

Automatic detection of volcanic surface deformation using deep learning

Jian Sun^{1,4}, Christelle Wauthier^{1,4}, Kirsten Stephens¹, Melissa Gervais^{2,4},
Guido Cervone^{2,3,4}, Peter La Femina¹, Machel Higgins¹

¹Department of Geosciences, The Pennsylvania State University, University Park, PA, USA

²Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University
Park, PA, USA

³Department of Geography, The Pennsylvania State University, University Park, PA, USA

⁴Institute for Computational and Data Sciences, The Pennsylvania State University, University Park, PA,
USA

Key Points:

- Novel application of CNN is capable of revealing both spatial and temporal dynamics of volcanic surface deformation.
- CNN automatically mitigates atmospheric noise and facilitates near real-time volcano monitoring.
- Results of the application of CNN to the Masaya volcano in Nicaragua, are consistent with continuous GPS data.

Abstract

Interferometric Synthetic Aperture Radar (InSAR) provides sub-centimetric measurements of surface displacements, which are key for characterizing and monitoring magmatic processes in volcanic regions. The abundant measurements of surface displacements in multitemporal InSAR data routinely acquired by SAR satellites can facilitate near real-time volcano monitoring on a global basis. However, the presence of atmospheric signals in interferograms complicates the interpretation of those InSAR measurements, which can even lead to a misinterpretation of InSAR signals and volcanic unrest. Given the vast quantities of SAR data available, an automatic InSAR data processing and denoising approach is required to separate volcanic signals that are cause of concern from atmospheric signals and noise. In this study, we employ a deep learning strategy that directly removes atmospheric and other noise signals from time-consecutive unwrapped surface displacements obtained through an InSAR time-series approach using an end-to-end Convolution Neural Network (CNN) with an encoder-decoder architecture, [Sun:modified U-net](#). The CNN is trained with simulated synthetic unwrapped surface displacement maps, and is then applied to real InSAR data. Our proposed architecture is capable of detecting dynamic spatio-temporal patterns of volcanic surface displacements. We find that an ensemble-average strategy is recommended to stabilize the detected results for varying deformation rates and signal-to-noise ratios (SNR). A case study is also presented where this method is applied to InSAR data covering the Masaya volcano in Nicaragua and the results are validated using continuous GPS data. The results confirm that our network can indeed efficiently suppress atmospheric and other noise to reveal the noise-free surface deformation.

1 Introduction

Volcanic unrest causes diverse natural hazards that can pose threats to local population, economy, infrastructures, and even trigger air traffic and climate perturbations globally. However, over 90 percent of volcanoes are not monitored consistently with ground-based networks (Loughlin et al., 2015). Massive Interferometric Synthetic Aperture Radar (InSAR) datasets routinely acquired by Synthetic Aperture Radar (SAR) satellites, such as Sentinel-1 or COSMO-SkyMed, provide an opportunity to globally measure surface displacements in active volcanic regions. While InSAR has the capacity to obtain subtle (mm to cm) geodetic measurements at volcanoes (Massonnet & Feigl, 1998; Hanssen, 2001), there are a number of challenges that complicate the interpretation of InSAR measurements. For example, coherent deformation signals that are cause of concern for volcanic unrest must be separated from other signals that can otherwise be misinterpreted as deformation including orbit, topography phase residuals, and atmospheric noise. Atmospheric phase delay signals in InSAR products are particularly challenging to mitigate due to the complex spatial and temporal variations in atmospheric parameters such as hydrostatic pressure, temperature, and concentrations of water vapor (Hanssen, 2001; Emardson et al., 2003; Remy et al., 2003). Stratified tropospheric phase contributions and turbulent atmospheric effects may greatly complicate the interpretation of deformation signals (Doin et al., 2009; Ebmeier, 2016; K. J. Stephens et al., 2020). In order to conduct near real-time volcano monitoring on a global basis, an automatic processing tool for detecting volcanic deformation signals through denoising interferograms and resulting InSAR time-series products is necessary.

Deep learning techniques have drawn significant attention across various fields due to their demonstrated skill for conducting complex tasks, especially those involving image processing, such as classification (Krizhevsky et al., 2012; Simonyan & Zisserman, 2014), segmentation (Long et al., 2015; Ronneberger et al., 2015; Häggström et al., 2019), denoising (Zhang et al., 2017), and superresolution (Kim et al., 2016). Efforts have been made to apply these data-driven methods to natural hazard classification and forecast-

ing using either seismic events or large numbers of SAR images. For example, Ebmeier (2016) and Gaddes et al. (2019) demonstrated that atmospheric noise and surface displacements in volcanic regions can be separated using a combination of traditional machine learning algorithms, independent component analysis (ICA) and cluster analysis. The advantage of ICA is that it is fully data-driven and thus does not require any prior information. However, similar to other linear signal separating methods, the order-agnostic of ICA marginalizes out either the spatial or temporal dynamics of the data depending on the assumption of independence. Thus, ^{Sum:}either spatial or temporal ICA can only allow for variability analysis in either spatially or temporally, but not both of them (Calhoun et al., 2001). ^{Sum:}Spatiotemporal ICA (stICA) is proposed to maximize the degree of independence of both spatial and temporal by allowing a tradeoff of the mutual independence of images and their corresponding time courses (Stone et al., 2002; Shi et al., 2018). ^{Sum:}Though stICA is powerful, it requires an expensive computational cost (Boergens et al., 2014).

There have been recent attempts at detecting volcanic surface deformation using modern deep neural networks. ^{Sum:}A large scale benchmark archive for remote sensing image classification were discussed in Sumbul et al. (2019). To improve the speed of manually distinguishing the volcanic deformation (i.e., fringes) from wrapped interferograms, Anantrasirichai et al. (2018, 2019) showed that interferometric fringes can be classified using a transfer learning strategy with AlexNet (Krizhevsky et al., 2012). However, atmospheric signals may also appear as “fringes” in wrapped interferograms, which would then be classified as volcanic deformation signals and lead to false positive identifications. ^{Sum:}Using variant combinations of training dataset and applying stratified atmospheric correction in advance may reduce the risk of false positives to some extent, and the detailed discussions can be found in Anantrasirichai et al. (2019). Atmospheric corrections are therefore still required to retrieve the noise-free surface displacements and those corrections remains a significant challenge (K. J. Stephens et al., 2020).

Unlike recent deep learning works applied to wrapped interferograms, in this study, we design a new approach using a deep neural network to extract the signal of interest, i.e., volcanic surface deformation, to distinguish the noise-free surface displacements from atmospheric signals or other sources of noise, and to automatically detect signs of unrest from a series of time-consecutive unwrapped surface displacement maps obtained through conventional InSAR or InSAR time-series. In section two, we briefly discuss the input InSAR data, the mathematical framework of our deep learning method, and the architecture proposed for volcanic surface displacements detection. In section three, we introduce the requirement of data preprocessing, training, and implementation strategy. In addition, the temporal and SNR dependency of our proposed method are also discussed. In section four, we implement the proposed architecture on a real InSAR data and examine our results’ consistency with continuous Global Positioning System (GPS) time-series data.

2 Methodology

2.1 Input InSAR Data

Conventional differential InSAR computes the phase difference between two complex SAR images spanning the same area but acquired at two different times (Massonnet & Feigl, 1998). An interesting extension of conventional InSAR are multi-temporal InSAR time-series techniques. InSAR time-series techniques involve the processing of multiple SAR acquisitions over the same area, allowing for the correction of uncorrelated phase noise terms, and reducing errors associated with the surface deformation measurements. Two classes of InSAR time-series algorithms are persistent scatterers (Ferretti et al., 2001; Kampes, 2005; Hooper, 2008) and Small Baseline Subset (SBAS) methods (Lundgren et al., 2001; Berardino et al., 2002; Lanari et al., 2007). ^{Sum:}One of the products obtained

~~through InSAR time series are time-consecutive unwrapped surface displacement maps. We will use such products.~~ In this paper, the InSAR time-series is performed using SBAS methods to obtain time-consecutive unwrapped surface displacement maps, which are used as input of our newly designed Convolution Neural Network (CNN).

2.2 Mathematical Framework

In this study, we consider volcanic surface deformation retrievals as a denoising problem. CNNs has been demonstrated to outperform many other framelets when applied to denoising problems (Ronneberger et al., 2015; Kim et al., 2016; Zhang et al., 2017) and is therefore a favorable framework of deep learning techniques to apply for this study. The deep learning mathematical expression for denoising time-consecutive unwrapped surface displacement maps can be written as:

$$\hat{\Phi}_{def} = \mathcal{F}(\Phi; \Theta) \quad (1)$$

where Φ_{def} is a time-series of m noise-free surface displacements with a fixed temporal interval, Φ is a time-series of m corresponding unwrapped surface displacement maps, Θ delineates all trainable parameters of the neural network that are applied to establish the nonlinear relationship of the unwrapped maps Φ and the corresponded surface displacements Φ_{def} , and $\mathcal{F}$ represents the projection between the input Φ and the output Φ_{def} using trainable parameters Θ . In a deep CNN, $\mathcal{F}$ is usually a series of convolution blocks where each block is composed of several operations among convolution, activation mapping, batch normalization, dropout, max/average pooling, upsampling, and transposed convolution.

The convolution process is implemented as an iterative element-wise product of the input and a set of k trainable convolution kernels. Taking the first convolution layer as an example, the input contains m time-consecutive unwrapped maps of cumulative surface displacements placed in various channels. Considering k special convolution kernels of size 1×1 , convolutions of inputs $\text{Sun:}(\underline{m \times w \times h})$ and k kernels are equivalent to stacking these unwrapped maps in different weights, Sun: where the result is in size of $k \times w \times h$. Therefore, to a certain extent, the stacking method can be considered as a convolution process with a special kernel of size 1×1 , which could be applied to eliminate most of temporally-uncorrelated atmospheric noise. However, stacking will not mitigate the effect of spatially- or temporally-correlated atmospheric noise such as stratified water vapor signals (Doin et al., 2009; Ebmeier, 2016).

To simultaneously mitigate spatially-correlated atmospheric noise, convolution kernels with receptive fields > 1 (Moody & Darken, 1988) must be adopted. For instance, convolving time-consecutive unwrapped surface displacement maps with kernels of size 3×3 can be treated as stacking them in a localized manner. The maximum spatially-correlated distance can be calculated by the size of convolution kernel and the spatial resolution of surface displacement maps. Furthermore, challenges remain in conducting atmospheric phase corrections over topography that varies dramatically over short spatial scales. To fit the complexity of projection between time-consecutive unwrapped surface displacement maps and noise-free deformation signals, the non-linearity of Sun: a CNN is also increased by introducing normalization, dropout, and activation mapping operations.

The training stage of the network ensures all trainable parameters are learned appropriately to fit a non-linear mapping from the input, i.e., time-consecutive unwrapped surface displacement maps, to the output, i.e., their corresponding noise-free surface displacements. The optimal weights of the neural network can be obtained by Sun: solving the optimization equation: minimizing the discrepancy between the network's output and its corresponding ground truth. ~~where Φ^i is the i th input of CNN which contains m time-consecutive~~

unwrapped surface displacement maps, Φ_{def}^i is the ground-truth target, i.e., a set of m corresponding noise-free surface deformations, P is the number of pixels in an individual unwrapped map, N is the total number of input and target pairs, $\mathcal{L}(\cdot)$ is a measure of the discrepancy between the ground-truth Φ_{def}^i and predicted values $\hat{\Phi}_{def}^i$. In our case, we adopted $l1$ -norm, i.e., mean-absolute-error (MAE), to compute the prediction error and to evaluate the performance of the neural network.

The minimization of the objective function of the neural network is commonly solved with the backpropagation methods using a gradient-based optimization algorithm, such as stochastic gradient descent (Bottou, 2010, or SGD) and adaptive moment (Kingma & Ba, 2014, or Adam). Training this tremendous number of parameters requires a very large set of input and ground-truth target pairs. This makes the computation of the gradient over the entire training dataset an impracticable task due to the high computation time and memory cost. As in other deep learning tasks, a minibatch size of n is applied for obtaining the objective function and for computing the gradient for parameter updates. The $l1$ -norm objective function with a minibatch size of n is expressed as:

$$\hat{\Theta} = \arg \min_{\Theta} \frac{1}{mnP} \sum_{i=1}^n \mathcal{L}_n = \arg \min_{\Theta} \frac{1}{mnP} \sum_{i=1}^n \|\Phi_{def}^i - \mathcal{F}(\Phi^i; \Theta)\| \quad (2)$$

Sum: where Φ^i is the i th input of CNN which contains m time-consecutive unwrapped surface displacement maps, Φ_{def}^i is the ground-truth target, i.e., a set of m corresponding noise-free surface deformations, P is the number of pixels in an individual unwrapped map, N is the total number of input and target pairs, $\mathcal{L}(\cdot)$ is a measure of the discrepancy between the ground-truth Φ_{def}^i and predicted values $\hat{\Phi}_{def}^i$.

During the training process, the entire training dataset will be shuffled and divided into batches using the given batch size n at every epoch. *Sum:* For instance, a batch of 10 time series each of 26 interferograms with size 200×300 consists of a tensor in size of $10 \times 26 \times 200 \times 300$. One epoch of training is defined as passing the entire training dataset forward and backward through the neural network a single time. In our experiments, due to the intricacy of atmospheric signals, the minibatch SGD algorithm is applied to compute updates of parameters and to prevent the objective function from converging into local minima (Li et al., 2014). In the prediction stage, m time-consecutive unwrapped surface displacement maps will be fed into the neural network to extract their corresponded noise-free displacements of the Earth's surface. Because of the fixed time-length of the CNN input (i.e., allowing m time-consecutive unwrapped surface displacement maps only), one can iteratively implement the process with a certain fold if the number of time-consecutive unwrapped surface displacement maps is larger than m . *Sum:* For the data that has less than m number of temporal series, one can repeat the last temporal series achieved to meet the fixed time-length requirement of the CNN. However, it is a trade-off between the number of temporal series and the accuracy, which is discussed in the latter context.

2.3 Architecture of Neural Network

Various architectures of CNN have been designed and proven very effective for image processing tasks. For example, Simonyan and Zisserman (2014) proposed a CNN architecture called VGG16 that includes a convolutional encoder and a classifier *Sum:* using three fully connected layers for image recognition of 1000 classes. The encoder network in VGG16 includes 13 convolution layers and 5 max-pooling layers that take the image dimensional pixels as input *Sum:* with size $224 \times 224 \times 3$ and outputs feature mapping representations *Sum:* with size $7 \times 7 \times 512$ in a lower dimension with increased depth. The classifier, consisting of 4096 neurons in each layer, then interprets these feature representations generated by the encoder and outputs *Sum:* a single value as the predicted class the predicted probability for each class of the input image. The study of VGG16 by Simonyan

and Zisserman (2014) indicates that the feature representations in a lower resolution can be generated by the encoder block. Meanwhile, these latent feature maps have abundant information to represent the input image. As a result of the skill of this architecture, VGG16 has become one of the most frequently applied architectures for image processing using deep neural networks.

A fully convolution neural network (FCN) was proposed by Long et al. (2015) for image segmentation that has the capability to produce the output in the same dimension and resolution of the input image. Instead of adopting a classifier in the image recognition task, Long et al. (2015) designed a decoder network using upsampling operators to reverse the output feature maps of the convolution encoder. The results showed these upsampling operations successfully increased the resolution of the final segmented output image. ^{Sun:}One problem with this architecture is that it has a large number of trainable parameters in the encoder network but a very small amount in the decoder. This results in a large overall size that makes the training process extremely difficult on an end-to-end task, such as extracting noise-free deformation signals from the unwrapped surface displacement maps.

To reduce the large training set requirement and to pursue a more precise segmentation output, a more elegant architecture known as U-net was designed by Ronneberger et al. (2015). U-net has a symmetric topological architecture to preserve the output segmentation dimension and the upsampling operator from ^{Sun:}the FCN was replaced by a combination of the upsampling and a convolution step ^{Sun:}sequence of an upsampling layer and a convolutionla layer, which makes the upscale process trainable. The malleable upscaling strategy can provide more precise localization and resolution of the output. Another important contribution of the U-net architecture is the bridge connection that allows the direct concatenation of the feature representations from the encoder and the upscaled feature maps from the decoder. ^{Sun:}The bridge connection is also known as the skip connection in some of FCNs architectures, for instance, ResNet (He et al., 2016). In a way, the concatenation operations provided by the bridge connection allows the CNN to conserve high resolution feature representations from the encoder even after many pooling layers.

We adopt a similar topological architecture as U-net for the noise-free surface deformation signals retrieval from the time-consecutive unwrapped surface displacement maps. Our architecture is shown in Figure 1. The only difference between our network and U-net in Ronneberger et al. (2015) is the upsampling strategy in the decoder, in which we replace the combination of upsampling and convolution with the transposed convolution (i.e., deconvolution). Compared to U-net, the deconvolution operation allows weighted re-distributions of the feature representations into higher resolution while improving its localization ability. Our network architecture also includes two classic subnets, a contracting path (i.e., a convolutional encoder), and an expansive path (i.e., a convolutional decoder).

As shown in Figure 1, the encoder subnet contains 4 combination blocks where each block has two convolution layers and one max-pooling layer. In the decoder network, 4 symmetric blocks are also designed where each block has a deconvolution layer followed by two convolution layers. A bottleneck connection containing two convolution layers is applied to bridge the encoder and decoder networks. Each convolution layer contains the convolution computation followed by Parametric Rectified-Linear-Unit (PReLU) activation mapping, batch normalization process, and a ^{Sun:}20 percentage dropout ^{Sun:}dropout layer with a constant probability of retention $p = 0.8$ for handling surface displacement maps in different levels of SNR and for improving the generalization capacity of the neural network. The parameterizations of convolution, max-pooling, and deconvolution layers are shown in Table 1, respectively.

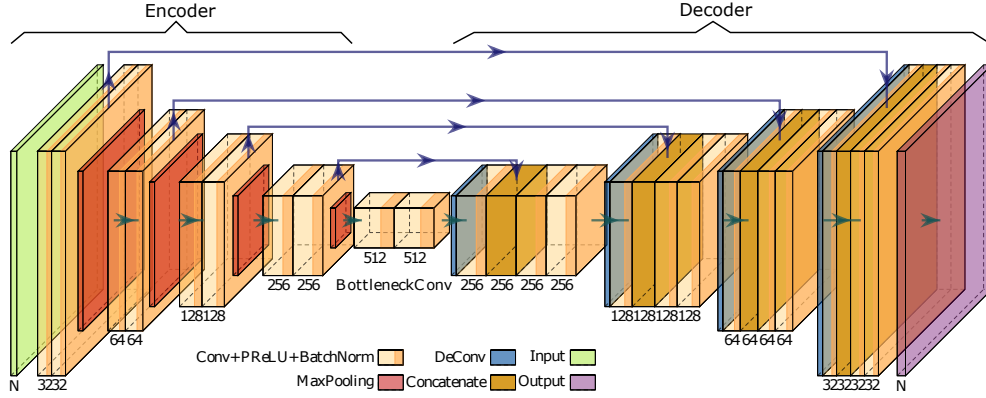


Figure 1. The architecture of the modified U-net for noise-free surface deformation signals separation from the time-consecutive unwrapped surface displacement maps.

Table 1. Summary of parameters used in the architecture shown in Figure 1

Operations	Kernel size	Stride	Zero-padding
Convolution	(3, 3)	(1, 1)	(1, 1)
Max-Pooling	(2, 2)	(2, 2)	(0, 0)
Deconvolution	(3, 3)	(2, 2)	(0, 0)

One of the responsibilities of the encoder network is to reduce the dimensionality of the input. The encoder takes the unwrapped surface displacement maps as the input and outputs feature representations in variant and lower Sun: dimensions spatial dimensions with enriched depth information. The decoder network will increase the dimensionality of the feature representations created by the encoder and convert them back to the original input dimension using the deconvolution operation. In addition, to achieve higher resolution output, the convolution is operated to the concatenation between the result from deconvolution and its corresponding feature maps from related stage of encoder passed by the bridge connection.

3 Network Training and Analysis

In this section, we will first discuss the data preparation, including synthetic dataset simulation and data preprocessing for both unwrapped surface displacement maps (input) and noise-free surface deformation signals (desired output). Secondly, the training and validation stages are discussed and the performance of the trained network using simulated data is examined by implementing it on synthetic test dataset. Finally, we analyze both the temporal and SNR dependencies of the trained network.

3.1 Data Preparation

In order to train our neural network, we first simulated sets of simple synthetic unwrapped surface displacement maps (interferograms) sharing similar properties to the Sentinel-1 InSAR data (Ebmeier, 2016). This synthetic data is designed to include signals of volcanic deformation, spatially correlated atmospheric variability, and errors in orbital estimation. The spatially correlated atmosphere (Hanssen, 2001; Lohman & Simons, 2005) is randomly generated under a normal distribution assumption with cen-

tral values of maximum variance=20mm² and characteristic length-scale exponent=0.5 (Emardson et al., 2003). With an average phase delay gradient of 1cm/km, Shuttle Radar Topography Mission (SRTM) of footprint 2760km², randomly cropped from Telica volcano (1061m, Nicaragua) surrounding area was used to mimic the signals of tropospherically correlated atmospheric variations (Remy et al., 2003; Bekaert et al., 2015). We also simulated the estimation errors of orbits using a linear ramps in form of $ax + by + c$, where $a \sim b \sim 0.01\text{km}^{-1}$. For synthetic deformation pattern construction, we simulate volume changes in spherical sources (Mogi, 1958) at depths ranging randomly from 4km to 15km. The volume change is randomly selected from one of the following six options: a) a linear increase over time, b) a linear decrease over time, c) sinusoidal variations, d) cosinusoidal variations, e) a “pulse” episode of source deflation, or f) inflation. Those six options are employed over a specific period of time such that they are only observed in a few consecutive unwrapped surface displacement maps. *Sun*: Compositions of a time-series simulated interferograms with 180m pixel size and 12 days interval are plotted in Figure 2, where the volume change of the Mogi source following a sinusoidal variation starting from day 40.

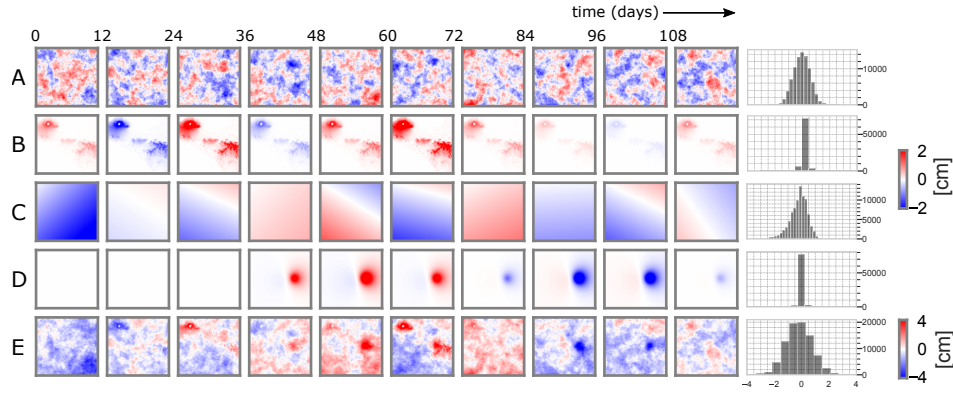


Figure 2. The signals used to simulate a synthetic interferograms following Ebmeier (2016). a) Spatially correlated atmospheric phase changes, b) topographically correlated atmospheric delay, c) linear orbital ramps of the form $\phi = ax + by + c$, where $a \sim b \sim 0.01 \text{ km}^{-1}$, d) Synthetic deformation using a Mogi source at 6km depth, for short-lived sinusoidal varying volume changes starting on day 40. e) Synthetic interferograms, from the sum of all signals above. Histograms show the average distribution of values for all interferograms in the sequence (days 0–120).

The simulated interferometric phase is then obtained by summing all atmospheric signals, orbit contributions, and synthetic Line-Of-Sight (LOS) deformation signals. 20,000 groups of unwrapped surface displacement maps are simulated, and each group contains 20 time-consecutive pairs of unwrapped surface displacement maps and noise-free surface deformation signals *Sun*: with 180m spatial resolution. To be consistent with deep learning terminology, the number of time-consecutive pairs feeding into CNN will subsequently be referred to as the “number of input channels”.

Instead of feeding the neural network with raw simulated unwrapped surface displacement maps, data preprocessing is generally required before the training stage of *Sun*: a neural network because of data inconsistency, incompleteness, and unpredictable range. Previous studies have demonstrated several advantages of using data preprocessing techniques (Kotsiantis et al., 2006), such as a simpler relationship between input and target, improvement of the stability, and enhancement of the generalization capacity of the neural network. In our experiments, a simple preprocessing is required for transform-

ing both time-consecutive unwrapped surface displacement maps and noise-free surface deformation signals into a comparable range.

First, outlier pixels, which are defined as those with a Z-score with an absolute value greater than 3, are replaced by a neighborhood value from unwrapped surface displacement map. Time-consecutive unwrapped surface displacement maps are then normalized by their maximum and minimum values, i.e., $\Phi_{in} = (\Phi - \Phi_{min}) / (\Phi_{max} - \Phi_{min})$. Thus, all the input value fed into the neural network are in the range of $[0, 1]$. The noise-free surface deformations are normalized using the pair of minimum and maximum values from their corresponding time-consecutive unwrapped surface displacement maps without subtracting their minimum, i.e., $\Phi_{out} = \Phi_{def} / (\Phi_{max} - \Phi_{min})$. As a result, smaller deformation signals will have a smaller influence on the objective function when back-propagating the neural network for weight parameter updating.

3.2 Training and SNR Dependency Analysis

All preprocessed data are randomly divided into training (80 percent of the total number of time-consecutive unwrapped surface displacement maps a total of 16,000) and test (the remaining 20 percent a total of 4,000) datasets for training and evaluation.

To increase the quantities and improve the diversity of training and test datasets, data augmentation, through random rotation, horizontal, and/or vertical flipping, was performed on both input unwrapped surface displacement maps and coherent noise-free surface deformation signals on training dataset during training stage. The designed encoder-decoder neural network (Figure 1) was trained using minibatch SGD optimization algorithm with $l1$ -norm discrepancy evaluation (equation 2) on the PyTorch platform (Paszke et al., 2019). To determine an appropriate time series length for the input unwrapped surface displacement maps, we trained the neural network eight times separately with different number of input channels. *Thus, we have eight trained network models accepting [6, 8, 10, 12, 14, 16, 18, 20] number of input channels, respectively.* The average training time over all the neural networks is approximately 150 minutes for 30 epochs with 20 NVIDIA Tesla P100 GPU accelerators.

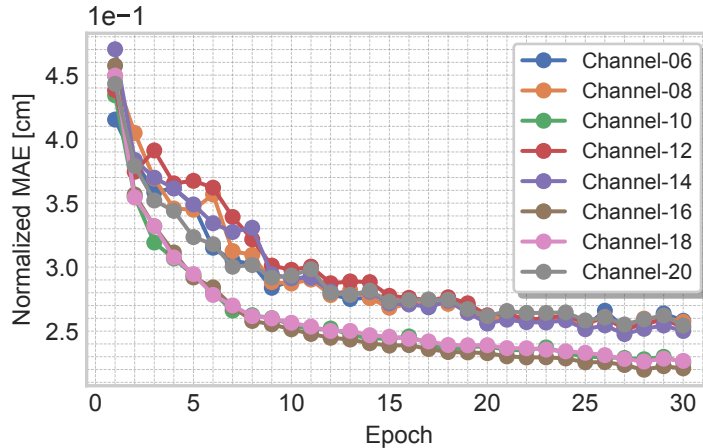


Figure 3. The normalized mean-absolute-errors (MAE) of predictions using networks with variant numbers of input channels during the training stage.

To visualize difference in the performance between the eight trained neural network models, we applied each of them to the entire test dataset after each training epoch and

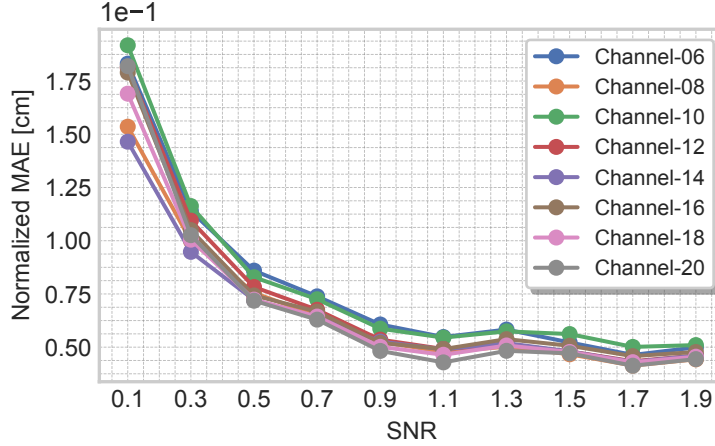


Figure 4. The normalized MAE of predictions, for InSAR data with different signal-to-noise ratios (SNRs), using networks with variant numbers of input channels after 30 epochs training.

computed their normalized mean-absolute-errors (MAE) (Figure 3). As the training progresses, the normalized MAEs of all eight networks are reduced with only small oscillations. That is to say, it is conceivable to build an end-to-end architecture of neural network for retrieving the interested (i.e., noise-free) surface deformation signals directly from time-consecutive unwrapped surface displacement maps.

In experiments of separating atmospheric noise from deformation signals, the SNR is a key limiting factor for the traditional methods. Therefore, the analysis of SNR dependency of the trained neural network is requisite to evaluate the generalization capacity and the stability of the trained network. We split the test dataset into 10 subsets with SNR ranging from $S_{un}:[0.1, 2.0]$ every 0.20.1 to 2.0 with interval of 0.2. Then the predicted $S_{un}:\text{normalized}$ MAEs of each subset are computed using each of the eight trained networks. The results are plotted in Figure 4. For each trained network, the predicted $S_{un}:\text{normalized}$ MAE decreases as the SNR of the input time-consecutive unwrapped surface displacement maps increases. However, it is worth noting that all predicted $S_{un}:\text{normalized}$ MAE for SNR in range of $[0.1, 2.0]$ are reasonably small.

3.3 Implementation and Temporal Dependency Analysis

To further intuitively examine the performance of these eight network models, in Figure 5, we plotted the predicted surface deformation signals from 20 time-consecutive unwrapped surface displacement maps (from left to right) associating with its noise-free deformation signals. $S_{un}:\text{Row A}$ The first row of Figure 5 shows 20 time-consecutive unwrapped surface displacement maps at 12 day intervals with a SNR of 0.95 as inputs. Their coherent noise-free surface deformation feature maps are arranged in the second row in the same order. From this comparison, it is clear that partial deformation signals can be observed from some unwrapped surface displacement maps, while others are distorted by atmospheric noise.

$S_{un}:$ Surface deformations were predicted using the eight trained networks by iteratively feeding 20 preprocessed time-consecutive unwrapped surface displacement maps until finished if the time length N of input required by the network is shorter than 20. For instance, using a network with the number of input channels of 8, 20 time-consecutive unwrapped surface displacement maps are required to be split into three temporal segments, i.e., 1-8, 9-16, and 13-20, and then are fed into the network,

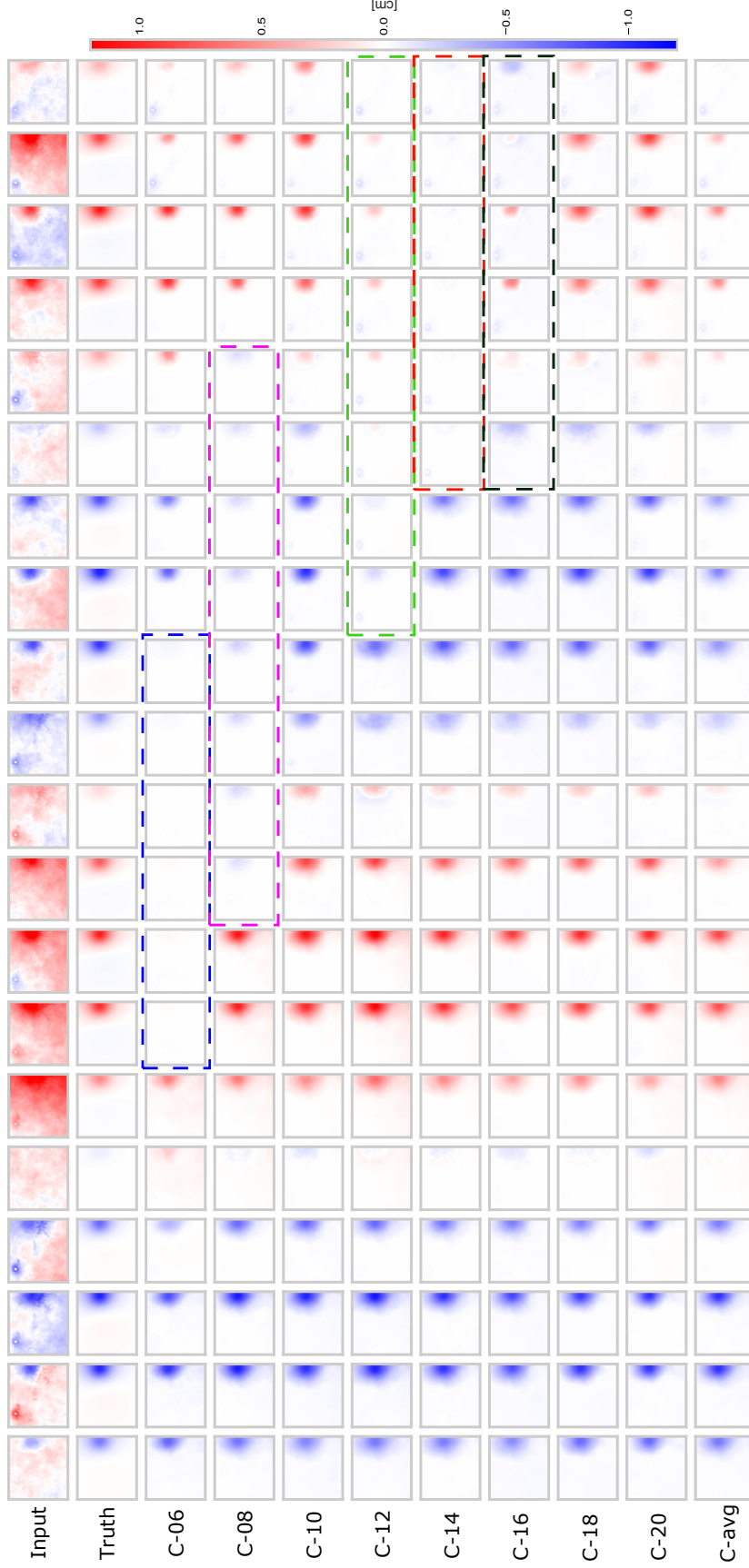


Figure 5. Surface deformation signals extraction using proposed networks trained with eight different numbers of input channels. The first row shows 20 time-consecutive unwrapped surface displacement maps with 12 days interval (i.e., inputs), the second row delineates the corresponding noise-free surface deformations (i.e., ground-truth targets). The rows C-XX (here, XX refers to the number of input channels and $XX \in [6, 8, 10, 12, 14, 16, 18, 20]$) represent extracted surface displacements using networks with variant input channels, respectively. The row C-avg is the average prediction of rows C-XX. The dashed boxes highlight the imprecise predictions.



Figure 6. The residuals of the ground-truth surface deformation signals and predictions using proposed networks with different numbers of input channels, i.e., the subtraction between the second row of Figure 5 and rows below it.

respectively. Note that the temporal segment from 13 to 16 are repetitively predicted, and therefore the average prediction is adopted as the final results. The results are illustrated in rows C-XX (XX refers to the number of input channels and $XX \in [6, 8, 10, 12, 14, 16, 18, 20]$) of Figure 5. ^{Sun:}In row C-06, the network is capable of retrieving the surface deformations at both the initial and the last stage, but fails to detect them during the middle phase (indicated as the blue dashed-rectangle). Similar behavior is seen in row C-08 (shown as the pink dashed-rectangle) when extraction is performed using network with input channel of 8. This demonstrates that the stability of the neural network for detecting the pattern of phase changes of deformation signals, is limited when the number of channels fed into the neural network is shorter than the time scale of the deformation.

One possible solution is to apply a multi-coverage strategy during the prediction stage. In other words, the time-step moving forward in prediction must be much shorter than the time length required by the neural network, which ensures most of the unwrapped surface displacement maps are fed into the network multiple times. The drawback of this approach is that it smooths the time-series plot of deformation signals, and it does not work when the network feeding time period is much shorter than the temporal period of deformation phase changes.

Another ^{Sun:}formula solution is to expand the time length covered by the network. Rows C-10/18/20 of Figure 5 shows the surface deformation predictions using networks with input channels of [10, 18, 20], respectively. The reconstructed signals capture all detailed phase changes of surface displacements. However, when increasing the number of input channels for the neural network there is a trade-off because it not only enhances the quality of interested signals, but introduces undesired noise which is destructive especially when $SNR < 1$. For example, with the input channels of [12, 14, 16], the detected surface deformations were also incomplete (shown as green, red, and black dashed-rectangles) because of the high impact of stratified atmospheric signals at bottom-left corner of the area. To get an intuitive observation, the residuals between the ground-truth surface deformation and all predictions using networks with different number of input channels are calculated and plotted in Figure 6.

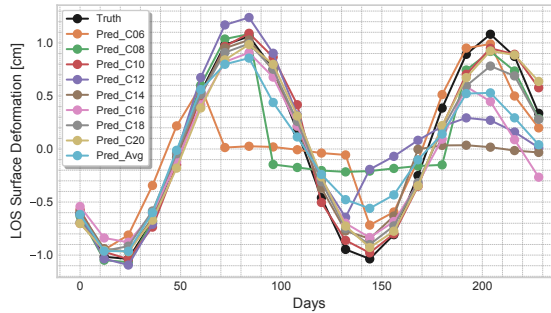


Figure 7. Time-series plots of surface deformations above the source location extracted from the ground-truth and all predictions in Figure 5.

We extracted the surface deformation above the source location from the second row in Figure 5 and plotted the time-series phase changes in black in Figure 7, which is acting in place of the noise-free continuous data collection from ground-based GPS station. The detected deformation signals using the neural networks with different input channels are also extracted at the same location from rows C-XX and C-avg in Figure 5, and plotted in Figure 7 to compare their abilities to extract the signal. As discussed, net-

works with input channels of [6, 8] fail to extract signals of interest during the middle of the time period, and networks with input channels of [12, 14, 16] fail to extract the signal at end of the time period. Predictions using networks with 10 and 20 input channels provide a good fit to the noise-free time-series plot. This suggests that the matching score of time-series plots between the prediction and reachable noise-free continuous GPS data may be a great indicator of finding the optimal number of input channels for network model on related InSAR data.

Without noise-free continuous GPS data for validation, to improve the stability of network, we recommend implementing the prediction with several networks using a variant numbers of input channels and then computing the average of them as the final predicted result. In this case we calculated the average prediction over networks with poor prediction skill (channels [6, 8, 12, 14, 16] or rows [C-06, C-08, C-12, C-14, C-16] in Figure 5) are plotted in the last row (row C-avg) Figure 5. Row C-avg shows that the averaged prediction signal still lack precision but successfully captured all primary phase changes of the surface displacements. In Figure 7, the extracted time-series plot of averaged predictions is not perfectly recovered, but well enough to be an indicator for detecting major volcanic unrest. As a result, to improve the stability of network in the absence of noise-free continuous GPS validation data, we recommend implementing the prediction with several networks using variant numbers of input channels and then computing the average of them as the final predicted result.

Figure 8 shows the cross-plots of noise-free surface deformations (the second row of Figure 5) compared to predictions using 10 input channels (row C-10 of Figure 5) and the average of predicted signal (row C-avg of Figure 5). The optimal relationship (i.e., $\Phi_{def} = \hat{\Phi}_{def}$) of predicted and ground-truth signals is plotted in the black dashed line. The distribution of predictions using 10 input channels are plotted as bright blue dots in Figure 5. The best-fit line, shown in dark blue, is very close to the optimal fitting line indicating that the 10 input channel prediction has high skill. The cross-mapping of average predictions and the ground-truth signals are plotted as orange dots with its best-fit line drawn in red. As expected, due to the smoothing influence, the absolute values of average predictions are smaller than the noise-free surface displacements.

To further investigate the influence of the number of input channels for the CNN, we compute the normalized MAE of surface deformation detection using the networks with different numbers of input channels on InSAR data with various SNR. In Figure 9, the normalized MAE with respect to the number of input channels on test subsets with SNR: [0.1, 0.3, 0.5, 0.7, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1] are plotted. For a very small SNR of 0.1 the mean absolute error remains high regardless of the number of input channels, consistent with our discussion above. As the SNR increases from 0.3 to 1.3, the MAE of the prediction reduces as the number of input channels increases. However, the number of input channels has limited impact on deformation detection when SNR ≥ 1.5 as shown in Figure 9. This is reasonable because, for high SNR, the neural networks are able to extract enough information within a short time period of unwrapped surface displacement maps.

4 Case Study: Masaya Volcano Unrest

4.1 Background on Masaya and Period of Study

Surface deformation maps of volcanic regions within the Central American Volcanic Arc are often plagued by atmospheric phase delays from the troposphere, and this distorts measurements of noise-free surface deformation associated with volcanic activity (Ebmeier et al., 2013; K. J. Stephens et al., 2020). For this study, we focus our efforts on Masaya caldera, a basaltic caldera located approximately 20 km SE of Managua City in Nicaragua. The summit is approximately 600 m in elevation, and consists of several

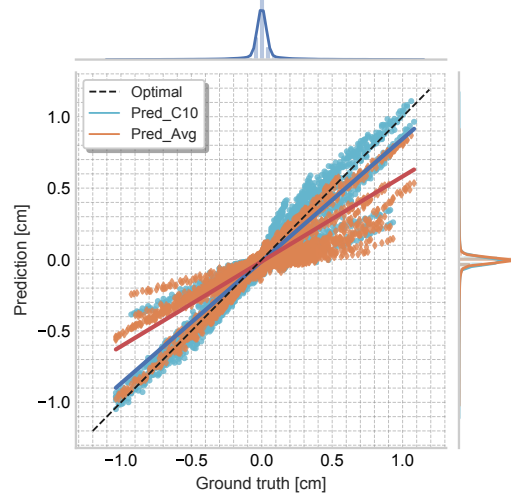


Figure 8. Cross-plots of noise-free surface deformation signals compared to the prediction using networks with 10 input channels and the averaged prediction, respectively. The dots are pixels of predictions C-10 of Figure 5 (blue dots) and pixels of the average prediction C-avg of Figure 5 (orange dots), respectively, with respects to the ground-truth signals (the second row of Figure 5). Their best-fit lines are plotted in dark-blue and red respectively.

Sum: ~~nesternested~~ pit craters, of which the Santiago pit crater has been the most active since its formation in 1858-1859 (Global Volcanism Program, 2013; McBirney, 1956; Rymer et al., 1998). Masaya has exhibited a variety of explosive and effusive activities, including phreatomagmatic to Plinian episodes, lava flows, and lava lakes hosted within the summit pit craters (Bice, 1980; Global Volcanism Program, 2013; Kutterolf et al., 2007; McBirney, 1956). Since December 2015, a lava lake has been hosted within the Santiago pit crater at the summit (Global Volcanism Program, 2013). Various geodetic observations spanning the appearance of the lava lake captured 8cm and 5.5 cm of vertical uplift using InSAR (K. Stephens & Wauthier, 2018) and precision levelling (Murray, n.d.; Rymer et al., n.d.) observations respectively. Additionally, since November 2015, one GPS station within the caldera (MAVC) has been operating continuously.

4.2 Data Processing and Time-series Analysis

We obtained 60 ascending SAR images from the Italian Space Agency COSMO-SkyMed (CSK) constellation to create surface deformation maps, spanning from August 2015 to October 2016. Potential interferogram pairs were evaluated using the criteria that the perpendicular baseline between potential pairs is less than 300 m, and the temporal spacing is less than 180 days. Following these criteria, 281 interferograms were processed using the GAMMA software (Werner et al., 2000), with topographic phase contributions removed using a 12 m spatial resolution TanDEM-X digital elevation model (Rizzoli et al., 2017; Wessel et al., 2018). An adaptive spectral filter was used to smooth interferograms (Goldstein & Werner, 1998) and unwrapping was performed using the minimum cost flow technique and triangular irregular network (Costantini, 1998). Unwrapped interferograms were converted from LOS phase radians to centimeters with respect to the ground, and re-referenced to a region ~~Sum: in~~ within the town of Masaya that was assumed to have minimal deformation. The interferograms were then downsampled to 90 m pixel spacing to reduce computational time. No coherence masks were applied to the interferograms in order for the full atmospheric noise field to be examined.

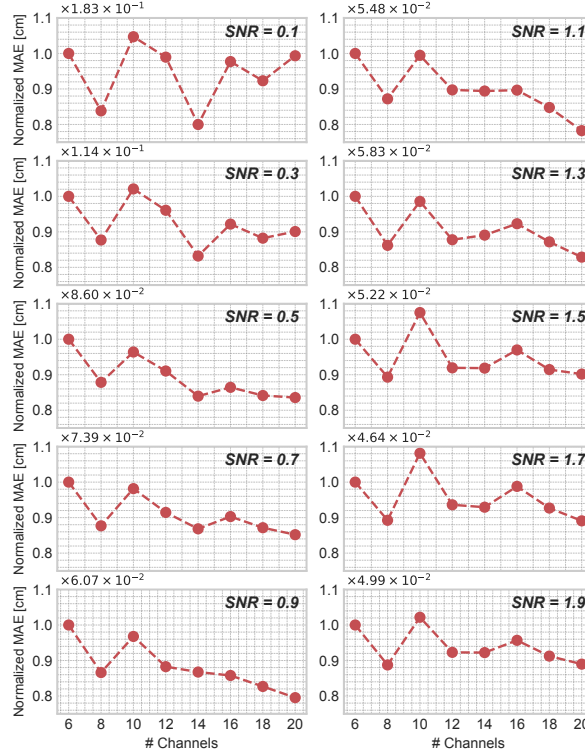


Figure 9. The normalized MAE of predictions with respect to the number of input channels for datasets with different SNRs.

The SBAS time-series technique creates LOS surface displacement maps between each acquisition date using a pixel-by-pixel linear least squares inversion method (Berardino et al., 2002; Lundgren et al., 2001). The first acquisition date (August 10th 2015) was typically set to be zero as the reference, and thus all surface displacement maps are with respect to this date, shown in Figure 10. All surface displacement maps were further downsampled to 180 m pixel spacing to match the input size requirement of the CNN. These time-consecutive unwrapped surface displacement maps generated by the time-series analysis are then used as the input to our CNN.

4.3 Implementations

Our pre-trained CNN requires time-consecutive unwrapped surface displacement maps with a constant date interval, which was 12 days during out training procedure. However, the ascending images obtained for the Masaya volcano have different date intervals. The date interval is calculated using dates between the current image and subsequent image, as shown in Figure 11. The majority of the time intervals are 4 or 8 days, which we will assume are consistent enough for our purposes. However, there is a large temporal gap in the measurements of 48 days from late January 2016 to March 2016. This gap divides the 60 images into two segments of roughly equal length. Therefore, instead of treating all 60 images as one piece of time-consecutive deformation maps, we implement the deformation detection using the neural network for each half of the Masaya dataset. The first 26 images spanning from August 2015 to January 2016 in Figure 10 are considered the first sub-dataset, and the remaining 34 spanning from March 2016 to November 2016 are considered the second sub-dataset. The varying and inconsistent time intervals of Masaya time series, as well as different deformation

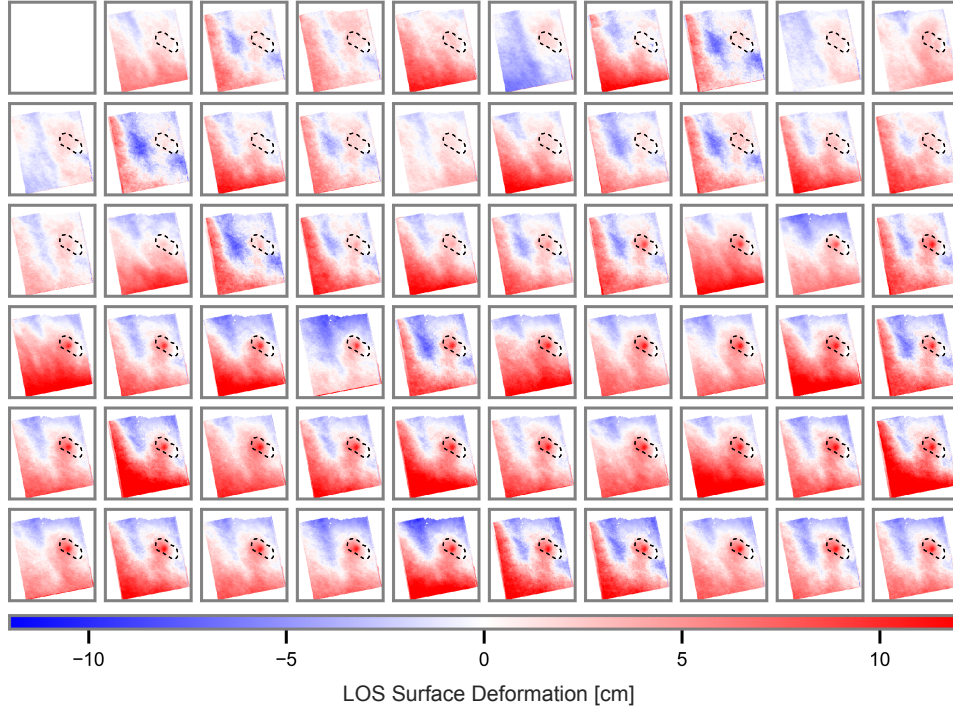


Figure 10. 60 consecutive CSK surface displacement maps obtained from SBAS time-series analysis, spanning from 10 August to 15 October 2016.

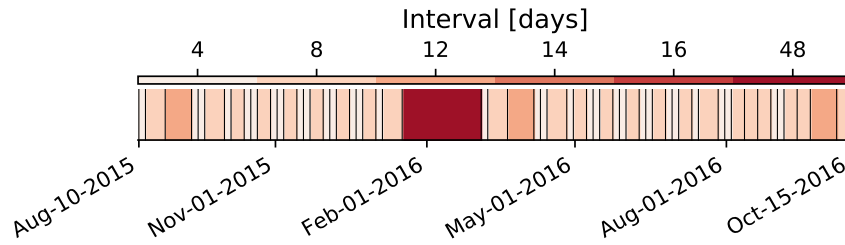


Figure 11. Time intervals among all 60 consecutive CSK acquisition dates, where the time-consecutive unwrapped deformation maps are shown in Figure 10.

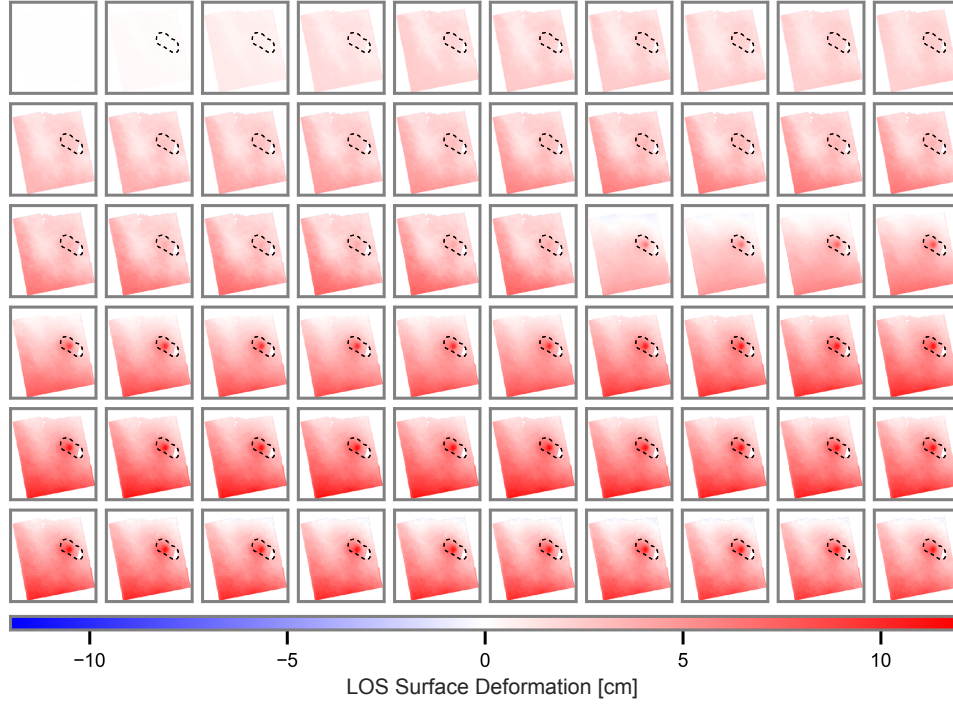


Figure 12. The average prediction of LOS surface deformation at Masaya, where the Masaya caldera is indicated in the dashed-line.

rates, may lead to varying suitable number of input channels for achieving the best prediction. Therefore, to stabilize the prediction of our network, as discussed in previous section, an ensemble-average strategy can be adopted to compensate their effects to some extent. A comprehensive analysis of the relationship among the network's performance, varying deformation rates, and inconsistent time intervals is our ongoing research.

In order to evaluate the performance of our approach and compare it to the processed continuous GPS data, we implement the average-ensemble prediction strategy discussed previously. Each sub-dataset is fed into a set of 8 networks with varying number of input channels of [6, 8, 10, 12, 14, 16, 18, 20]. The average detection result is calculated a posterior. The final surface displacements detection results for all images are plotted in Figure 12. ^{Sun:}Comparing to the original SBAS time-series (Figure 10), the surface deformations at Masaya obtained with the CNN (Figure 12) are able to more clearly identify volcanic surface deformations as .

4.4 Time-Series Analysis

The GPS station MAVC is located within the Masaya caldera at the Masaya Visitor's Center (location is shown in Figure 13) and is used as a validation tool in this study. In order to compare the GPS time-series data to that of the InSAR LOS results, the three vertical and horizontal components of the continuous GPS data were converted into the satellite LOS using the satellite incidence angle (θ) and azimuth of the satellite heading vector(α), as well as the LOS vector components of the surface displacement field d_N (north-south horizontal), d_E (east-west horizontal), and d_U (vertical) (Fialko et al., 2001):

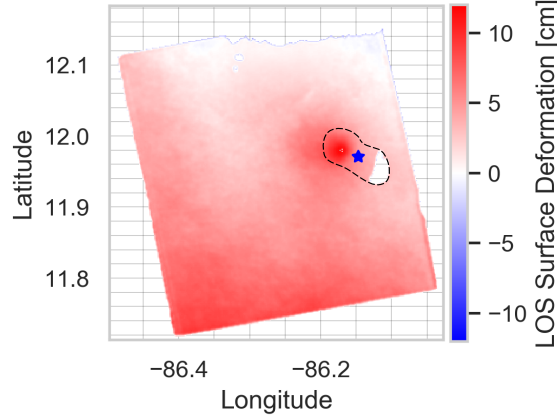


Figure 13. The average prediction of LOS surface deformation on 15 October 2016, extracted from the last time segment in Figure 12. The blue star represents the location of MAVC GPS station within Masaya caldera indicated in the dashed-line.

$$d_{LOS} = (d_N \sin \alpha - d_E \cos \alpha) \sin \theta + d_U \cos \theta \quad (3)$$

With the GPS, CSK InSAR time-series and CNN prediction all in the same units, time-series of surface displacements at the GPS location can be extracted from average-ensemble detected maps and the InSAR time-series for direct comparison. Figure 14 shows the LOS displacements in the GPS (blue dots), CSK InSAR time-series (green triangles), and CNN prediction (red diamonds).

The GPS data shows changes in LOS displacement within the region of the offset magma reservoir, which is associated with the appearance of the lava lake at the summit. Over the first 6 months of lava lake activity, the offset magma reservoir continued to inflate, however the magnitude of LOS displacement is less than that observed at the centre of the deforming region (K. J. Stephens et al., 2020). The InSAR time-series closely follows the GPS LOS displacement within the first two months of lava lake activity whereas the CNN over-estimates the LOS displacement. These initial errors in the CNN may be caused by poor spatial sampling. Beginning around April/May 2016, the InSAR time-series under-estimates the LOS displacement. This may be due to the large temporal gap in the InSAR data from late January to March 2016 or the poor temporal sampling of the InSAR data compared to the daily GPS solutions. However, The CNN LOS displacement becomes more consistent with the GPS station starting from March 2016, demonstrating that our proposed approach has the capacity of revealing the Masaya volcano unrest.

5 Conclusions

Abundant routine acquisitions of SAR using satellites make global near real-time volcano monitoring a reality. However, two major questions need to be answered before its practical application: 1) how to build a straightforward tool for mitigating atmospheric noise and detecting surface displacements signals, 2) is this tool fast enough to accommodate near real-time volcano monitoring on a global basis. In this paper, we discussed the possibility of directly detecting volcanic surface deformation without atmospheric noise from time-consecutive unwrapped surface displacement maps (interferograms) using neural network. We demonstrated the use of an end-to-end CNN with an encoder-

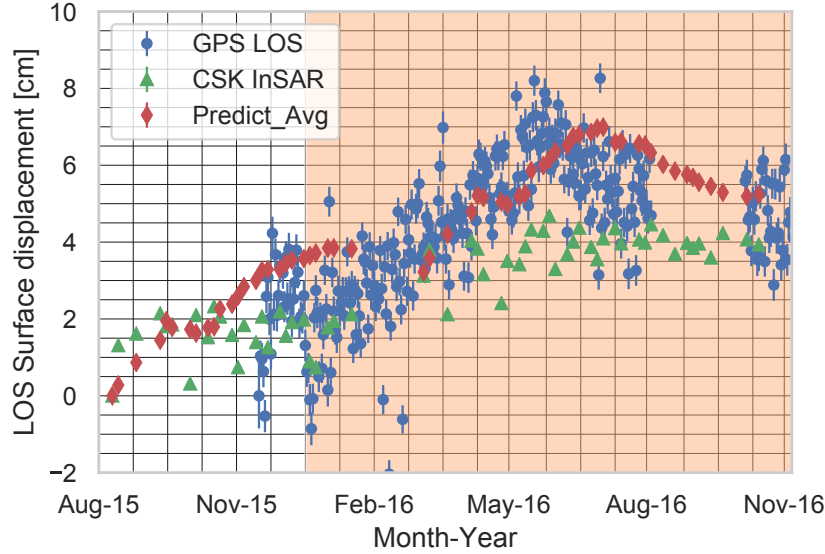


Figure 14. Time-series comparison between the LOS displacement observed from MAVC GPS station (blue), CSK InSAR time-series analysis (green), and the average-detected deformation (red) using CNN extracted from Figure 12. Orange background indicates temporally when the lava lake is present at the summit.

decoder architecture as a potential tool for practical near real-time volcano monitoring on a global scale. We trained the proposed CNN with simulated unwrapped surface displacement maps only, but it may be further optimized with more realistic InSAR datasets.

Analysis using synthetic benchmark shows that the CNN is capable of revealing noise-free surface deformation signals from unwrapped surface displacement maps with variant SNRs. However, the precision of detected results depends on the time scale that time-consecutive unwrapped surface displacement maps covered. Depending on data availability we recommended one of two options to determine the best number of input channels for the CNN: 1) ground-based continuous GPS data can be used as an indicator if available, or 2) an ensemble of CNN detections with various input channels can be computed where the the final detected result is the average over all members.

A case study is conducted where the CNN ensemble average approach is applied to detect surface displacements at the Masaya volcano. The results demonstrate much more distinct surface deformation patterns than in the raw images. Time-series from the CNN detection extracted at the same location as available GPS observation are used to validate the CNN detection, showing that the CNN is able to detect Masaya volcano unrest. The performance of our proposed CNN on both synthetic and real InSAR datasets shows great potentials of automatic noise-free surface deformation detection for ~~Sun: global~~ globally near real-time volcano monitoring.

Acknowledgments

All SAR COSMO-SkyMed data obtained over Masaya and Telica were provided through the Committee on Earth Observation Satellites (CEOS) Volcano Pilot and Demonstrator working group programs. TanDEM-X 12 m resolution data were obtained through the German Aerospace Center (DLR Proposal ID 1552). This study was funded by National Aeronautics and Space Administration (NASA) Earth Surface and Interior (ESI)

grants (NNX17AD70G) to P.C.L. and C.W., (ESI NNX16AK87G and ESI 80NSSC20K0490) to C.W., and NSF (EAR-0911546) to P.C.L. A

References

- Anantrasirichai, N., Biggs, J., Albino, F., & Bull, D. (2019). The application of convolutional neural networks to detect slow, sustained deformation in insar timeseries. *Geophysical Research Letters*, *46*, 11850–11858.
- Anantrasirichai, N., Biggs, J., Albino, F., Hill, P., & Bull, D. (2018). Application of machine learning to classification of volcanic deformation in routinely generated insar data. *Journal of Geophysical Research: Solid Earth*, *123*(8), 6592–6606.
- Bekaert, D., Hooper, A., & Wright, T. (2015). A spatially variable power law tropospheric correction technique for insar data. *Journal of Geophysical Research: Solid Earth*, *120*(2), 1345–1356.
- Berardino, P., Fornaro, G., Lanari, R., & Sansosti, E. (2002). A new algorithm for surface deformation monitoring based on small baseline differential sar interferograms. *IEEE Transactions on geoscience and remote sensing*, *40*(11), 2375–2383.
- Bice, D. (1980). Tephra stratigraphy and physical aspects of recent volcanism near managua. *Nicaragua {Ph. D. Thesis}: Berkeley, California, University of California*.
- Boergens, E., Rangelova, E., Sideris, M. G., & Kusche, J. (2014). Assessment of the capabilities of the temporal and spatiotemporal ica method for geophysical signal separation in grace data. *Journal of Geophysical Research: Solid Earth*, *119*(5), 4429–4447.
- Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of compstat'2010* (pp. 177–186). Springer.
- Calhoun, V. D., Adali, T., Pearlson, G., & Pekar, J. J. (2001). Spatial and temporal independent component analysis of functional mri data containing a pair of task-related waveforms. *Human brain mapping*, *13*(1), 43–53.
- Costantini, M. (1998). A novel phase unwrapping method based on network programming. *IEEE Transactions on geoscience and remote sensing*, *36*(3), 813–821.
- Doin, M.-P., Lasserre, C., Peltzer, G., Cavalié, O., & Doubre, C. (2009). Corrections of stratified tropospheric delays in sar interferometry: Validation with global atmospheric models. *Journal of Applied Geophysics*, *69*(1), 35–50.
- Ebmeier, S. (2016). Application of independent component analysis to multitemporal insar data with volcanic case studies. *Journal of Geophysical Research: Solid Earth*, *121*(12), 8970–8986.
- Ebmeier, S., Biggs, J., Mather, T., & Amelung, F. (2013). Applicability of insar to tropical volcanoes: insights from central america. *Geological Society, London, Special Publications*, *380*(1), 15–37.
- Emardson, T., Simons, M., & Webb, F. (2003). Neutral atmospheric delay in interferometric synthetic aperture radar applications: Statistical description and mitigation. *Journal of Geophysical Research: Solid Earth*, *108*(B5).
- Ferretti, A., Prati, C., & Rocca, F. (2001). Permanent scatterers in sar interferometry. *IEEE Transactions on geoscience and remote sensing*, *39*(1), 8–20.
- Fialko, Y., Simons, M., & Agnew, D. (2001). The complete (3-d) surface displacement field in the epicentral area of the 1999 mw7. 1 hector mine earthquake, california, from space geodetic observations. *Geophysical research letters*, *28*(16), 3063–3066.
- Gaddes, M., Hooper, A., & Bagnardi, M. (2019). Using machine learning to automatically detect volcanic unrest in a time series of interferograms. *Journal of Geophysical Research: Solid Earth*, *124*.

- Global Volcanism Program. (2013). *Masaya (344100) In Volcanoes of the World; venzke, e., ed.; smithsonian institution.* <https://doi.org/10.5479/si.GVP.VOTW4-2013>.
- Goldstein, R. M., & Werner, C. L. (1998). Radar interferogram filtering for geophysical applications. *Geophysical research letters*, 25(21), 4035–4038.
- Häggström, I., Schmidlein, C. R., Campanella, G., & Fuchs, T. J. (2019). Deeppet: A deep encoder–decoder network for directly solving the pet image reconstruction inverse problem. *Medical image analysis*, 54, 253–262.
- Hanssen, R. F. (2001). *Radar interferometry: data interpretation and error analysis* (Vol. 2). Springer Science & Business Media.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).
- Hooper, A. (2008). A multi-temporal insar method incorporating both persistent scatterer and small baseline approaches. *Geophysical Research Letters*, 35(16).
- Kampes, B. (2005). *Displacement parameter estimation using permanent scatterer interferometry* (Unpublished doctoral dissertation). Delft University of Technology.
- Kim, J., Kwon Lee, J., & Mu Lee, K. (2016). Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1646–1654).
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Data preprocessing for supervised learning. *International Journal of Computer Science*, 1(2), 111–117.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).
- Kutterolf, S., Freundt, A., Perez, W., Wehrmann, H., & Schmincke, H.-U. (2007). Late pleistocene to holocene temporal succession and magnitudes of highly-explosive volcanic eruptions in west-central nicaragua. *Journal of Volcanology and Geothermal Research*, 163(1-4), 55–82.
- Lanari, R., Casu, F., Manzo, M., Zeni, G., Berardino, P., Manunta, M., & Pepe, A. (2007). An overview of the small baseline subset algorithm: A dinsar technique for surface deformation analysis. In *Deformation and gravity change: Indicators of isostasy, tectonics, volcanism, and climate change* (pp. 637–661). Springer.
- Li, M., Zhang, T., Chen, Y., & Smola, A. J. (2014). Efficient mini-batch training for stochastic optimization. In *Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 661–670).
- Lohman, R. B., & Simons, M. (2005). Some thoughts on the use of insar data to constrain models of surface deformation: Noise structure and data downsampling. *Geochemistry, Geophysics, Geosystems*, 6(1).
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431–3440).
- Loughlin, S. C., Sparks, R. S. J., Sparks, S., Brown, S. K., Jenkins, S. F., & Vye-Brown, C. (2015). *Global volcanic hazards and risk*. Cambridge University Press.
- Lundgren, P., Usai, S., Sansosti, E., Lanari, R., Tesauro, M., Fornaro, G., & Berardino, P. (2001). Modeling surface deformation observed with synthetic aperture radar interferometry at campi flegrei caldera. *Journal of Geophysical Research: Solid Earth*, 106(B9), 19355–19366.
- Massonnet, D., & Feigl, K. L. (1998). Radar interferometry and its application to changes in the earth’s surface. *Reviews of geophysics*, 36(4), 441–500.

- McBirney, A. R. (1956). The nicaraguan volcano masaya and its caldera. *Eos, Transactions American Geophysical Union*, 37(1), 83–96.
- Mogi, K. (1958). Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. *Bull. Earthq Res Inst*, 36, 99–134.
- Moody, J., & Darken, C. (1988). *Learning with localized receptive fields*. Yale Univ., Department of Computer Science.
- Murray, A. (n.d.). Recent inflation at masay volcano, nicaragua. In *IAVCEI 2017 Scientific Assembly: Fostering Integrative Studies of Volcanoes; IAVCEI: Portland, Oregon*, 739.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in neural information processing systems 32* (pp. 8024–8035). Curran Associates, Inc. Retrieved from <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>
- Remy, D., Bonvalot, S., Briole, P., & Murakami, M. (2003). Accurate measurements of tropospheric effects in volcanic areas from sar interferometry data: Application to sakurajima volcano (japan). *Earth and Planetary Science Letters*, 213(3-4), 299–310.
- Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Tridon, D. B., Bräutigam, B., ... others (2017). Generation and performance assessment of the global tandem-x digital elevation model. *ISPRS Journal of Photogrammetry and Remote Sensing*, 132, 119–139.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International conference on medical image computing and computer-assisted intervention* (pp. 234–241).
- Rymer, H., de Vries, B. v. W., Stix, J., & Williams-Jones, G. (1998). Pit crater structure and processes governing persistent activity at masaya volcano, nicaragua. *Bulletin of Volcanology*, 59(5), 345–355.
- Rymer, H., Williams-Jones, G., Murray, J., Delmelle, P., Reid, K., & Caravantes Gonzalez, G. (n.d.). Precursors to the current activity at masaya volcano, nicaragua. In *IAVCEI 2017 Scientific Assembly: Fostering Integrative Studies of Volcanoes; IAVCEI: Portland, Oregon*, 939.
- Shi, Y., Zeng, W., Wang, N., & Zhao, L. (2018). A new constrained spatiotemporal ica method based on multi-objective optimization for fmri data analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26(9), 1690–1699.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Stephens, K., & Wauthier, C. (2018). Satellite geodesy captures offset magma supply associated with lava lake appearance at masaya volcano, nicaragua. *Geophysical Research Letters*, 45(6), 2669–2678.
- Stephens, K. J., Wauthier, C., Bussard, R. C., Higgins, M., & LaFemina, P. C. (2020). Assessment of mitigation strategies for tropospheric phase contributions to insar time-series datasets over two nicaraguan volcanoes. *Remote Sensing*, 12(5), 782.
- Stone, J., Porrill, J., Porter, N., & Wilkinson, I. (2002). Spatiotemporal independent component analysis of event-related fmri data using skewed probability density functions. *NeuroImage*, 15(2), 407–421.
- Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. (2019). Bigearthnet: A large-scale benchmark archive for remote sensing image understanding. In *Igarss 2019-2019 ieee international geoscience and remote sensing symposium* (pp. 5901–5904).

- 756 Werner, C., Wegmüller, U., Strozzi, T., & Wiesmann, A. (2000). Gamma sar and in-
757 terferometric processing software. In *Proceedings of the ers-envisat symposium,*
758 *gothenburg, sweden* (Vol. 1620, p. 1620).
- 759 Wessel, B., Huber, M., Wohlfart, C., Marschall, U., Kosmann, D., & Roth, A.
760 (2018). Accuracy assessment of the global tandem-x digital elevation model
761 with gps data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 139,
762 171–182.
- 763 Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). Beyond a gaussian de-
764 noiser: Residual learning of deep cnn for image denoising. *IEEE Transactions*
765 *on Image Processing*, 26(7), 3142–3155.