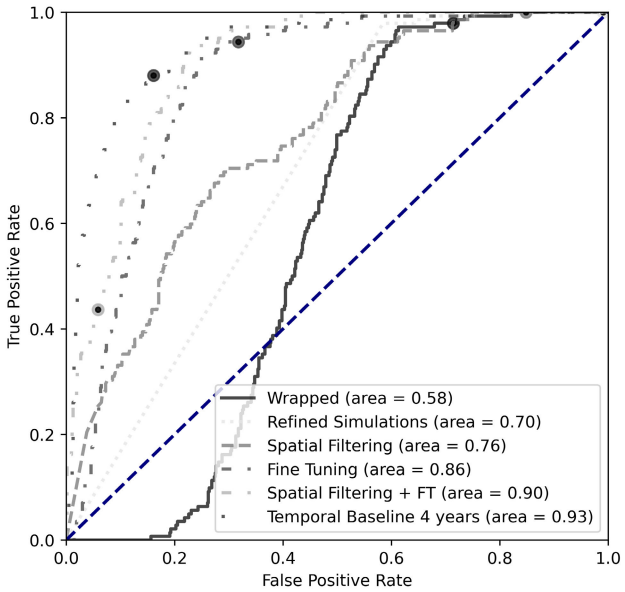



Fig. 8. ROC of the models on real test data, showing incremental improvements of each technique. Finally, the improvement that the custom temporal baseline adjustment would give is presented.

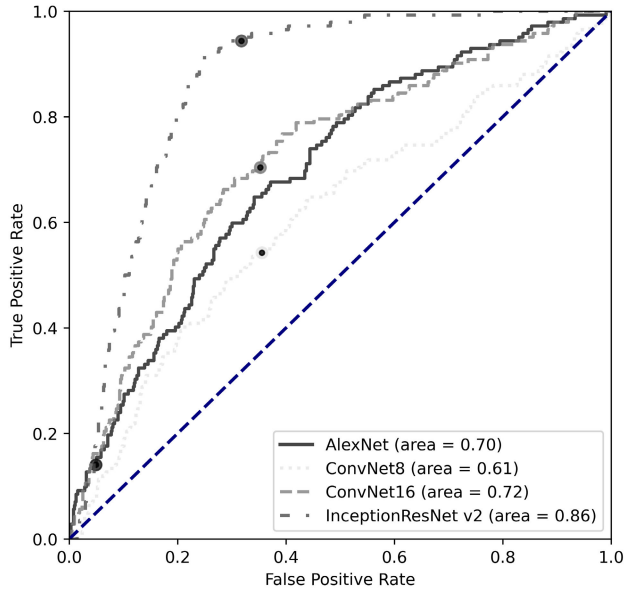



Fig. 9. Comparison of the ROCs on the real test set of models using the FT approach. InceptionResNet v2 architecture demonstrates significantly better performance in comparison to other tested models.

methods suffer from the independence of model randomization and label permutation, and Grad-CAM does not [70]. It is also one of the most reliable and least resource and time-consuming methods for Earth observation, as shown in [71].

Grad-CAM [34] is a local post-hoc feature attribution saliency map generation technique, which shows where in the image the model is "looking at." To produce a Grad-CAM, the gradient of the output class score is determined with respect to feature maps of a selected convolution layer (the last convolutional layer in our case). Then, by global average pooling the gradients, weights are obtained. Then, a weighted combination of feature maps is followed by a rectified linear unit (ReLU), giving a coarse Grad-CAM heatmap. Furthermore, the heatmap is normalized and upscaled for visualization.

Grad-CAM visualizations display the model's attention map but do not directly show the conclusion of whether the model is well-trained or not, which can be misleading. The model's classification output must be taken into consideration as well. Therefore, we scale the Grad-CAM activation maps with the models' probability of containing the volcanic deformation, combining them into a fully informative visualization.

*2) Visualization of Learned Features:* Comparison of extracted features from the FT and test sets via trained model can better understand the model's function and discover potential pitfalls. It is performed by visually comparing the similarity of the distributions of extracted latent features from the two sets. We use the last convolutional layer output as the extracted features, with a dimension of 1536. To visualize it in a low-dimensional space and keep its original modality, a widely used nonparametric nonlinear dimension reduction tool t-SNE [35] is employed.

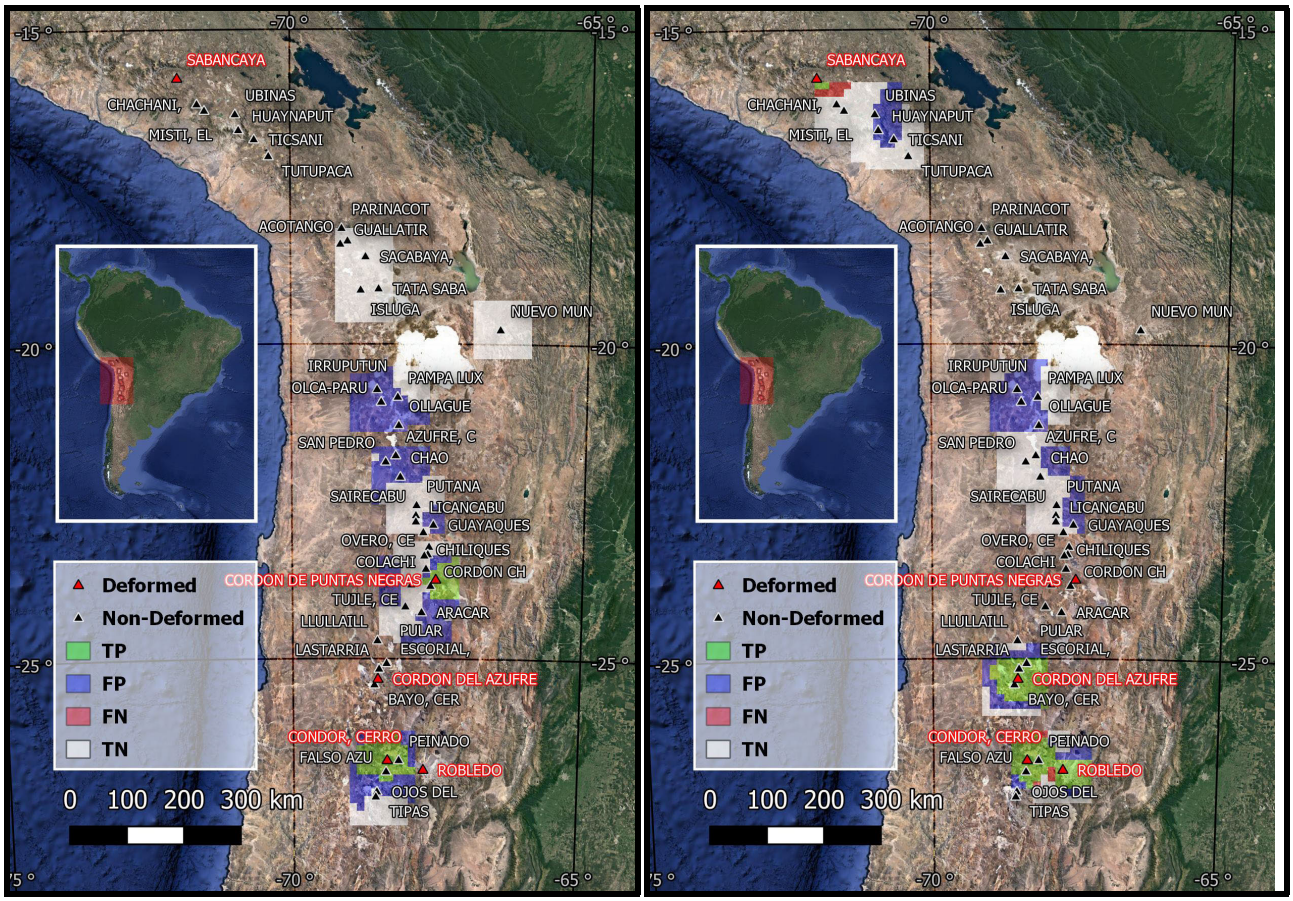
It starts by converting the Euclidean distances of feature pairs $x_i$ and $x_j$ into conditional probabilities $p_{ij}$ where points $x_i$ and $x_j$ are neighbors using student's t distribution

$$p_{ij} = \frac{\exp\left(-\|x_i - x_j\|^2/2\sigma^2\right)}{\sum_{k\neq l}\exp\left(-\|x_k - x_l\|^2/2\sigma^2\right)} \tag{6}$$

where $\sigma$ is a Gaussian variance centered on the point $x_i$. Similarly, the conditional probabilities $q_{ij}$ in feature pairs $y_i$ and $y_j$ in low-dimensional space are calculated using t-student distribution as

$$q_{ij} = \frac{\left(1 + \|y_i - y_j\|^2\right)^{-1}}{\sum_{k\neq l}\left(1 + \|y_k - y_l\|^2\right)^{-1}}. \tag{7}$$

To keep the modalities of the high-dimension data in the low-dimensional space, the Kullback–Leibler divergence (8)




(a) (b)

Fig. 10. Wide area demonstration split by the satellite track. The model detected all five deforming volcanoes that are visible in the data. TP, FP, FN, and TN represent true-positive, false-positive, false-negative, and true-negative detections, respectively. *The visualization does not reflect the statistical results because of the two levels of weighted averaging of overlapping patches and frames. (a) Ascending. (b) Descending.

between conditional probabilities $p_{ij}$ and $q_{ij}$ is minimized, which is achieved by minimizing $\mathrm{KL}(P\|Q)$ using the gradient descent algorithm

$$\mathrm{KL}(P\|Q) = \sum_{i\neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}. \tag{8}$$

Then, the original high-dimension data $x_i$ are, thus, embedded into a low-dimensional (2-D in our case) space and are converted to $y_i$. Now, we can directly compare the simulated and RD using the converted features in 2-D space.

## V. Experiments

### A. Initial Detection Results

The models are trained from random initialization. The labeled synthetic training dataset is used for training; 5% of the synthetic data is used for validation and synthetic testing each. The models are trained using Adam optimizer and cross-entropy loss. For the best model weights' selection during training, checkpoint and early stopping mechanisms are employed. All the models are trained using a maximum of 100 epochs, with the early stop mechanism set to 10 epochs without loss improvement before training ends. Having a sizeable synthetic training set, most models did not need more than 30 epochs of training.

The models are compared on unwrapped and wrapped data. The wrapped case performed better, which the presence of fringes can explain. Fringe is a clear border created by the wrapped deformation over the value of $2\pi$ (a contour between $\pi$ and $-\pi$). These contours can be detected more easily by convolution layers in models [24].

Even though the model's performance on the wrapped data is better, AUC ROC is low. The performance on the synthetic validation set is similar to the training set but not satisfactory on a real set. To correct this, the deformation scale of simulations is checked, and a refined synthetic set is created. The refined synthetic set does not obscure smaller volcanic deformations; therefore, training the model is easier. This improvement brought a 12% increase in AUC ROC between the best models of both approaches.

### B. Grad-CAM Visualization

After training the model, we generated a Grad-CAM visualization map of each test sample. These models' activation maps are overlayed with the original images. Further analysis is conducted to validate the model: 1) check whether high activations are distributed over the volcano area in images classified as volcanic deformation and 2) find the patterns over high activation areas in FP samples.

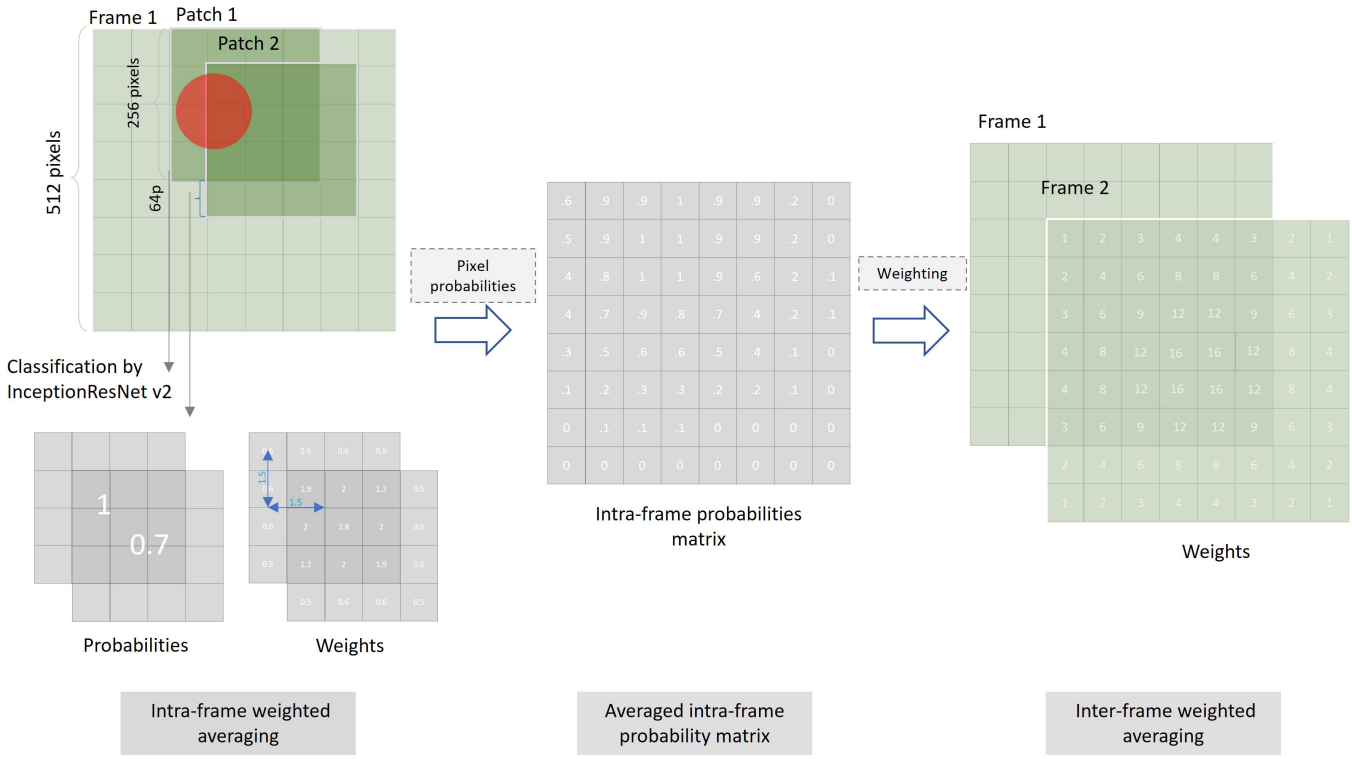



Fig. 11. Illustration of used voting system to transform the patch level probabilities into geographic areal representation. It consists of two steps: intraframe and interframe weighted averaging. The result is a unified map showing the presence and probability of volcanic deformations.

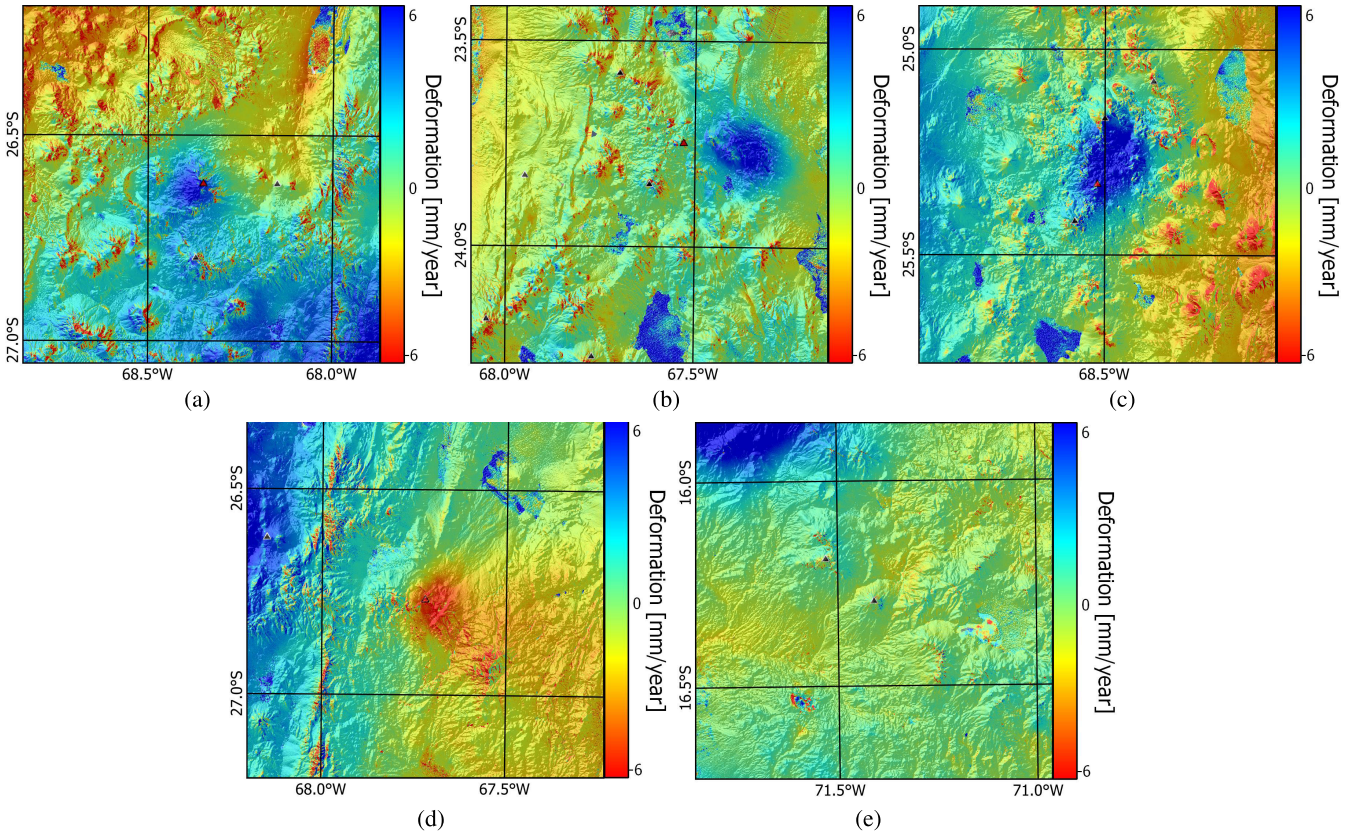
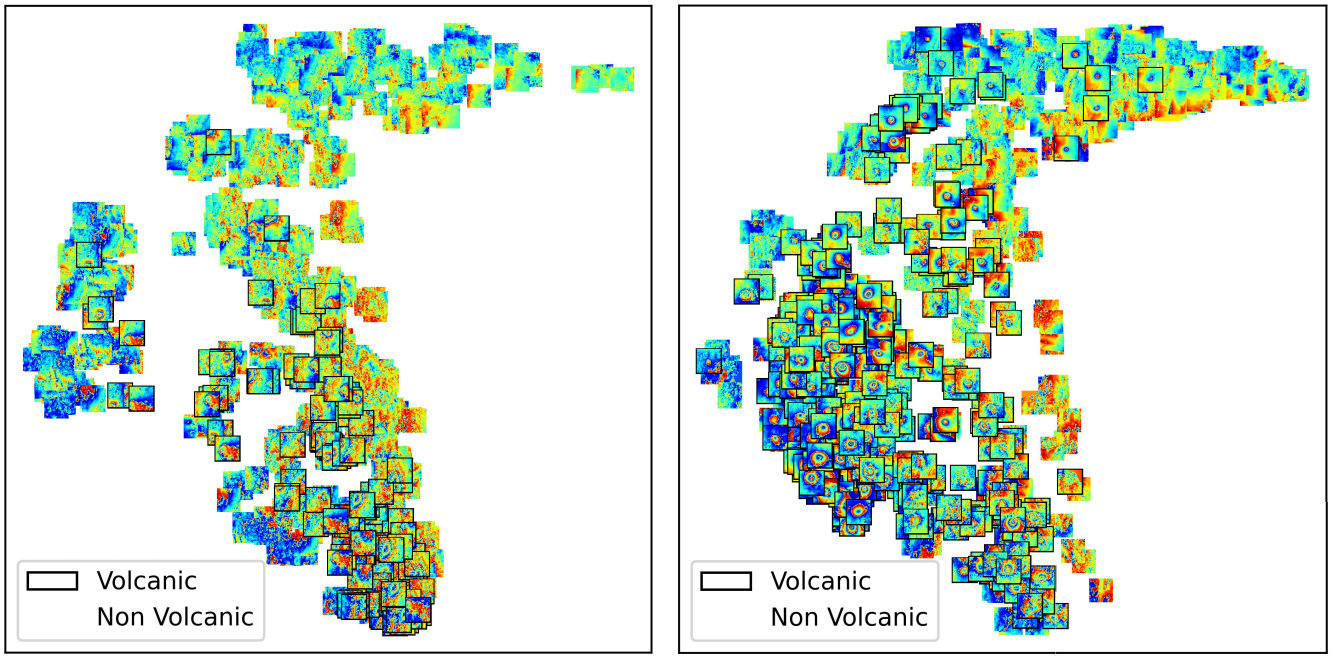



Fig. 12. Showcases of the unwrapped deformation maps of five deforming volcanoes. (a) Cerro El Có. (b) Cordon de Puntas Negras. (c) Lazufre. (d) Robledo. (e) Sabancaya.

Fig. 7 shows a correct (bottom) and two false classified samples. The top visualization shows a strong activation signal for a salt lake, and the mid shows how the slope-induced signal confuses the model. This is not an isolated case of these patterns being misidentified as volcanic deformation, which must be corrected. Slope-induced signals and salt lakes




(a) (b)

Fig. 13. t-SNE transformation of feature space of fine-tuned model on the real test data and FT set. The distribution follows the one presented in Fig. 14. It is noticeable that most of the volcanic deformations are grouped well together. In the real set, there are two volcanic clusters and about ten examples of partial deformations further away from these clusters. In the FT set, there is a big volcanic cluster with smoother patterns than the real test set cluster and a cluster and outliers with high similarity to real samples. (a) Real test set. (b) FT set.

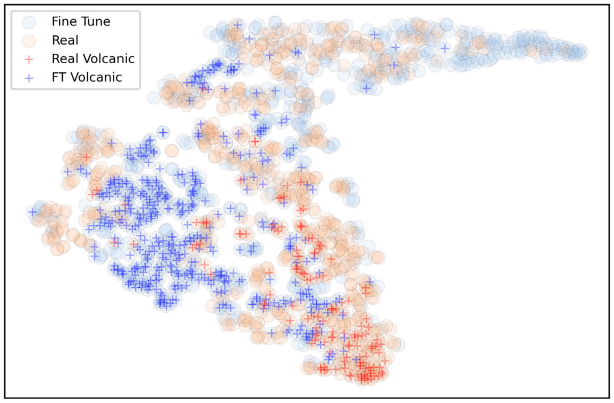



Fig. 14. Comparison of samples of real and FT sets visualized using t-SNE. The FT data are different from RD, but there is a significant overlap of distributions.

are widely present in the explored region but have not been addressed in the synthetic data. To overcome this obstacle, an FT step is performed.

### C. Fine-Tuning Using Slope-Induced Signal

We identified and explored two approaches to solve the problem of unaccounted patterns: to make the training dataset more similar to the RD or, the other way around, to make the RD more similar to the simulated data.

LSF makes RD less noisy and more similar to synthetic data. The median filter using different kernel sizes is tested and improved the results. The kernel size from 20 to 50 can filter out the noise and keep the volcanic deformation signal due to their spatial scale difference. Volcanic deformations in our data ranged from about 15–50 km or 75–250 pixels and should not be filtered out by smaller filter sizes.

However, improving the training set is more challenging because simulating the newly uncovered nonvolcanic deformation patterns is difficult. Thus, patches with similar deformation patterns of slope-induced signal and salt lakes are extracted directly from RD over parts of the Andes where no volcanoes exist. A combination of the 836 cropped patches, half with and half without synthetic volcanic data, is henceforth referred to as the FT set. It is used to fine-tune the last layer of the model using a small learning rate to improve the detection performance [72], [73].

Table II listed and compared various classification metrics achieved by different approaches, where fine-tuning and testing the model on the filtered RD set give the best performance. As shown in Fig. 8, the ROC also suggests that the proposed approach obtains a significantly higher detection rate over others when the FPR is small (FPR $< 0.2$).

After fine-tuning, InceptionResNet v2 increases its lead in performance compared to other architectures. The comparison of ROC curves between different fine-tuned architectures, seen in Fig. 9, shows that InceptionResNet v2 has 14% more considerable AUC ROC value than other models.

### D. Results Fusion

Since there may be multiple frames around one volcano, and each frame was cut into several patches, a result-fusion step is applied to the detection results to identify the deforming volcanoes. It fuses the intraframe and

TABLE III
LIST OF THE VOLCANIC DEFORMATIONS (MM/YEAR) IN NINE FRAMES WITH DIFFERENT PSEUDOTEMPORAL BASELINES. THE MODEL HAS SUCCESSFULLY DETECTED THE BOLDED VALUES. THE FINAL COLUMN GIVES THE MINIMAL DETECTED DEFORMATIONS BY THE MODEL (TRAINED ON CUMULATIVE DEFORMATIONS OF ≈5 YEARS), AND UNDERLINED VALUES HIGHLIGHT THE MINIMAL DETECTION OVERALL

| Frame | Volcano | $\approx 5.0y$ | 4.5 y | 4.0 y | 3.5 y | 3.0 y | 2.5 y | 2.0 y | 1.5 y | 1 y | Min Defo |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Cerro El Cóndor | **10.0** | **8.8** | **7.8** | **6.9** | **5.9** | **4.9** | 3.9 | 2.9 | 2.0 | 4.9 |
| 2 | Cerro El Cóndor | **9.9** | **8.7** | **7.8** | **6.8** | **5.8** | **4.9** | 3.9 | 2.9 | 1.9 | 4.9 |
| 3 | Cerro El Cóndor | **17.5** | **16.8** | **14.9** | **13.0** | **11.2** | **9.3** | **7.5** | 5.6 | 3.7 | 7.5 |
| 4 | Cordon De Puntas Negras | **9.6** | **9.0** | **8.0** | **7.0** | **6.0** | **5.0** | 4.0 | 3.0 | 2.0 | 5.0 |
| 5 | Cordon De Puntas Negras | **7.7** | **7.2** | **6.4** | **5.6** | **4.8** | **4.0** | 3.2 | 2.4 | 1.6 | 4.0 |
| 6 | Lazufre | **5.3** | **4.5** | **4.0** | **3.5** | **3.0** | **2.5** | **2.0** | 1.5 | 1.0 | <u>2.0</u> |
| 7 | Lazufre | **10.5** | **9.0** | **8.0** | **7.0** | **6.0** | **5.0** | **4.0** | 3.0 | 2.0 | 4.0 |
| 8 | Lazufre | **9.0** | **7.7** | **6.9** | **6.0** | **5.1** | **4.3** | **3.4** | 2.6 | 1.7 | 4.3 |
| 9 | Robledo | **-11.4** | **-11.0** | **-9.8** | **-8.5** | **-7.3** | -6.1 | -4.9 | -3.7 | -2.4 | -7.3 |
| | Min Defo | 5.0 | 4.5 | 4.0 | 3.5 | 3.0 | 2.5 | 2.0 | / | / | |

These deformation values presented are relative to the spatial and temporal reference point for each stack. They refer to the deformations along the LOS; the nine stacks of InSAR data also have slightly different time frames.

interframe detection results around a volcano. The resulting visualization allows checking the potential deforming volcanoes in a single visualization, instead of many individual patches.

Fig. 10 shows the geographical distributions of detection results. Here, green, blue, red, and white rectangles indicate TP, FP, FN, and TN detected areas within a frame, respectively. In general, all the deforming volcanoes are identified by the model. However, there exist significant numbers of FPs. The FP distributions are independent of track directions (ascending and descending) but follow ridges of mountains and significant elevation changes.

*1) Fusion Method:* In the first step, an empty probability matrix is created to collect the probabilities from each patch to get the intraframe votes. The probability matrix, $8 \times 8$ pixels in size, spatially represents the frame, each pixel covering $64 \times 64$ pixels in the patches/frames. As the stride is also 64 pixels, each patch matches and overlaps $4 \times 4$ probability matrix pixels. Overlapping patch probabilities are weighted with the Euclidean distance to the center of the overlapping pixels in the probability matrix and summed up. Finally, the sum of weighted probabilities is divided by the sum of weights giving the final probabilities for each probability matrix pixel within the frame.

A similar approach is to use interframes. Frames probability matrices are forced into a common grid, and each pixel is given a weight equal to the number of overlapping patches over the pixel within that frame. In addition, overlapping pixels' interframes are weight-averaged, providing a single probability for each pixel in the whole area (see Fig. 11). Because of these two levels of weighted averaging of overlapping patches and frames, the visualization does not fully reflect the statistical results in Table II.

This approach gives lower weights to the edges of frames, prioritizing frames with centrally positioned and fully visible volcanoes, where the classification decisions are made more confidently. The voting system maintained the detection of all visible deforming volcanoes (see Fig. 10).

*2) Detected Deforming Volcanoes:* The models identified five volcanic deformations in the region, as shown in Fig. 12. First, Cerro El Cóndor (Antuco) has been reported to release gas and steam on April 20, 2013 [74]. While Cerro El Condor is estimated to be a young and faster growing volcano in the region [75], up until now, to the best of our knowledge, the growth has not been practically confirmed. This is the first case of measurement confirmation of the Cerro El Condor uplift. During our study from 2014 to 2020, it deforms at 9.9–17.5 mm/year in the LOS. Second, Lazufre was observed to uplift ≈20 mm/year from 1998 to 2011 [11] and about 30 mm/year from 2003 to 2010 [76], which slowed down to 5.2–10.4 mm/year during our observations. Third, deformation ≈25 km east of Cordon de Puntas Negras was previously reported as deformation near Cerro Overo [11]. It subsided at a rate of 4 mm/year up to 2003–2005 and then began uplifting 5 mm/year to 2010. During our study, it is uplifting 7.7–9.6 mm/year. Fourth, Robledo, also known as Cerro Blanco, subsided in the survey from 1992 to 1996/7 26 mm/year, slowing to 18 mm/year to 2000, 12 mm/year from 2003 to 2007, and 8.7 mm/year from 2005 to 2010 [11], [77]. Based on our research, from 2015 to 2020, it continued subsiding at a rate of 11.4 mm/year. Fifth, Sabancaya [78], of which only an edge of deformation is presented in our data, has erupted from 2015 to 2022 (most of our study period), deforming at a rate of 29.6 mm/year. In addition, while the Uturuncu volcano is close to our study region and has been reported to continue its deformation [11], [79], it is not contained in our dataset.

## VI. DISCUSSION

### A. Evaluation of Learned Features

To better understand the model and the datasets and to identify what the problematic regions found in the geographic demonstration have in common, feature visualization is used. It is performed by visually comparing the similarity of the distributions of extracted latent features from the two sets. We use the last convolutional layer output as the extracted features, with a dimension of 1536. To visualize it in a low-dimensional space and keep its original modality, a widely used nonparametric nonlinear dimension reduction tool t-SNE [35] is employed.

First, the FT and RD test set latent feature representations are compared. From the second last layer of a fine-tuned

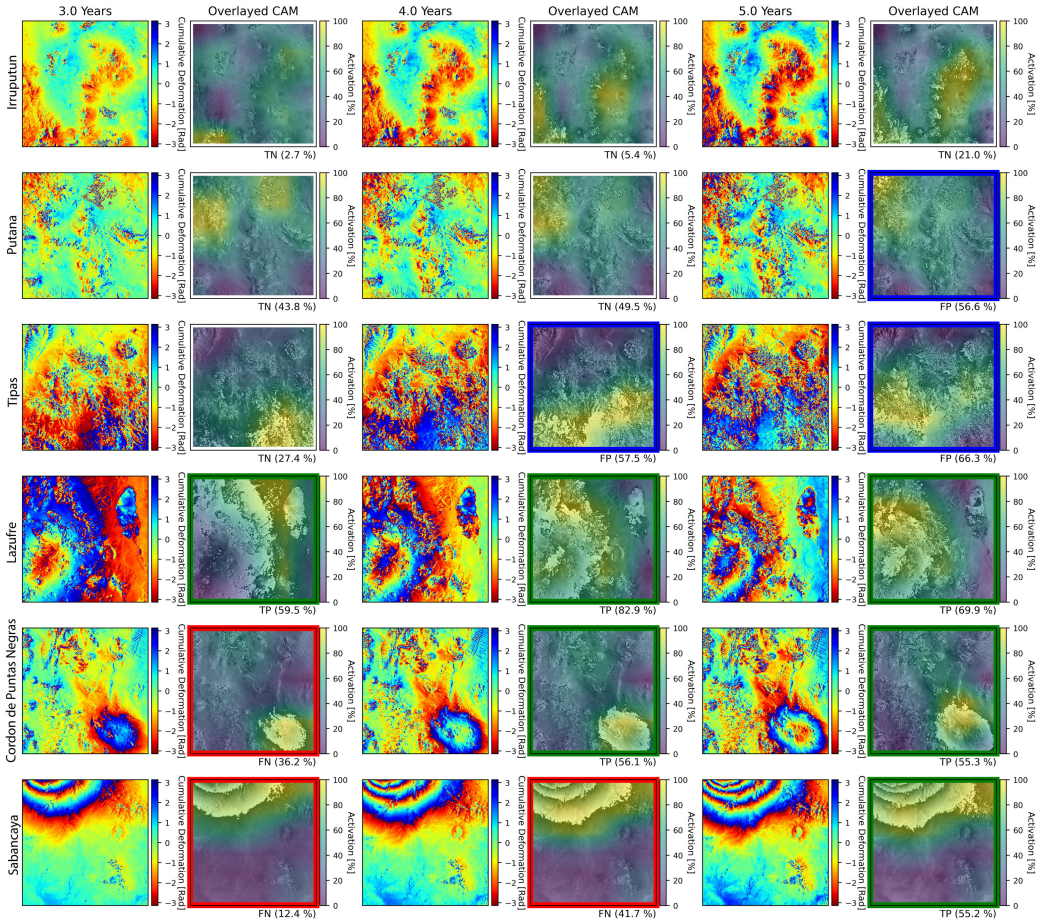



Fig. 15. Results of six cases and their GradCAM related to temporal baseline change. The top cases show the slope-induced signals, some of which start being detected as FPs by increasing the temporal baseline. The bottom three cases show volcanic deformations that benefit from an increase in temporal baseline, becoming detectable by the model. Below each GradCAM, the modeled probability of volcanic deformations and the classification result are shown. TP, FP, FN, and TN represent true-positive, false-positive, false-negative, and true-negative detections, respectively.

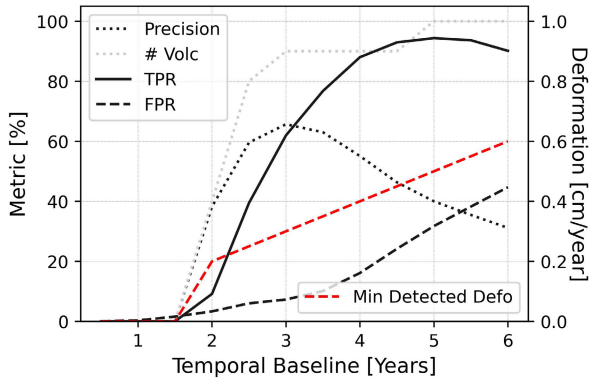



Fig. 16. Sensitivity analysis showing model performances with the change of temporal baseline. # Volc shows the ratio of frames containing volcanic deformations with at least one patch flagged as positive. Min Detected Defo is taken from the bottom row in Table III.

model, we extract 1536 latent features of each patch. These features represent the high-level visual features of the data. The latent features are then coded to two dimensions using the t-SNE technique [36], [37], as shown in Figs. 13 and 14. While the FT set is similar to the RD set, they have slightly different distributions. Most of the volcanic deformations in the FT set are grouped, containing only a couple of RD set samples in each cluster. To a lesser extent, the same goes for the RD test set volcanic deformations, implying an observable difference between the sets. It is also noticeable that, besides the patterns, the intensity of the deformations plays a significant role in grouping the data.

Feature analysis verifies and expands the data's geographic visualization findings. Fig. 13 shows that the FT set has fewer slope-induced signals and a smoother background than the RD set because of the elevation component [53] coming from the residual tropospheric noise. The elevation component has a higher intensity along the mountain ridges because of the more significant elevation differences and the coarse resolution of the atmospheric correction data. Since the background of the FT set is collected surrounding the volcanic region but never overlapping it, it typically does not cover the mountain ridge and, therefore, the more prominent elevation component. As a result, the elevation component seems to confuse the model.

Theoretically, the elevation (stratified) noise component should be sufficiently excluded by the applied atmospheric corrections and PSI processing [60]. In contrast, the residual noise patterns in our data seem to correlate to terrain elevation. In addition, it has been shown that residual tropospheric noise cannot be removed entirely in some volcanic

regions [53], [80], thus creating difficulties for simulations and CNN training. Furthermore, the residual elevation noise is exaggerated by increasing the temporal baseline to ≈5 years and poses a problem when detecting subtle deformations reflected in false-positive detections.

### B. Sensitivity Analysis

Only four fully present deforming volcanoes exist in the RD test set, and the lowest deformation rate is 5 mm/year. Therefore, we cannot fully explore the model's capabilities using the RD set only. Thus, we use a sensitivity analysis to expose the model to more data and find the minimal detectable volcanic deformation. It is done by testing the trained model with the rescaled and rewrapped cumulative deformation maps using different pseudotemporal baselines from one to five years, by increments of half a year. The original RD test set has a temporal baseline of ≈5 years wrapped cumulative deformation. The minimal detectable deformation is observed only on confirmed volcanic deformations. The scale of volcanic deformations is the maximum pixel value in volcanic deformation extracted from InSAR observations of each volcano.

Because of the fringes created by the wrapping process, the data with changed temporal baseline do not only transform in intensity but also in spatial features without affecting the SNR. Reducing the intensity of slope-induced signal or volcanic deformation signal can reposition or remove fringes, creating different conditions for classification, as seen in Figs. 15 and 16. An increase in intensity (via temporal baseline) leads to a rise in positive classifications and, hence, more true-positive detections and vice versa. As tested with these data, we find that the model can detect deformations down to 2 mm/year, as shown in Table III and Fig. 16.

The 2-mm/year limitation in volcanic deformations scale detections appears because of the significance of deformation intensity in decision-making, set by our synthetic training set and identified in the t-SNE analysis. In addition, when tested on the RD with the adjusted temporal baseline of four years, the FT model increases AUC ROC to 93.2% (see Fig. 8). This result confirms the t-SNE analysis findings that residual background noise is lower in the FT set.

## VII. Conclusion

The state-of-the-art in the field has been overviewed and categorized according to the approaches taken. The main challenges in detecting volcanic deformations in InSAR data relate to the atmospheric noise and the lack of data for model training. To respond to these challenges, we use the velocity maps for increased detail and reduction in atmospheric noise effect and synthetic data for model training. In addition, the variogram modeling was used to statistically extract the parameters used for simulations of residual atmospheric noise, thus exactly covering the whole range of residual noise present in the real test data.

A significant amount of the terrain deformation signal has accumulated over the observation period. Slope-induced signals and salt lakes are the region's most frequent terrain deformation patterns. These signals create difficulties for simulations and DL model training. They are exaggerated by increasing the temporal baseline to five years and play a significant role when detecting subtle deformations. In addition, while, theoretically, the topological component should be excluded by the applied atmospheric corrections and PSI processing, a noise pattern is present in this work, correlating to terrain elevation.

We proposed a DL model for detecting subtle long-term volcanic deformation, which has not been previously explored. Our model achieved AUC ROC values similar to that found in other state-of-the-art short-term deformation detection models [22], [24], [38] while detecting more than an order of magnitude smaller deformations. All of the volcanic deformations in the test set are detected, the smallest one being 5 mm/year. These results are extended by the sensitivity analysis, which shows that the minimal detectable deformation is in the range of 2 mm/year. It is an improvement over previous work, having the smallest detectable deformation of 35 mm/year [39].

The following phase is to utilize the trained model on a larger scale. Adding a residual elevation component [53], [80] to the synthetic set would make the model more resilient and generalizable to wider regions and reduce the false-positive rate. In addition, the data-centric analysis conducted in this article can be utilized to uncover other region-specific nonvolcanic deformation patterns. Further improvements in the simulation of these nonvolcanic patterns should boost the detection accuracy.

## Acknowledgment

This article is based on research performed as a part of the TecVolSA project. Modified Copernicus Sentinel data are used. Persistent scatterer interferometry (PSI)/distributed scatterer interferometry (DSI) processing is performed with the integrated wide area processor (IWAP) by the InSAR Team, German Aerospace Center (DLR). The authors especially thank Robert Shau and Victor Diego Navarro Sanchez for their patience and help regarding the data. They also especially thank the German Research Center for Geosciences (GFZ) and especially Rene Mania for the simulation of volcanic deformations and Thomas Walter for their help with the domain knowledge on volcano deformations and regional geologic processes. The manuscript would not be the same without Beatrice Hildebrand's help. For support with their discussions and advice on deep learning (DL), they thank Yuanyuan Wang, Syed Mohsin Ali, and Ivica Obadić.

## References

[1] E. Venzke, "Volcanoes of the World (v 4.9.1; Feb 2022)," Global Volcanism Program, 2023. Distributed by Smithsonian Institution, doi: 10.5479/si.GVP.VOTW4-2013.

[2] J. Biggs et al., "Global link between deformation and volcanic eruption quantified by satellite imagery," *Nature Commun.*, vol. 5, no. 1, pp. 1–7, Apr. 2014.

[3] M. A. Furtney et al., "Synthesizing multi-sensor, multi-satellite, multi-decadal datasets for global volcano monitoring," *J. Volcanol. Geothermal Res.*, vol. 365, pp. 38–56, Oct. 2018.

[4] T. Giachetti, R. Paris, K. Kelfoun, and B. Ontowirjo, "Tsunami hazard related to a flank collapse of Anak Krakatau Volcano, Sunda Strait, Indonesia," *Geol. Soc., London, Special Publications*, vol. 361, no. 1, pp. 79–90, Jan. 2012.

[5] R. Williams, P. Rowley, and M. C. Garthwaite, "Reconstructing the Anak Krakatau flank collapse that caused the December 2018 Indonesian Tsunami," *Geology*, vol. 47, no. 10, pp. 973–976, Oct. 2019.

[6] M. Manga and E. Brodsky, "Seismic triggering of eruptions in the far field: Volcanoes and geysers," *Annu. Rev. Earth Planet. Sci.*, vol. 34, no. 1, pp. 263–291, May 2006.

[7] R. Scarpa and P. Gasparini, "A review of volcano geophysics and volcano-monitoring methods," in *Monitoring and Mitigation of Volcano Hazards*. Berlin, Germany: Springer, 1996, pp. 3–22, doi: 10.1007/978-3-642-80087-0_1.

[8] R. Lara, D. Benítez, A. Caamaño, M. Zennaro, and J. L. Rojo-Álvarez, "On real-time performance evaluation of volcano-monitoring systems with wireless sensor networks," *IEEE Sensors J.*, vol. 15, no. 6, pp. 3514–3523, Jun. 2015.

[9] M. E. Pritchard and M. Simons, "An InSAR-based survey of volcanic deformation in the southern Andes," *Geophys. Res. Lett.*, vol. 31, no. 15, 2004, Art. no. L15610.

[10] J. Biggs, I. D. Bastow, D. Keir, and E. Lewi, "Pulses of deformation reveal frequently recurring shallow magmatic activity beneath the Main Ethiopian Rift," *Geochem., Geophys., Geosyst.*, vol. 12, no. 9, Sep. 2011, Art. no. Q0AB10.

[11] S. T. Henderson and M. E. Pritchard, "Decadal volcanic deformation in the central Andes volcanic zone revealed by InSAR time series," *Geochem., Geophys., Geosyst.*, vol. 14, no. 5, pp. 1358–1374, May 2013.

[12] D. Massonnet and K. L. Feigl, "Radar interferometry and its application to changes in the Earth's surface," *Rev. Geophys.*, vol. 36, no. 4, pp. 441–500, Nov. 1998.

[13] A. Ferretti et al., "Submillimeter accuracy of InSAR time series: Experimental validation," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1142–1153, May 2007.

[14] A. Ferretti, C. Prati, and F. Rocca, "Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2202–2212, Sep. 2000.

[15] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci, "A new algorithm for processing interferometric data-stacks: SqueeSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3460–3470, Sep. 2011.

[16] M. Crosetto, O. Monserrat, M. Cuevas-González, N. Devanthéry, and B. Crippa, "Persistent scatterer interferometry: A review," *ISPRS J. Photogram. Remote Sens.*, vol. 115, pp. 78–89, May 2016.

[17] Y. Wang and X. X. Zhu, "Robust estimators for multipass SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 968–980, Feb. 2016.

[18] H. Ansari, F. De Zan, and R. Bamler, "Sequential estimator: Toward efficient InSAR time series analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5637–5652, Oct. 2017.

[19] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 37–44, Jan. 2005.

[20] X. X. Zhu and R. Bamler, "Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 735–739, Jul. 2011.

[21] X. X. Zhu et al., "Deep learning meets SAR: Concepts, models, pitfalls, and perspectives," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 4, pp. 143–172, Dec. 2021.

[22] N. Anantrasirichai, J. Biggs, F. Albino, P. Hill, and D. Bull, "Application of machine learning to classification of volcanic deformation in routinely generated InSAR data," *J. Geophys. Res., Solid Earth*, vol. 123, no. 8, pp. 6592–6606, Aug. 2018.

[23] N. Anantrasirichai, F. Albino, P. Hill, D. Bull, and J. Biggs, "Detecting volcano deformation in InSAR using deep learning," 2018, *arXiv:1803.00380*.

[24] M. Gaddes, A. Hooper, and F. Albino, "Simultaneous classification and location of volcanic deformation in SAR interferograms using deep learning and the VolcNet database," Tech. Rep., 2021, doi: 10.31223/X5CW2J.

[25] B. Rouet-Leduc, R. Jolivet, M. Dalaison, P. A. Johnson, and C. Hulbert, "Autonomous extraction of millimeter-scale deformation in InSAR time series using deep learning," *Nature Commun.*, vol. 12, no. 1, pp. 1–11, Nov. 2021.

[26] J. Sun et al., "Automatic detection of volcanic surface deformation using deep learning," *J. Geophys. Res., Solid Earth*, vol. 125, no. 9, Sep. 2020, Art. no. e2020JB019840.

[27] M. E. Gaddes, A. Hooper, M. Bagnardi, H. Inman, and F. Albino, "Blind signal separation methods for InSAR: The potential to automatically detect and monitor signals of volcanic deformation," *J. Geophys. Res., Solid Earth*, vol. 123, no. 11, pp. 10–226, Nov. 2018.

[28] S. Valade et al., "Towards global volcano monitoring using multisensor sentinel missions and artificial intelligence: The MOUNTS monitoring system," *Remote Sens.*, vol. 11, no. 13, p. 1528, Jun. 2019.

[29] J. Biggs et al., "Baseline monitoring of volcanic regions with little recent activity: Application of Sentinel-1 InSAR to Turkish volcanoes," *J. Appl. Volcanol.*, vol. 10, no. 1, pp. 1–14, Dec. 2021.

[30] H. Ansari, F. De Zan, and R. Bamler, "Efficient phase estimation for interferogram stacks," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 7, pp. 4109–4125, Jul. 2018.

[31] H. Ansari, F. De Zan, and R. Bamler, "Distributed scatterer interferometry tailored to the analysis of big InSAR data," in *Proc. 12th Eur. Conf. Synth. Aperture Radar (EUSAR)*. Frankfurt, Germany: VDE, Jun. 2018, pp. 1–5. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8438005

[32] H. Ansari, F. De Zan, G. Gomba, and R. Bamler, "EMI: Efficient temporal phase estimation and its impact on high-precision InSAR time series analysis," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2019, pp. 270–273.

[33] A. Parizzi et al., "Processing and performance of large scale deformation monitoring with InSAR stacks," *Geophys. Res. Abstr.*, vol. 21, p. 1, Jan. 2019.

[34] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.

[35] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, 2008. [Online]. Available: http://www.jmlr.org/papers/v9/vandermaaten08a.html

[36] J.-Y. Kim and S.-B. Cho, "Electric energy consumption prediction by deep learning with state explainable autoencoder," *Energies*, vol. 12, no. 4, p. 739, Feb. 2019.

[37] P. Lübbecke, N. Mehdiyev, and P. Fettke. (2019). *Substitution of Hazardous Chemical Substances Using Deep Learning and t-SNE*. [Online]. Available: https://aisel.aisnet.org/wi2019/track12/papers/5

[38] N. Anantrasirichai, J. Biggs, F. Albino, and D. Bull, "A deep learning approach to detecting volcano deformation from satellite imagery using synthetic datasets," *Remote Sens. Environ.*, vol. 230, Sep. 2019, Art. no. 111179.

[39] N. Anantrasirichai, J. Biggs, F. Albino, and D. Bull, "The application of convolutional neural networks to detect slow, sustained deformation in InSAR time series," *Geophys. Res. Lett.*, vol. 46, no. 21, pp. 11850–11858, Nov. 2019.

[40] M. E. Gaddes, A. Hooper, and M. Bagnardi, "Using machine learning to automatically detect volcanic unrest in a time series of interferograms," *J. Geophys. Res., Solid Earth*, vol. 124, no. 11, pp. 12304–12322, Nov. 2019.

[41] B. Silva, J. J. Sousa, M. Lazecky, and A. Cunha, "Deformation fringes detection in SAR interferograms using deep learning," *Proc. Comput. Sci.*, vol. 196, pp. 151–158, Jan. 2022.

[42] C. M. J. Brengman and W. D. Barnhart, "Identification of surface deformation in InSAR using machine learning," *Geochem., Geophys., Geosyst.*, vol. 22, no. 3, Mar. 2021, Art. no. e2020GC009204.

[43] N. I. Bountos, I. Papoutsis, D. Michail, and N. Anantrasirichai, "Self-supervised contrastive learning for volcanic unrest detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[44] N. I. Bountos, D. Michail, and I. Papoutsis, "Learning from synthetic InSAR with vision transformers: The case of volcanic unrest detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4509712.

[45] N. I. Bountos, I. Papoutsis, D. Michail, A. Karavias, P. Elias, and I. Parcharidis, "Hephaestus: A large scale multitask dataset towards InSAR understanding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2022, pp. 1453–1462.

[46] F. Albino, J. Biggs, C. Yu, and Z. Li, "Automated methods for detecting volcanic deformation using Sentinel-1 InSAR time series illustrated by the 2017–2018 unrest at Agung, Indonesia," *J. Geophys. Res., Solid Earth*, vol. 125, no. 2, Feb. 2020, Art. no. e2019JB017908.

[47] A. Parizzi, R. Brcic, and F. De Zan, "InSAR performance for large-scale deformation measurement," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8510–8520, Oct. 2021.

[48] L. A. Jeni, J. F. Cohn, and F. D. L. Torre, "Facing imbalanced data-recommendations for the use of performance metrics," in *Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact.*, Sep. 2013, pp. 245–251.

[49] S. L. De Silva and P. Francis, *Volcanoes of the Central Andes*, vol. 219. Berlin, Germany: Springer-Verlag, 1991.

[50] D. J. Varnes, "Slope movement types and processes," *Special Rep.*, vol. 176, no. 176, pp. 11–33, 1978. [Online]. Available: http://onlinepubs.trb.org/Onlinepubs/sr/sr176/176-002.pdf

[51] A. J. T. Guerra, M. A. Fullen, M. D. C. O. Jorge, J. F. R. Bezerra, and M. S. Shokr, "Slope processes, mass movement and soil erosion: A review," *Pedosphere*, vol. 27, no. 1, pp. 27–41, Feb. 2017.

[52] M. Eineder, C. Minet, P. Steigenberger, X. Cong, and T. Fritz, "Imaging geodesy—Toward centimeter-level ranging accuracy with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 661–671, Feb. 2011.

[53] X. Cong, "SAR interferometry for volcano monitoring: 3D-PSI analysis and mitigation of atmospheric refractivity," Ph.D. dissertation, Dept. Civil, Geo Environ. Eng., Technische Universität München, Munich, Germany, 2014. [Online]. Available: http://nbn-resolving.org/urn:nbn:de:bvb:91-diss-20140606-1198042-0-8

[54] C. Gisinger, "Sar imaging geodesy-towards absolute coordinates with centimeter accuracy," Ph.D. dissertation, Dept. Civil, Geo Environ. Eng., Technische Universität München, Munich, Germany, 2019.

[55] H. Ansari, F. De Zan, and A. Parizzi, "Study of systematic bias in measuring surface deformation with SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1285–1301, Feb. 2021.

[56] B. M. Kampes, *Radar Interferometry*, vol. 12. Dordrecht, The Netherlands: Springer, 2006, doi: 10.1007/978-1-4020-4723-7.

[57] G. Petit et al., "IERS conventions," *IERS Tech. Note*, vol. 36, no. 1, p. 2010, 2010.

[58] M. Eineder, U. Balss, S. Suchandt, C. Gisinger, X. Cong, and H. Runge, "A definition of next-generation SAR products for geodetic applications," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 1638–1641.

[59] M. Even and K. Schulz, "InSAR deformation analysis with distributed scatterers: A review complemented by new advances," *Remote Sens.*, vol. 10, no. 5, p. 744, May 2018.

[60] J. Jung, D.-J. Kim, and S.-E. Park, "Correction of atmospheric phase screen in time series InSAR using WRF model for monitoring volcanic activities," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2678–2689, May 2014.

[61] A. M. Hoyt, E. Chaussard, S. S. Seppalainen, and C. F. Harvey, "Widespread subsidence and carbon emissions across Southeast Asian peatlands," *Nature Geosci.*, vol. 13, no. 6, pp. 435–440, Jun. 2020.

[62] M. Nikkhoo, T. R. Walter, P. R. Lundgren, and P. Prats-Iraola, "Compound dislocation models (CDMs) for volcano deformation analyses," *Geophys. J. Int.*, vol. 208, no. 2, Feb. 2017, Art. no. ggw427.

[63] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[64] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 1–7.

[65] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[66] S. Targ, D. Almeida, and K. Lyman, "ResNet in ResNet: Generalizing residual architectures," 2016, *arXiv:1603.08029*.

[67] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[68] T. Beker, H. Ansari, S. Montazeri, Q. Song, and X. X. Zhu, "Fine-tuning CNNs for decreased sensitivity to non-volcanic deformation velocity signals," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 3, pp. 85–92, May 2022.

[69] W. Samek, T. Wiegand, and K.-R. Müller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models," 2017, *arXiv:1708.08296*.

[70] J. Adebayo, J. Gilmer, M. Muelly, I. Goodfellow, M. Hardt, and B. Kim, "Sanity checks for saliency maps," 2018, *arXiv:1810.03292*.

[71] I. Kakogeorgiou and K. Karantzalos, "Evaluating explainable artificial intelligence methods for multi-label deep learning classification tasks in remote sensing," 2021, *arXiv:2104.01375*.

[72] N. Tajbakhsh et al., "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.

[73] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 1, pp. 1–40, Dec. 2016.

[74] S. K. E. Sennert. (2013). *Report on Antuco (Chile)*. Accessed: Apr. 18, 2022. [Online]. Available: https://volcano.si.edu/showreport.cfm?doi=GVP.WVAR20130424-357080

[75] P. Grosse, Y. Orihashi, S. R. Guzmán, H. Sumino, and K. Nagao, "Eruptive history of Incahuasi, Falso Azufre and El Cóndor Quaternary composite volcanoes, southern central Andes," *Bull. Volcanol.*, vol. 80, no. 5, pp. 1–26, May 2018.

[76] D. Remy et al., "Persistent uplift of the Lazufre volcanic complex (Central Andes): New insights from PCAIM inversion of InSAR time series and GPS data," *Geochem., Geophys., Geosyst.*, vol. 15, no. 9, pp. 3591–3611, Sep. 2014.

[77] C. A. Brunori, C. Bignami, S. Stramondo, and E. Bustos, "20 years of active deformation on volcano caldera: Joint analysis of InSAR and AInSAR techniques," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 23, pp. 279–287, Aug. 2013.

[78] P. MacQueen et al., "Volcano-tectonic interactions at Sabancaya volcano, Peru: Eruptions, magmatic inflation, moderate earthquakes, and fault creep," *J. Geophys. Res., Solid Earth*, vol. 125, no. 5, 2020, Art. no. e2019JB019281.

[79] N. Lau, E. Tymofyeyeva, and Y. Fialko, "Variations in the long-term uplift rate due to the Altiplano–Puna magma body observed with Sentinel-1 interferometry," *Earth Planet. Sci. Lett.*, vol. 491, pp. 43–47, Jun. 2018.

[80] K. J. Stephens, C. Wauthier, R. C. Bussard, M. Higgins, and P. C. LaFemina, "Assessment of mitigation strategies for tropospheric phase contributions to InSAR time-series datasets over two nicaraguan volcanoes," *Remote Sens.*, vol. 12, no. 5, p. 782, Mar. 2020.

**Teo Beker** (Member, IEEE) received the B.Sc. and M.Sc. degrees in geodesy and geomatics from the University of Novi Sad, Novi Sad, Serbia, in 2014 and 2019, respectively, and the M.Sc. degree in geoinformatics from the Politecnico di Milano, Milan, Italy, in 2018. He is currently pursuing the Ph.D. degree with the Chair of Data Science in Earth Observation (SiPEO), Technical University of Munich (TUM), Munich, Germany and the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Weßling, Germany.

He worked as a Research Associate at the Institute of Lowland Forestry and Environment, Novi Sad, from 2018 to 2020, applying deep learning for habitat suitability studies. His research interests include applied machine and deep learning, explainable artificial intelligence (AI), remote sensing, and geospatial data analysis.

**Homa Ansari** received the B.Sc. degree in geodesy and geodetic engineering from the University of Isfahan, Isfahan, Iran, in 2011, and the M.Sc. degree (cum laude) in Earth-oriented space science and technology, specializing in precise global navigation satellite systems and remote sensing, and the Ph.D. degree in radar remote sensing from the Technical University of Munich (TUM), Munich, Germany, in 2013 and 2018, respectively.

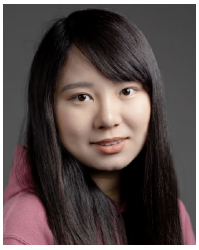
In 2012, she joined the German Geodetic Research Institute, Munich, as a Research Assistant. In 2013, she started as a Research Assistant at the German Aerospace Center (DLR), Weßling, Germany. Since 2014, she has been with the Earth Observation Center, Remote Sensing Technology Institute, DLR. Since 2014, her work at DLR has been contributing to the German Helmholtz Alliance on Remote Sensing and Earth System Dynamics and the performance studies of the Tandem-L mission. In 2017, she visited the Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands, as a Visiting Researcher. Her work is focused on the design and accuracy assessment of efficient algorithms for processing big data from modern synthetic aperture radar (SAR) missions. Her research interests include the application of statistical signal processing, estimation theory, geodetic time-series analysis, machine learning, and deep neural networks to differential SAR interferometry.

Dr. Ansari was a recipient of the IEEE Mikio Takagi Prize, awarded for the Best Student Paper at the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Texas, USA, in 2017. The prize was awarded to a novel algorithm tailored to efficient time-series analysis for SAR interferometry.

**Sina Montazeri** received the B.Sc. degree in geodetic engineering from the University of Isfahan, Isfahan, Iran, in 2011, the M.Sc. degree in geomatics from the Delft University of Technology, Delft, The Netherlands, in 2014, and the Ph.D. degree in radar remote sensing from the Technical University of Munich (TUM), Munich, Germany, in 2019, with a dissertation on geodetic synthetic aperture radar (SAR) interferometry.

In 2012, he spent two weeks at the Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection, University of Strasbourg, Strasbourg, France, as a Junior Researcher working on thermal remote sensing. From 2013 to 2015, he was a Research Assistant with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Weßling, Germany, where he was involved in the absolute localization of point clouds obtained from SAR tomography. From 2015 to 2019, he was a Research Associate with the Signal Processing in Earth Observation Research Group, TUM, and IMF, DLR, where he worked on the automatic positioning of ground control points from multiview radar images. He is currently a Senior Researcher with the Department of Earth Observation Data Science, IMF, DLR, where he has been focusing on developing machine learning algorithms applied to radar imagery. His research interests include advanced interferometric SAR techniques for the deformation monitoring of urban infrastructure, image and signal processing relevant to radar imagery, and applied machine learning.

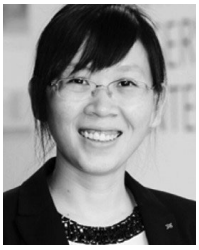
Dr. Montazeri received the DLR Science Award and the IEEE Geoscience and Remote Sensing Society Transactions Prize Paper Award in 2016 and 2017, respectively, for his work on geodetic SAR tomography.



**Qian Song** (Member, IEEE) received the B.E. degree (Hons.) from the School of Information Science and Technology, East China Normal University, Shanghai, China, in 2015, and the Ph.D. degree (Hons.) from Fudan University, Shanghai, in 2020.

From 2020 to 2022, she was a Post-Doctoral Fellow with the Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Weßling, Germany. Since 2023, she has been a Post-Doctoral Fellow with the Chair of Data Science in Earth Observation, Technical University of Munich (TUM), Munich, Germany. Her research interests include advanced deep learning technologies and their applications in synthetic aperture radar image interpretation and forest monitoring.

Dr. Song was awarded the International Union of Radio Science (URSI) Young Scientist Award in 2020.



**Xiao Xiang Zhu** (Fellow, IEEE) received the M.Sc., Dr.Ing., and Habilitation degrees in signal processing from the Technical University of Munich (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

She is the Chair Professor for Data Science in Earth Observation at Technical University of Munich (TUM) and was the founding Head of the Department "EO Data Science" at the Remote Sensing Technology Institute, German Aerospace Center (DLR). She is also a Visiting AI Professor at ESA's Phi-lab. Since May 2020, she is the PI and Director of the international future AI lab "AI4EO – Artificial Intelligence for Earth Observation: Reasoning, Uncertainties, Ethics and Beyond", Munich, Germany. Since October 2020, she also serves as the Director of the Munich Data Science Institute (MDSI), Neuherberg, Germany, TUM. From 2019 to 2022, Zhu has been a co-coordinator of the Munich Data Science Research School (www.mu-ds.de) and the head of the Helmholtz Artificial Intelligence – Research Field "Aeronautics, Space and Transport". She was a Guest Scientist or Visiting Professor at the Italian National Research Council (CNR-IREA), Naples, Italy, Fudan University, Shanghai, China, the University of Tokyo, Tokyo, Japan and University of California, Los Angeles, United States in 2009, 2014, 2015 and 2016, respectively. Her main research interests are remote sensing and Earth observation, signal processing, machine learning, and data science, with their applications in tackling societal grand challenges, e.g., Global Urbanization, UN's SDGs, and Climate Change.

Dr. Zhu was a member of the Young Academy (Junge Akademie/Junges Kolleg) at the Berlin-Brandenburg Academy of Sciences and Humanities and the German National Academy of Sciences Leopoldina and the Bavarian Academy of Sciences and Humanities. She serves on the Scientific Advisory Board in several research organizations, including the German Research Center for Geosciences (GFZ, 2020-2023) and the Potsdam Institute for Climate Impact Research (PIK), among others. She is an Associate Editor of IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, PATTERN RECOGNITION. She serves as an Area Editor responsible for special issues of *IEEE Signal Processing Magazine*.