Application of Machine Learning to Classification of Volcanic Deformation in Routinely-Generated InSAR data

N. Anantrasirichai[1], J. Biggs[2], F. Albino[2], P. Hill[1], D. Bull[1]

[1]Visual Information Laboratory, University of Bristol, UK
[2]School of Earth Sciences, University of Bristol, UK

Key Points:

- We present a machine learning framework to detect volcanic ground deformation in wrapped interferograms using convolutional neural networks.
- The classification model is initialised with Envisat dataset, then tested and retrained with Sentinel-1 dataset covering over 900 volcanoes.
- This framework can reduce the number of interferograms for manual inspection from more than 30,000 to approximately 100.

Corresponding author: N. Anantrasirichai, n.anantrasirichai@bristol.ac.uk

Abstract

Recent improvements in the frequency, type and availability of satellite images mean it is now feasible to routinely study volcanoes in remote and inaccessible regions, including those with no ground-based monitoring. In particular, Interferometric Synthetic Aperture Radar (InSAR) data can detect surface deformation, which has a strong statistical link to eruption. However, the dataset produced by the recently-launched Sentinel-1 satellite is too large to be manually analysed on a global basis. In this study, we systematically process >30,000 short-term interferograms at over 900 volcanoes and apply machine learning algorithms to automatically detect volcanic ground deformation. We use a convolutional neutral network (CNN) to classify interferometric fringes in wrapped interferograms with no atmospheric corrections. We employ a transfer learning strategy, and test a range of pretrained networks, finding that AlexNet is best suited to this task. The positive results are checked by an expert and fed back for model updating. Following training with a combination of both positive and negative examples, this method reduced the number of interferograms to ~100 which required further inspection, of which at least 39 are considered 'true positives'. We demonstrate that machine learning can efficiently detect large, rapid deformation signals in wrapped interferograms, but further development is required to detect slow or small deformation patterns which do not generate multiple fringes in short duration interferograms. This study is the first to use machine learning approaches for detecting volcanic deformation in large datasets, and demonstrates the potential of such techniques for developing alert systems based on satellite imagery.

1 Introduction

Globally 800 million people live within 100 km of a volcano [*Loughlin et al.*, 2015]. Improvements in monitoring and forecasting have been shown to reduce fatalities due to volcanic eruptions [*Auker et al.*, 2013; *Mei et al.*, 2013] but a significant proportion of the ~1500 holocene volcanoes have no ground-based monitoring. Interferometric Synthetic Aperture Radar (InSAR) is a satellite remote sensing technique used to measure ground displacement at the centimeter-scale over large geographic areas and has been widely applied to volcanology [e.g. *Biggs and Pritchard*, 2017; *Pinel et al.*, 2014]. Furthermore, InSAR measurements of volcanic deformation have a significant statistical link to eruption [*Biggs et al.*, 2014]. Modern satellites provide large coverage with high resolution signals, generating large datasets. For example, the two-satellite constellation, Sentinel-1 A and B, offers a 6 day repeat cycle and acquires data with a 250-km swath at a 5 m by 20 m spatial resolution (single look). This amounts to >10 TB per day or about 2 PB collected between its launch in 2014 and June 2017 [*Fernández et al.*, 2017]. The explosion in data has brought major challenges associated with manual inspection of imagery and timely dissemination of information. Moreover, many volcano observatories lack the expertise needed exploit satellite datasets, particularly those in developing countries.

Machine learning technologies have been widely implemented in the field of computer science, where the computers use statistical techniques to learn a specific and complex task from given data. In the Earth Sciences, machine learning has been employed in several applications [*Lary et al.*, 2016], such as predicting earthquake magnitudes [*Adeli and Panakkat*, 2009], land surface classification [*Li et al.*, 2014], vegetation indices [*Brown et al.*, 2008], landslide susceptibility mapping [*Yilmaz*, 2010], etc. The techniques used previously include tree-based methods [*Wei et al.*, 2013], artificial neural networks [*Conforti et al.*, 2014], support vector machines [*Tien Bui et al.*, 2017] and Bayesian methods [*Totaro et al.*, 2016].

Here, we present a novel approach to detect volcanic ground deformation automatically from InSAR images. This approach brings together satellite-based volcano geodesy and machine learning algorithms to develop new ways of automatically searching through large volumes of InSAR images to detect patterns that may be related to volcanic activity. The proposed method works on 'wrapped' interferograms displayed as fringes each representing a set amount of displacement, equal to half the radar wavelength. However, these interferograms also contain artifacts associated with atmospheric conditions, and at volcanoes the effect of stratified atmospheric water vapour can be particularly difficult to distinguish from ground deformation [e.g. *Ebmeier et al.*, 2013b; *Parker et al.*, 2015a]. In this paper, we extract the spatial characteristics of the interferograms using deep convolutional neural networks (CNN) – biologically-inspired architectures that comprise multiple layers of neural connections that have learnable weights and biases [*Krizhevsky et al.*, 2012]. Similar approaches have been highly successful when applied to the analysis of visual imagery, image classification, object detection and tracking [*LeCun et al.*, 2015], and could ultimately be used in near-real time to detect volcano deformation and inform the local volcano observatories.

2 Background: Machine Learning Algorithms

Machine learning is a generic term for the automatic discrimination of input patterns into learnt or defined classes, originally introduced in the 1950s [*Samuel*, 1959]. For the case of volcanic unrest classification, the input is InSAR interferograms and the output will be one of two classes: unrest or no unrest (or the likelihood of each). Machine learning techniques can be separated into two categories, supervised and unsupervised methods. Supervised methods learn representations of the output classes using labelled ground truth examples of those classes [*Kotsiantis*, 2007] (i.e. in this case volcanic unrest and no volcanic unrest), whereas unsupervised methods cluster together similar groups in the data without any ground truth [e.g. *Zanero and Savaresi*, 2004]. In this study, we focus on supervised methods, particularly deep Convolutional Neural Networks (CNNs) [*Rumelhart et al.*, 1986; *Krizhevsky et al.*, 2012; *Szegedy et al.*, 2016] and Support Vector Machines (SVMs) [*Christianini and Shawe-Taylor*, 2000].

Support Vector Machines (SVMs) typically use hand-defined inputs such as intensity distributions and Gabor features extracted from the input images [*Chang and Lin*, 2011]. SVMs classify using a "maximum margin" technique and are able to linearly distinguish two or more classes. However, using the "kernel trick" the input domain is projected into a (possibly infinite) higher dimensional space to provide very effective non-linear classification [*Burges*, 1998]. The main advantages of SVMs are that the training process does not require a truly large dataset (large in this context can be considered to be on the order of 10,000 or more data points). The SVM process is also fast even for machines without a graphics processing unit (GPU). However, in many supervised classification problems with large ground truth datasets, deep networks such as CNNs often outperform shallow machine learning algorithms such as SVMs [*Goodfellow et al.*, 2016].

Convolutional Neural Networks (CNNs) are a class of neural networks that employ locally connected layers that apply convolution between a kernel (filter matrix) and an internal signal and are most commonly used for image recognition and classification. The deep, hierarchical and densely connected nature of CNNs enable them, not only to classify, but also to generate discriminating features of progressive complexity from the input to the output layers [*Jia et al.*, 2014]. For image based classification, the first layers convolve small spatial regions with learnt blocks of weights. These weight blocks can be considered to be feature extractors and often resemble early vision basis functions found in the human visual cortex (i.e. similar to 2D Gabor functions) [*Matsugu et al.*, 2003]. The output of these layers are often integrated (or "pooled") before connection to lower layers. The convolutional layers are commonly then connected to dense layers of fully connected neurons leading to a final classification (often using an output activation function such as *softmax* [*Goodfellow et al.*, 2016]). All neurons within the convolutional and fully connected layers are defined by weights and a bias from the connected neurons one layer above. Depending on the architecture, all layers use activation functions such as *tanh* (the hyperbolic tangent) or *ReLU* (Rectified Linear Unit) to introduce non-linearity into the networks [*Agostinelli et al.*, 2015]. The weights in all layers are initiated in training with non-zero random or pseudo-random values.

All weights are then modified using a batch based iterative back propagation method using a testing dataset with associated ground truth. To prevent overfitting, regularisation techniques such as "dropout" are used to ensure the network is able to effectively generalise [*Srivastava et al.*, 2014]. The effective training of deep CNN networks is both extremely computationally expensive and requires very large training datasets [*Simonyan and Zisserman*, 2014]. It is therefore common to pretrain convolutional layers in an unsupervised fashion, followed by supervised fine-tuning [*Erhan et al.*, 2010]. Several high-performance pretrained models have been employed to serve a specific purpose, such as AlexNet [*Krizhevsky et al.*, 2012] and ResNet [*He et al.*, 2016]. Figure 1 illustrates the architecture of an example CNN (Alexnet). This figure shows the 2D convolution layers and the 1D output linear layers and how they are connected giving a hierarchical representation across all the layers. The input to Alexnet (as shown in the figure) is a 2D image of 224x224 pixels.

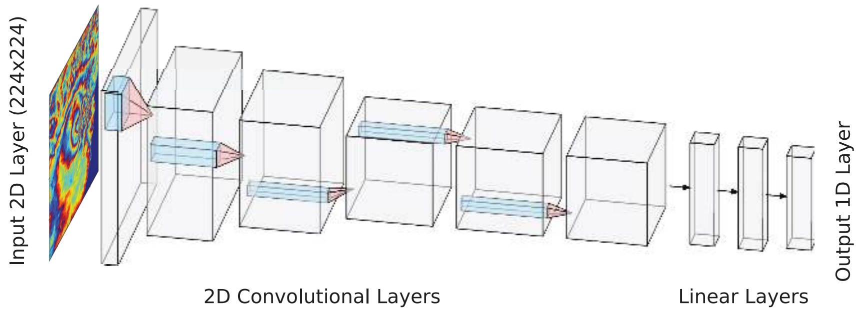

Figure 1. AlexNet CNN for InSAR input. The first Block from the left is the input image with a size of 224×224. It is followed by five 2D convolutional layers (filter sizes of 11×11, 5×5, 3×3, 3×3 and 3×3) with ReLU and max-pooling (filter size of 3×3). The last three are fully connected layers (linear layers generate features with a length of 4096). The blue block shows the neighbouring pixels associated in each convolution to produce one value for the next layer.

3 InSAR Dataset

The first Sentinel-1 satellite (S1A) was launched in 2014 and the mission ensures Earth's observations for the next 25 years with repeat intervals of 6-24 days globally. The data is freely-available in near real-time making it ideal for routine volcano monitoring. The global dataset used in this study consists of 30,249 interferograms covering ~900 volcanoes in 2016-2017. The interferograms were processed with the automated InSAR processing system LiCSAR (http://comet.nerc.ac.uk/COMET-LiCS-portal/) developed by the Centre for Observation and Modelling of Earthquakes, Volcanoes and Tectonics (COMET). Each acquisition is connected to the three preceding acquisitions, forming a trio of interferograms of increasing time-span. We crop the images to a region spanning 0.5° in latitude and longitude for each of the ~900 volcanoes. We include volcanoes in temperate, tropical and arid environments with morphologies ranging from steep stratovolcanoes, to large calderas and small islands (Figure 2a). The dataset is weighted towards the European Volcanoes where images are acquired every 6 days and the LiCSAR system has been running the longest (2016-2017). Temporal baseline of the interferograms ranges from 6 to 120 days, including one third of the dataset with timespans of 6 and 12 days (Figure 2b).

A major challenge for both manual and automated InSAR monitoring systems is distinguishing deformation signals from atmospheric artifacts which can also generate concentric fringes around volcanoes, particularly those with steep topography [e.g. *Ebmeier et al.*, 2013b; *Pinel et al.*, 2014]. Several approaches have been proposed to correct these artifacts, with external data sources such weather models, or GPS tropospheric delays, or by apply-

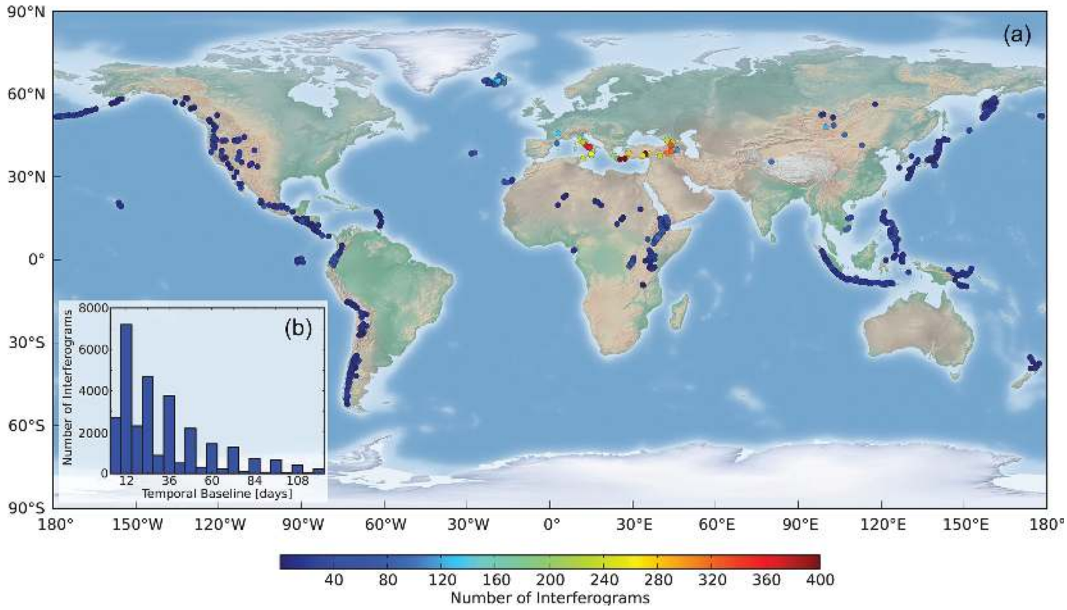

Figure 2. (a) Worldmap showing the spatial distibution of the dataset: colour dots indicate the number of Sentinel-1 interferograms calculated for each volcano. Notice that the number of data is largest for European volcanoes. (b) Histogram showing the distribution of the temporal baseline (timespan between the master and the slave acquisition) of our dataset.

ing statistical approaches to phase-elevation correlations or time-series [e.g. *Bekaert et al.*, 2015; *Li et al.*, 2005; *Jolivet et al.*, 2014]. The quality of atmospheric correction is highly dependent on geographical location and is hence variable [*Parker et al.*, 2015a]. Furthermore, atmospheric corrections can only be applied to unwrapped interferograms, and unwrapping is computationally expensive, slow and can introduce phase errors. For our initial, proof-of-concept study, we chose to use wrapped, uncorrected interferograms and test the ability of our approach to discriminate between deformation and atmospheric signals.

To provide ground-truth information for training and verification of supervised classification systems, it is necessary to manually identify a selection of interferograms where several fringes can be attributed to volcanic deformation. Even though there are >30,000 interferograms in our Sentinel-1 dataset, the majority are short-duration inteferograms covering volcanoes that are not deforming, or are deforming slowly. Identifying a sufficient number of positive images in the Sentinel-1 dataset is challenging, so we pretrain the network using an older archive of interferograms from the European Space Agency's Envisat satellite. Several possible datasets exist, including over the Main Ethiopian Rift [*Biggs et al.*, 2011], the Kenyan Rift [*Biggs et al.*, 2009], the Central Andes [*Pritchard and Simons*, 2004a] and the Southern Andes [*Pritchard and Simons*, 2004b]. All of these contain 1) multiple volcanic systems displaying persistent deformation at variable rates, and 2) areas which are not deforming but show a range of features including incoherence and atmospheric artifacts (Figure 3). We chose to use a dataset over the Main Ethiopian Rift (MER) for convenience. The Envisat background mission (2003-2010) acquired three to four images per year over the Main Ethiopian Rift, and has been used to identify deformation at 4 volcanoes previously considered dormant: Alutu, Corbetti, Bora and Haledebi [*Biggs et al.*, 2011]. These interferograms are a good test case The rates of deformation are several centimetres per year, which means that over the time period of the interofergrams (variable, but typically several months), the interferograms show several fringes of deformation.

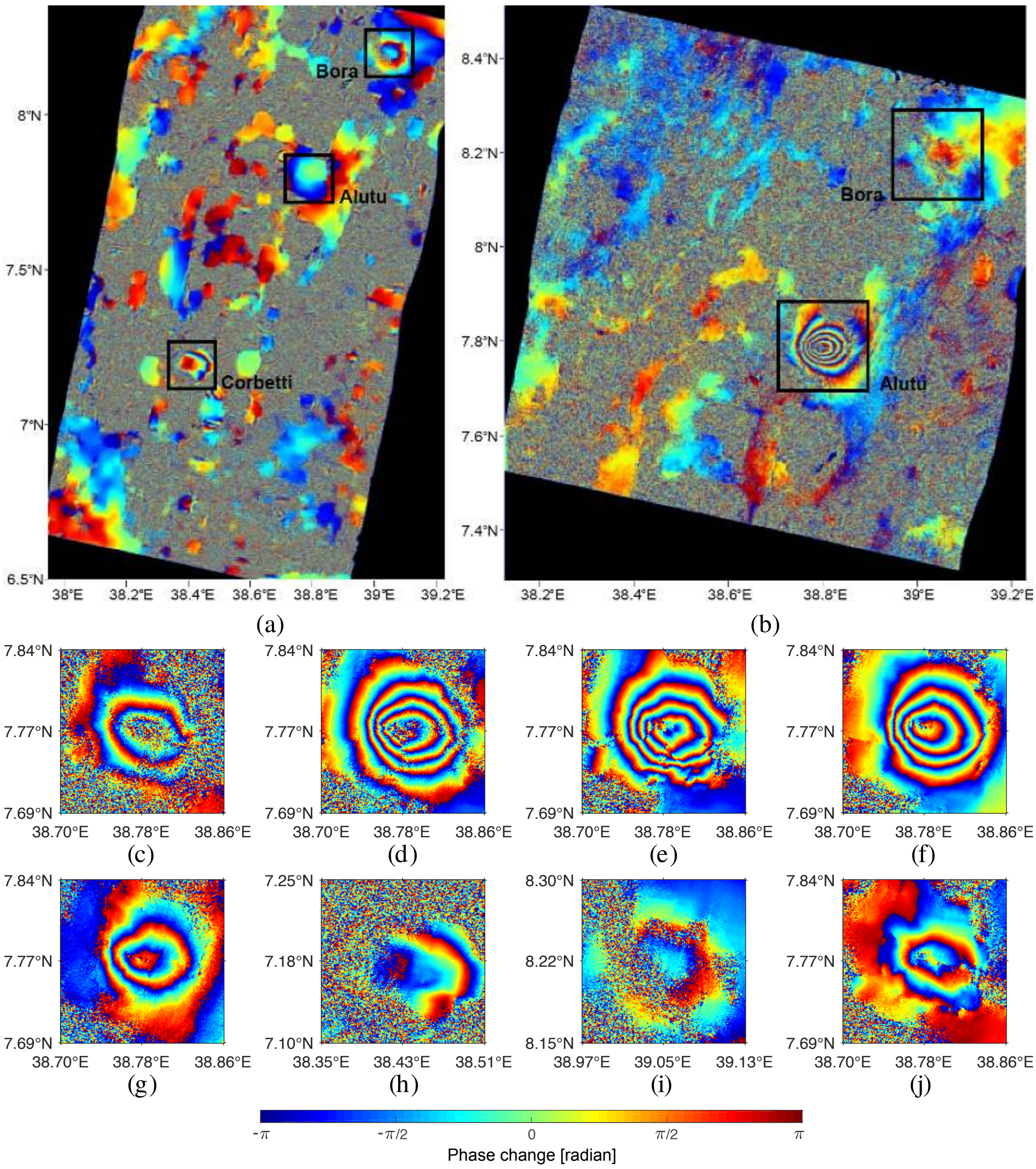

Figure 3. Archive dataset over the Main Ethiopian Rift produced using the Envisat satellite. This is used to increase the number of positive samples for training. (a) covering Bora, Alutu and Corbetti (20080827-20100623), (b) showing uplift at Alutu (20071226-20081210). The bottom row shows ground deformation signals at (c) Alutu (20040922-20080514), (d) Alutu (20071226-20090114), (e) Alutu (20071226-20081105), (f) Alutu (20080514-20090114), (g) Alutu (20080827-20081210), (h) Corbetti (20080827-20100623), (i) Bora (20080827-20100623) (j) Alutu (20081105-20100728) [*Biggs et al.*, 2011]. Each colour cycle (fringe) represents 2.8 cm of displacement in the satellite line-of-sight.

Despite the small number of examples, it is important to train the network using some Sentinel-1 data to account for differences in processing strategy and atmospheric behaviour. A small dyke intrusion at Erte Ale volcano (Ethiopia) occurred in January 2017 associated with the overflow of the lava lake [*Xu et al.*, 2017] and interferograms spanning this event shows 4 fringes of deformation (Figure 4 a). Interferograms of Etna volcano (Italy) spanning October 2016 show fringes potentially related to an intrusive event (Figure 4 b-c); the National Institute of Geophysics and Volcanology (INGV) reported the openning of a new volcanic vent on 7 August and an explosion at Bocca Nova on 10 October. Interferograms from other time periods at Erte Ale and Etna show multiple fringes that are atmospheric in origin (Figure 4 e-f). Cerro Azul and Fernandina volcanoes (Galapagos) have been deforming during 2017 [*Bagnardi*, 2017] and typically show several fringes of deformation in a single interferogram. Several other volcanoes are known to be deforming slowly during this time period, for example, Medicine Lake, USA which has been subsiding for ~60 years at ~10 mm/yr [*Parker et al.*, 2015b] and Laguna del Maule, Chile [*Singer et al.*, 2014] and Corbetti Ethiopia [*Lloyd et al.*, 2018], which are uplifting at rates of >6 cm/yr. However, in short interferograms, these slow rates of deformation are not sufficient to produce multiple fringes of deformation, and we do not attempt to identify them in the current study. We use interferograms spanning the intrusions at Erte Ale and Etna and to train the network (Figure 4 e-f), and include the Galapagos volcanoes in the test dataset to assess detection capability. For our initial runs, we do not flag interferograms with atmospheric artifacts as negative results, instead testing the ability of to distinguish deformation patterns based on positive examples alone.

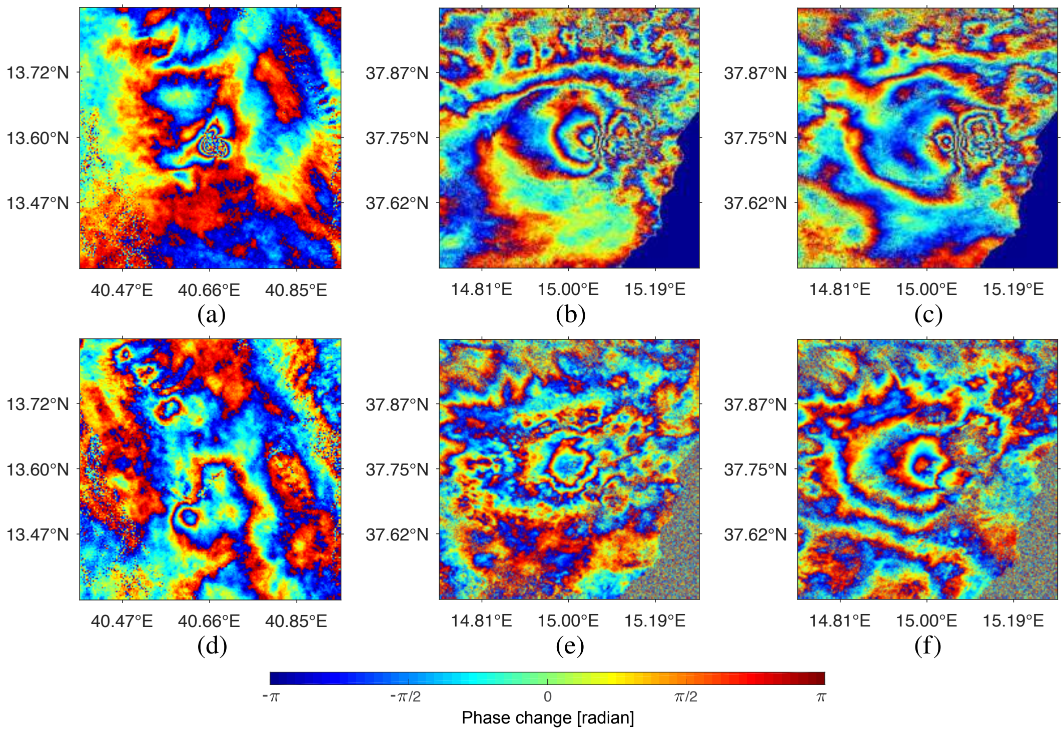

Figure 4. (a-c): Volcanic ground deformation signals in Sentinel-1 inteferograms at a) Erta Ale (20170104-20170209) [*Xu et al.*, 2017], b) Etna (20161003-20161015) and c) Etna (20161003-20161021). (d-f): Atmospheric signals at d) Erta Ale (20170925-20171031), e) Etna (20170916-20171010) and f) Etna (20170916-20170928) . Each colour cycle (fringe) represents 2.8 cm of displacement in the satellite line-of-sight.

Table 1 shows the list of volcanoes used as positive samples in the training process (Section 4.2). The negative samples are generated from both non-deformation and background as described in Section 4.1.

Table 1. List of volcanoes showing deformation and used in the training process. Note that the number of interferograms is before applying data augmentation (which is the process of increasing the number of positive samples to be balanced with that of the negative samples in the training dataset).

Training process	Volcano name	Type	period	# interferograms
Initial (Envisat)	Alutu	Stratovolcano	2003-2010	158
	Bora	Pyroclastic cone	2003-2010	52
	Corbetti	Caldera	2003-2010	44
	Haledebi	Fissure vent	2003-2010	46
Retrained (Sentinel)	Etna	Stratovolcano	2016	2
	Ale Bagu	Stratovolcano	2017	3
	Bora Ale	Stratovolcano	2017	3
	Cerro Azul	Shield	2017	8
	Erta Ale	Shield	2017	3
	Hayli Gubbi	Shield	2017	3
	Sierra Negra	Shield	2017	17

4 Method Development

The proposed framework for using machine learning to identify volcanic deformation in interferograms is shown in Figure 5. For the training process, each image is processed as described in Section 4.1 and then fed into the CNN to learn ground deformation characteristics (positive class) against those of background, atmosphere and noise (negative class). We conducted initial tests on a range of pretrained CNNs and SVMs using small archive and test datasets from Envisat and Sentinel-1 respectively.

4.1 Data preparation

The values of wrapped interferograms vary between $-\pi$ and π and they are typically displayed with colours (red, green, and blue intensities). For the purposes of machine learning, we first convert the wrapped interferogram into grayscale image, i.e. the pixel value in the range of $[-\pi, \pi]$ is scaled to $[0, 255]$ or $[-125, 125]$ if zero-centre normalisation is required (Figure 6b). Subsequently, each training image is divided into patches equal to the input size the CNN (e.g. 224×224 pixels for AlexNet [*Krizhevsky et al.*, 2012]). The patches overlap by half their size (Figure 6c).

We then employ Canny edge detection [*Canny*, 1986], where a Gaussian filter is firstly applied to remove noise, then double thresholding is applied to the intensity gradients of the image. As the wrapped-phase interferograms shows strong edges where the phase jumps between -π and π, the Canny operator can straightforwardly extract fringes occurring from volcano deformation (Figure 6 c). As the number of background areas (negative samples) is significantly larger than those associated with volcano deformation (positive samples), only the patches in which strong edges have been detected are used. Since areas without strong edges are unlikely to contain volcanic deformation they are instantly defined as background without classification by the CNN.

For machine learning, balancing the number of training samples between classes is very important but we have only 300 positive examples. There are over 100 times more negative patches containing strong edges than positive patches. Therefore, we increase the number positive patches for training using a data augmentation approach [*Krizhevsky et al.*, 2012]. We generate more positive patches by i) shifting every 10 pixels around the volcano; ii) flipping horizontally and vertically; iii) rotation through angles of 22.5°, 45°, 67.5° and

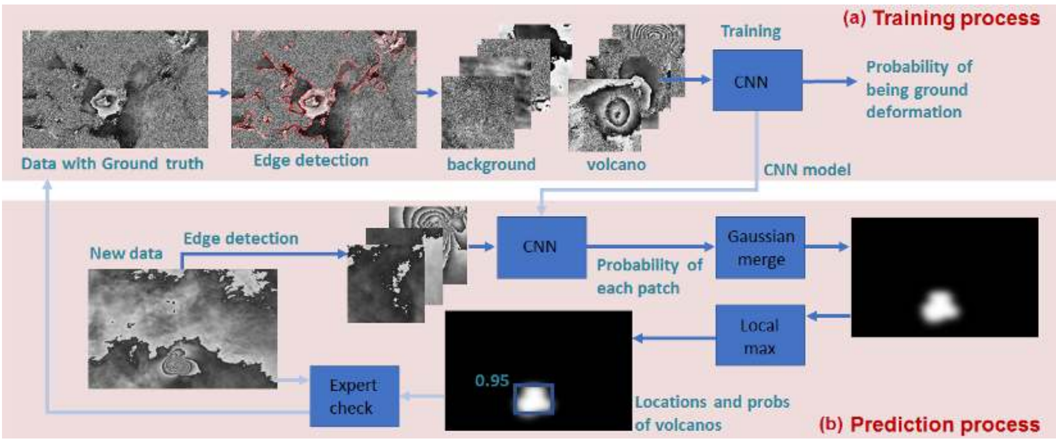

Figure 5. Diagram of the proposed framework, showing (a) the training process and (b) the prediction process. The training process starts with data with ground truth (labelled as "1" or "positive", where deformation is present and "0" or "negative" in other areas, e.g. Figure 6a). Then, edge detection is applied so that only the areas with large phase changes are considered. These areas are subsequently divided into 2 classes of patches and fed to the CNN for training. For the prediction process, the new interferogram is divided into overlapping patches and the patches with strong edges are tested with the trained CNN model, giving the probability P of being ground deformation. The probabilities of all patches are merged with Gaussian weights. The highest probability P_{max} and its location are provided for the development of an alert system. Finally the expert checks the result and the positives are employed to retrain the CNN for better performance.

90°; and iv) distorting the shape of deformation by varying scales along the x and y axes of the affine transformation. This data augmentation technique increases the 300 positive samples initially identified in the Envisat dataset to approximately 10,000 positive patches (Figure 6d). We randomly select negative patches so that the numbers are balanced.

4.2 Initial tests

We employ a transfer learning strategy by fine-tuning a pretrained network. This approach is faster and easier than training a network with randomly initialised weights from scratch (which could take months for training). Parameters and features of these networks have been learnt from a very large dataset of natural images thereby being applicable to numerous specific applications. The last two layers are replaced with a fully connected layer and a softmax layer to give a classification output related to volcanic unrest. The learning rates of the new layers are defined to be faster than the transferred layers. We set the maximum number of epochs to 50 and the batch size to 100. The output of the softmax layer, the top layer of the CNN, is the probability P of the patch being a positive result. The probabilities for each patch are merged with Gaussian weights ($\mu = 0, \sigma = 1$), where μ and σ are the mean and the standard deviation, respectively.

Initially, we use the Envisat archive to test three popular pretrained CNN architectures: AlexNet [*Krizhevsky et al.*, 2012], ResNet50 [*He et al.*, 2016] and InceptionV3 [*Szegedy et al.*, 2016]. We also test a support vector machine (SVM) classifier based on textural features following *Anantrasirichai et al.* [2013]. The objective results were evaluated using a receiver operating characteristic curve (ROC curve), which illustrates the performance of the identification method by comparing true positive rates (TPR) and false positive rates (FPR). The TPR (or sensitivity or recall) is the ratio between the number of positive samples correctly categorised as positive and the total number of actual positive samples. The false positive rate is calculated as the ratio between the number of negative samples wrongly cat-

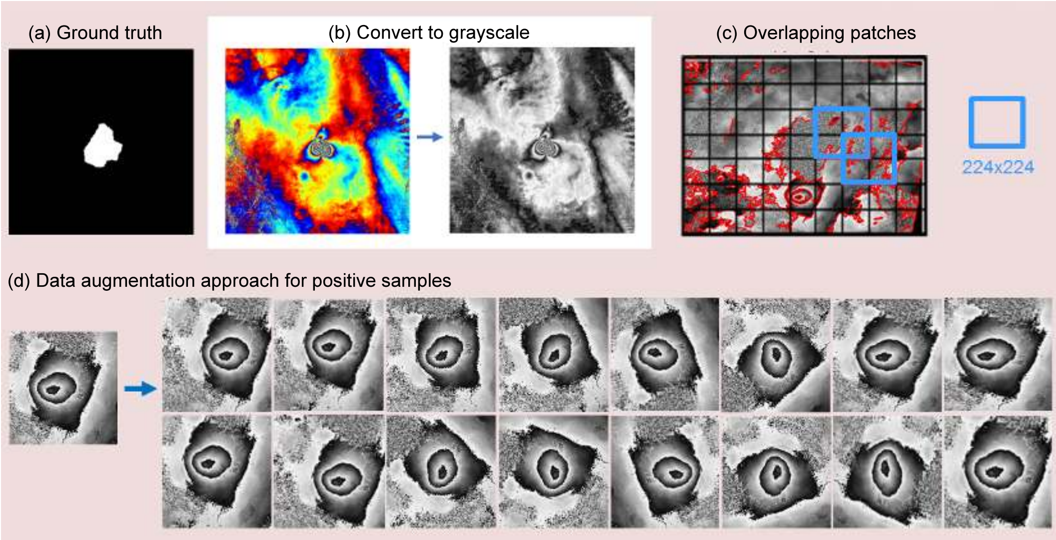

Figure 6. Data preparation process, comprising (a) ground truth, labelled by the experts – the white and the black areas are the positives and the negatives, respectively; (b) value scaling, where the phase varying between $-\pi$ to π is converted to grayscale values between 0 to 255, which suit the CNN inputs; (c) overlapping patches, generated by dividing the image with the size of the CNN input (e.g. 224×224 pixels) and positioning by overlapping by half this size (e.g. blue boxes in the figure). The patches without edges are defined to be negatives instantly and will not be used for training; (d) data augmentation, where the number of positive patches is increased to match that of the negative patches, which is done using rotations, flips, distortions and pixel shifts.

egorized as positive and the total number of actual negative samples. The area under curve (AUC) is the integrated area under the ROC curve. Better performance results in higher AUC values (maximum = 1), achieved through a high TPR and low FPR, such that most true ground deformations are correctly identified and only a few backgrounds are falsely identified as positive results.

Figure 7 shows the ROC curve for a 2-fold cross validation, where half of the dataset is employed for training and the other half is used for testing, then they are swapped, and the results are averaged. We also calculate the accuracy and true negative rate (TNR), for comparison: the accuracy is the proportion of correctly predicted results amongst all testing samples, whilst the TNR measures the proportion of negative samples that are correctly identified. AlexNet achieves 0.995, 0.925, 0.899 and 0.992 for AUC, accuracy, TPR and TNR respectively. It outperforms ResNet50, InceptionV3 and texture features with SVM by approximately 8%, 5% and 11% on the average of these four metrics, respectively.

Next, we employ the initial model based on the AlexNet CNN and the SVM, trained by Envisat described above, and retrain it by including a subset of the Sentinel-1 dataset. We use interferograms covering Erta Ale, Ethiopia, and Etna, Italy, which include both deformation and atmospheric signals as previously discussed (Figure 4). We evaluate these tests using 2-fold cross validation and compute the accuracy, TPR and TNR as before (Table 2). For both Erta Ale and Etna, the AlexNet CNN outperforms the SVM. The results for Erta Ale show exceptional performance, with an accuracy of 0.994 for the CNN, whilst those for the Etna area are less good (accuracy of 0.871), probably due to atmospheric interference.

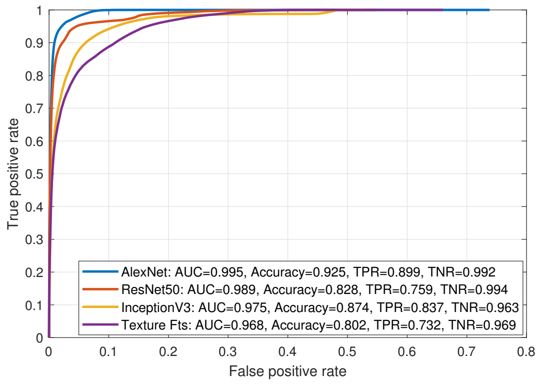

Figure 7. ROC curves for the 2 folds of cross validation using Envisat dataset to train the networks. These compare classification performances between AlexNet, ResNet50, InceptionV3 and texture features (Texture Fts). Four metrics are also computed, namely the area under the curve (AUC), the accuracy, the true positive rate (TPR), and the true negative rate (TNR). AlexNet achieves the best performance, followed by ResNet50, InceptionV3 and texture features, respectively.

Table 2. The average results of 2-fold cross validation of the AlexNet and the SVM when trained with the Envisat dataset and the Sentinel-1 dataset of Erta Ale, Ethiopia, and Etna, Italy.

Region	methods	Accuracy	TPR	TNR
Erta Ale	CNN	0.994	1.000	0.988
	SVM	0.985	0.982	0.985
Etna	CNN	0.871	0.747	0.981
	SVM	0.742	0.654	0.783

5 Application to the Global Dataset

In the previous sections, we have demonstrated that deep learning with CNNs has significant potential to capture the characteristics of volcano deformation present in interferograms despite the challenges of large scale, heterogeneity, and non-stationary distribution that such problems typically present for deep learning [*Chen and Lin*, 2014]. In this section, we apply our pre-trained CNN to the global dataset of ~900 volcanoes and 30,249 interferograms described in section 3, using the framework illustrated in Figure 5. Following an initial run, we use expert analysis of the results to retrain the model and rerun it.

The CNN-training process was run on a graphics processing unit (GPU) at the High Performance Computing facility (BlueCrystal) at the University of Bristol. The initial and retrained models were completed in approximately 38 hours and 26 hours, respectively. The retraining process was faster, despite using a larger training dataset (Envisat dataset plus some positive results of Sentinel dataset), because the weights and biases of the network are initialised with values closer to the optimum. The prediction process for each 500×500 pixel interferogram took ~1.5 seconds (~10 hours for 30,249 interferograms). In theory, the CNN model can be retrained whenever a new result is confirmed by an expert, a process that would likely focus on true positive and false negative results (i.e. if a real deformation event is missed). However, the training dataset requires balanced numbers of positive and negative samples, and since false positives occur more frequently than false negatives, care is required to augment the deformation samples, ensuring that data points are positioned to prevent overfitting.

Table 3. Classification results of 30,249 interferograms showing the total number of predicted positives, the numbers of confirmed true positives, confirmed false positives, and the number of results required further analysis. This shows that the performances of the CNN model is improved from 2.85% to 37.5% in term of the positive predictive value (PPV – the fraction of true positives among all retrieved positives) when the model is retrained with the confirmed positives of the Sentinel-1 dataset.

Model	# Positives	# True Positives	# False Positives	# Unconfirmed
Initial	1369	39	894	435
Retrained	104	39	-	65

For each run, we calculate the number of total positive results (Positive), confirmed true positives (TP), confirmed false positive (FP) and results requiring further analysis (Unconfirmed) (Table 3). The initial model run identified 1368 positive results, of which 39 were considered to be true positives, including the examples at Sierra Negra and Cero Azul in the Galapagos that were included as a test, and additional interferograms showing deformation at Etna (Figure 8 a-c). These examples all have detection probabilities >0.999. Of the remaining 1329 "positive" interferograms, 894 were quickly identified as false positives, mostly small islands and turbulent atmospheric artifacts, which typically have detection probabilities less than 0.85 (Figure 8 d-f). The true positive and false positive results were then fed back to the CNN to retrain the model.

The retrained model identified 104 positive results, including the 39 true positives identified initially. The other 65 examples all contained concentric fringes around the volcano, and even experts were unable to determine from a single inteferogram whether the fringes were caused by volcanic deformation or atmospheric artefacts. This includes Tambora Indonesia, Alayta Ethiopia, Adwa Ethiopia and Etna Italy (Figure 9) which are all high relief stratovolcanoes. The merged probabilities assigned to these detections are 0.965, 0.867, 0.733 and 0.953 respectively, slightly lower than those assigned to the true positives. By calculating the correlation between the phase and the elevation and looking at pair-wise logic in the time series [*Ebmeier et al.*, 2013b], we finally conclude that these 65 signals were caused by atmospheric artifacts.

The CNN identified over 30,000 negative results, but manually searching through all these for false negatives is not feasible. However, we have checked all scenes associated with reported eruptions during this time period [*Global Volcanism Program*, 2013]. The only example with a visible fringe pattern was detected at Ulawun, Papua New Guinea (20170604-20170722), which erupted between 11 June 2017 - 3 November 2017 (Fig. 10). The full interferogram and a zoomed-in version showing the fringes are shown in Figure 10c and 10d, respectively. Our framework did not detect this signal because the visible fringe area is relatively small compared to those in the training positive patches, and it is surrounded by noise. After applying several convolutions and pooling in the CNN, the features of the noise become dominant and it is classified as a negative result.

6 Discussion

The majority of volcanoes worldwide have little or no ground-based monitoring. Satellite systems, such as InSAR, have the potential to measure surface deformation at volcanoes globally, but until now the utility of these systems has been limited by the acquisition strategy and data policy of the space agencies. The launch of Sentinel-1 is providing unprecedented data access, but poses new challenges, as more data is available than can be analysed by manual inspection. This paper demonstrates that machine learning using deep convolutional neural networks (CNNs) has the capability to identify rapid deformation signals from

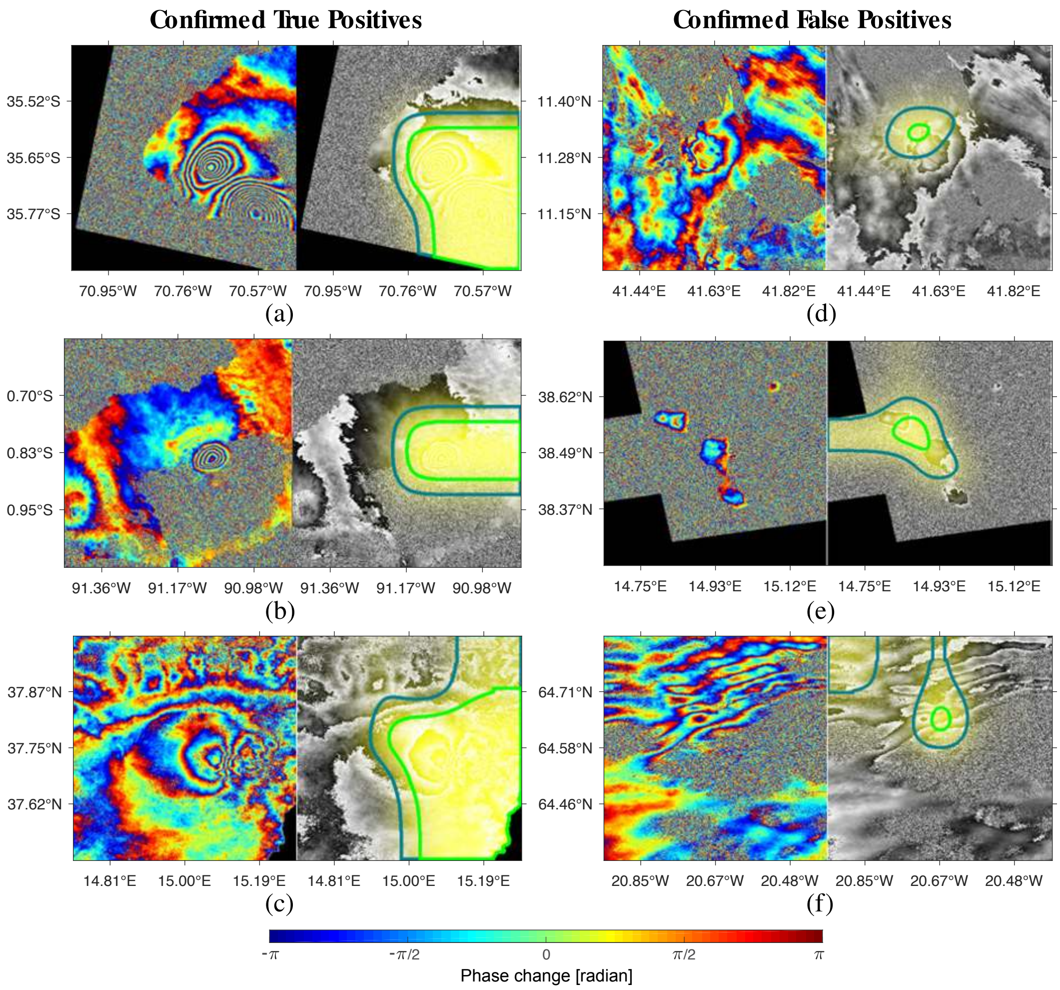

Figure 8. Results of the initial model showing the original image (left) and overlaid with probability of being volcanic deformation (right). Confirmed true positive results from (a) Cerro Azul, Galapagos (20170320-20170401) P_{max} =1, (b) Sierra Negra, Galapagos (20170519-20170718) P_{max} =1, (c) Etna, Italy (20161003-20161015) P_{max} =0.999. Confirmed false positive results from (d) Dama Ali, Ethiopia (20170511-20170710) P_{max}=0.812, (e) Lipari, Italy (20170212-20170224) P_{max} =0.824, (f) Prestahnukur, Iceland (20170517-20170722) P_{max}=0.815. The brighter yellow means higher probability. Areas inside dark and bright green contours are where P>0.5 and P>0.8, respectively. Each colour cycle (fringe) represents 2.8 cm of displacement in the satellite line-of-sight.

a large dataset of interferograms. This is a proof-of-concept, and further development is still required to develop an operational global alert system for volcanic unrest based on satellite observations of surface deformation. In this section, we discuss the limitations of the current process and outline future developments that would lead to the development of an operational system.

The first component of any automated alert system is the automatic processing of satellite data. The currently available Sentinel-1 dataset has a relatively small number of interferograms that show deformation, meaning a limited number of positive samples are available for training. For this initial test, we have resorted to using examples from Envisat and data augmentation approached to increase the number of available positive results for training.

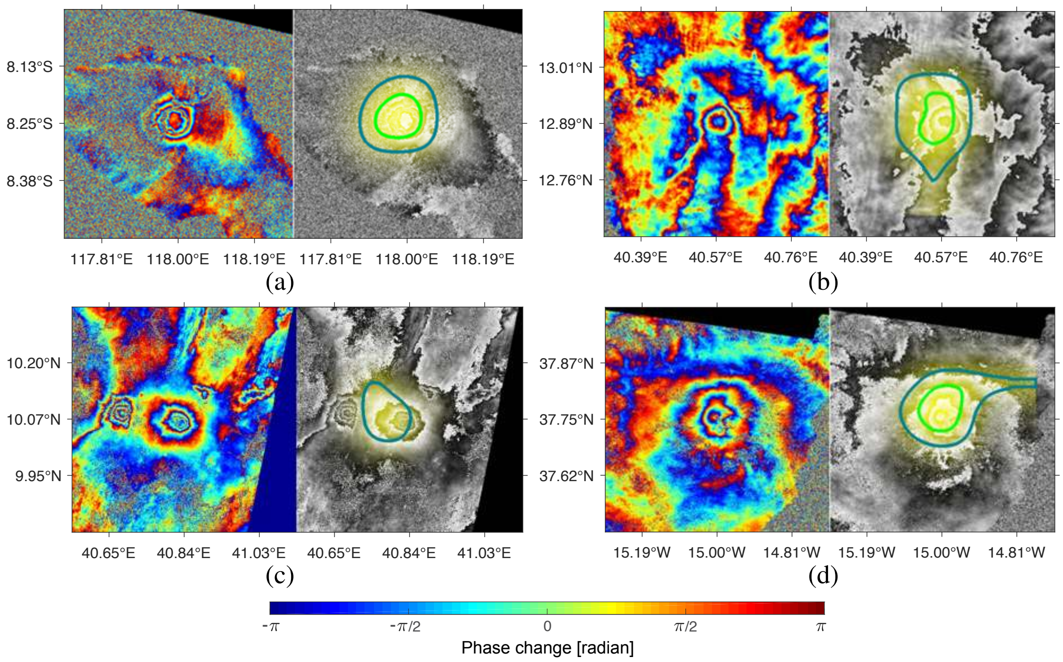

Figure 9. Unconfirmed positive results from the retrained model showing the original image (left) and overlaid with probability of being volcanic deformation (right) of (a) Tambora, Indonesia (20170718-20170730) P_{max}=0.965, (b) Alayta, Ethiopia (20170329-20170528) P_{max}=0.867, (c) Adwa, Ethiopia (20170516-20170609) P_{max}=0.733 (Note: Ayelu on the left of Adwa is not identified as deformed ground as P=0.06), (d) Etna, Italy (20170922-20170928) P_{max}=0.953. The brighter yellow means higher probability. Areas inside dark and bright green contours are where P>0.5 and P>0.8, respectively. Each colour cycle (fringe) represents 2.8 cm of displacement in the satellite line-of-sight.

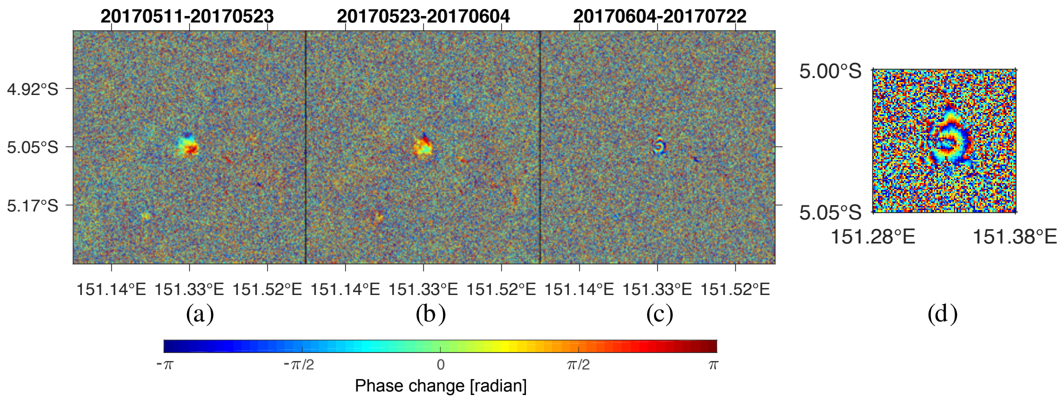

Figure 10. Three interferograms at Ulawan, Papua New Guinea, (a) 20170511-201705523, (b) 20170523-20170604, (c) 20170604-20170722. The last one (c) shows possible deformation. (d) The zoom-in area of (c). Each colour cycle (fringe) represents 2.8 cm of displacement in the satellite line-of-sight.

However, these may not truly reflect the characteristics of global volcanic deformation. As the system continues running, more positive samples will become available and the model is retrained, the system performance will improve.

The European Space Agency posts raw Sentinel-1 data to their website within hours of acquisition, but limited bandwidth makes this data access route unsuitable for automated systems operating on a global scale. The LiCSAR system uses the archive held by the UK Centre for Environmental Data Analysis (CEDA) which typically has a latency period of a few weeks. This latency is well suited for routine surveys of ground deformation [e.g. *Pritchard and Simons*, 2004c; *Biggs et al.*, 2011; *Chaussard and Amelung*, 2012; *Ebmeier et al.*, 2013a], which can be used for motivating changes in long-term monitoring strategies, but would be too slow for crisis response [*Ebmeier et al.*, 2018]. Automated processing of archived data could be supplemented by direct download for a limited number of volcanoes which are considered to be high threat because of changes in behavior identified by other methods, such as seismic swarms. Once trained, the CNN runs in a matter of seconds, and would not add noticeably to the time taken for data to be communicated. The retraining process is slower and could be undertaken periodically, or when particularly significant events are detected, such as a new type of deformation pattern.

The current proof-of-concept study demonstrates the ability of CNNs to identify rapidly deforming systems that generate multiple fringes in wrapped interferograms. For a 12-day C-band interferogram this corresponds to a deformation rate of 1.8 m/yr. Such high rates are typically only observed for very short periods, and are often associated with dyke intrusions or eruptions [*Biggs and Pritchard*, 2017]. There are several possible adaptations that would enable a machine learning system to detect slower rates of deformation associated with sustained unrest. The first option is to generate long time-span interferograms, which will increase the number of fringes per image where deformation is sustained. For example, in a year-long interferogram, the average deformation rate required to generate two fringes is only 6 cm/yr. The second option is to develop a machine learning approach capable of detecting deformation in unwrapped data. However, fringes are ideally suited to machine-learning approaches because the high-frequency content is easy to identify using edge-detection methods and provides strong features for distinguishing deformation from other signals. Anomaly detection techniques may be suitable for classifying unusual events in unwrapped data [e.g *Gaddam et al.*, 2007], but care needs to be taken when scaling the unwrapped data to the settings of the pretrained network (e.g. 0-255), as clipping large magnitudes may loose information.

For our initial tests, we have chosen to use wrapped interferograms, as although several unwrapping algorithms exist [e.g. *Chen and Zebker*, 2001; *Goldstein et al.*, 1988], they are computationally expensive, particularly in areas of low or patchy coherence. Although the automatic processing of unwrapped interferograms on a global basis is challenging, there are several advantages. In general, stacking multiple short-time period interferograms will produce more coherent results than directly processing longer time-span interferograms [e.g. *Biggs et al.*, 2007]. However, there are exceptions, particularly where the level of coherence is seasonally variable, for example, due to snowfall, and further analysis of global patterns of coherence is required in order to determine the most appropriate strategy for automating this. Once the interferogram has been unwrapped, it can be re-wrapped at any chosen interval, meaning that higher fringe rates can be artificially generated. The optimal fringe rate will depend on the ability of the CNN to distinguish the spatial patterns of fringes as increasing the rate will also increase the number of fringes associated with turbulent atmospheric artifacts. Using unwrapped interferograms also improved the ability to identify atmospheric signals, either by applying a direct correction or as a secondary stage. Weather models are available globally and services such as the Generic Atmospheric Correction Online Service (GACOS) exist, but are not yet routinely applied on a global basis [*Yu et al.*, 2018]. A more efficient approach would be to use the weather models as a secondary stage, once the CNN has identified a smaller subset of 'positive' results.

The final challenge is ensuring that information is provided to the appropriate authorities in a timely and useful manner. The proof-of-concept algorithm reduces the number of interferograms that require manual inspection from >30,000 to 104, but expert analysis is still required to distinguish deformation from some types of atmospheric artifacts, and to interpret the deformation patterns in terms of source processes. Although there is a strong statistical link between satellite observations of deformation and eruptions, *Biggs et al.* [2014]'s global study found that only about half of deforming volcanoes erupted on a decadal timescale. Therefore, these alerts should be considered flags for further investigation using complementary datasets, rather than indicators of impending eruption. The ability of volcano observatories to interpret InSAR data is highly variable between countries. The algorithms developed here provide a probability that a given interferogram contains surface deformation, but further capacity building will be required before many volcano observatories, particularly those in developing countries, are able to use this information to influence alert levels or long-term monitoring strategies. Identifying all volcanic ground deformation signals will expand our understanding of the behaviour of a wide range of magmatic systems, and improve eruption forecasting in the future.

7 Conclusions

This paper is the first to demonstrate the capability of machine learning algorithms for detecting volcanic ground deformation in large sets of InSAR data. The proposed method was developed using a current popular machine learning algorithm for image classification – convolutional neural network (CNN). Our classification model was initialised with archive data from the Envisat mission using the pretrained CNN, AlexNet. It was then applied to a Sentinel dataset consisting of over 30,000 images at 900 volcanoes. After an initial run, expert classification of the positive results were used to retrain the network and the classification performance was improved, increasing the proportion of correctly identified deformations amongst all positive results from 2.85% to 37.5%. This retrained network reduced the number of interferograms that required manual inspection from >30,000 to ~100 and more training is likely to improve the performance yet further. These results indicate that machine learning algorithms combined with automated processing systems have the potential to form an alert system for volcanic unrest in remote and inaccessible regions.

Acknowledgments

This work was supported by the EPSRC Global Challenges Research Fund, the NERC BGS Centre for Observation and Modelling of Earthquakes, Volcanoes and Tectonics (COMET), the NERC large grant Looking into the Continents from Space (LiCS) - grant code NE/K010913/1, the EPSRC Platform Grant - Vision for the Future (EP/M000885/1), and a seed grant from the University of Bristol Cabot Institute. The InSAR datasets are available at `https://volcanodeformation.blogs.ilrt.org/` and `http://comet.nerc.ac.uk/COMET-LiCS-portal/` and the training dataset is available at `https://seis.bristol.ac.uk/~eexna/download.html`. We would like to thank the Advanced Computing Research Centre, University of Bristol, for free access the High Performance Computing machine (BlueCrystal) used for training machine learning algorithms. Also, we would like to thank Dr Karsten Spaans and Emma Hatton, School of Earth and Environment, University of Leeds, for Sentinel dataset.

References

Adeli, H., and A. Panakkat (2009), A probabilistic neural network for earthquake magnitude prediction, *Neural Networks*, *22*(7), 1018 – 1024, doi:https://doi.org/10.1016/j.neunet.2009.05.003.

Agostinelli, F., M. Hoffman, P. Sadowski, and P. Baldi (2015), Learning activation functions to improve deep neural networks, in *Proceedings of International Conference on Learning Representations*, pp. 1–9.

Anantrasirichai, N., A. Achim, J. E. Morgan, I. Erchova, and L. Nicholson (2013), Svm-based texture classification in optical coherence tomography, in *IEEE 10th International*

Symposium on Biomedical Imaging, pp. 1332–1335, doi:https://doi.org/10.1109/ISBI.2013.6556778.

Auker, M. R., R. S. J. Sparks, L. Siebert, H. S. Crosweller, and J. Ewert (2013), A statistical analysis of the global historical volcanic fatalities record, *Journal of Applied Volcanology*, *2*(1), 2.

Bagnardi, M. (2017), Intrusive activity at Cerro Azul volcano, Galápagos islands (Ecuador), *Available from: http://comet.nerc.ac.uk/intrusive-activity-cerro-azul-volcano-galapagos-islands-ecuador/*, Accessed 26 March 2018.

Bekaert, D., R. Walters, T. Wright, A. Hooper, and D. Parker (2015), Statistical comparison of insar tropospheric correction techniques, *Remote Sensing of Environment*, *170*, 40–47, doi:https://doi.org/10.1016/j.rse.2015.08.035.

Biggs, J., and M. E. Pritchard (2017), Global volcano monitoring: what does it mean when volcanoes deform?, *Elements*, *13*(1), 17–22, doi:https://doi.org/10.2113/gselements.13.1.17.

Biggs, J., T. Wright, Z. Lu, and B. Parsons (2007), Multi-interferogram method for measuring interseismic deformation: Denali Fault, Alaska, *Geophysical Journal International*, *170*(3), 1165–1179, doi:https://doi.org/10.1111/j.1365-246X.2007.03415.x.

Biggs, J., E. Anthony, and C. Ebinger (2009), Multiple inflation and deflation events at kenyan volcanoes, east african rift, *Geology*, *37*(11), 979–982.

Biggs, J., I. Bastow, D. Keir, and E. Lewi (2011), Pulses of deformation reveal frequently recurring shallow magmatic activity beneath the Main Ethiopian Rift, *Geochemistry, Geophysics, Geosystems*, *12*, 1–11, doi:https://doi.org/10.1029/2011GC003662.

Biggs, J., S. K. Ebmeier, W. P. Aspinall, Z. Lu, M. E. Pritchard, R. S. J. Sparks, and T. A. Mather (2014), Global link between deformation and volcanic eruption quantified by satellite imagery, *Nature Communications*, *5*, 3471–3471, doi:https://doi.org/10.1038/ncomms4471.

Brown, M., D. Lary, A. Vrieling, D. Stathakis, and H. Mussa (2008), Neural networks as a tool for constructing continuous ndvi time series from AVHRR and MODIS, *International Journal of Remote Sensing*, *29*(24), 7141âĂŞ7158, doi:https://doi.org/10.1080/01431160802238435.

Burges, C. J. (1998), A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery*, *2*(2), 121–167, doi:https://doi.org/10.1023/A:1009715923555.

Canny, J. (1986), A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.*, *8*(6), 679 – 698, doi:https://doi.org/10.1109/TPAMI.1986.4767851.

Chang, C.-C., and C.-J. Lin (2011), LIBSVM: A library for support vector machines, *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 27:1–27:27, doi:https://doi.org/10.1145/1961189.1961199.

Chaussard, E., and F. Amelung (2012), Precursory inflation of shallow magma reservoirs at west sunda volcanoes detected by insar, *Geophysical Research Letters*, *39*(21).

Chen, C. W., and H. A. Zebker (2001), Two-dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization, *JOSA A*, *18*(2), 338–351, doi:https://doi.org/10.1364/JOSAA.18.000338.

Chen, X. W., and X. Lin (2014), Big data deep learning: Challenges and perspectives, *IEEE Access*, *2*, 514–525, doi:https://doi.org/10.1109/ACCESS.2014.2325029.

Christianini, N., and J. Shawe-Taylor (2000), *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge UK: Cambridge University Press.

Conforti, M., S. Pascale, G. Robustelli, and F. Sdao (2014), Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the turbolo river catchment (northern calabria, italy), *CATENA*, *113*, 236 – 250, doi:https://doi.org/10.1016/j.catena.2013.08.006.

Ebmeier, S., J. Biggs, T. Mather, and F. Amelung (2013a), On the lack of insar observations of magmatic deformation at central american volcanoes, *Journal of Geophysical Research: Solid Earth*, *118*(5), 2571–2585.

Ebmeier, S., B. Andrews, M. Araya, D. Arnold, J. Biggs, C. Cooper, E. Cottrell, M. Furtney, J. Hickey, J. Jay, et al. (2018), Synthesis of global satellite observations of magmatic and volcanic deformation: implications for volcano monitoring & the lateral extent of magmatic domains, *Journal of Applied Volcanology*, *7*(1), 2, doi:https://doi.org/10.1186/s13617-018-0071-3.

Ebmeier, S. K., J. Biggs, T. A. Mather, and F. Amelung (2013b), Applicability of InSAR to tropical volcanoes: insights from Central America, *Geological Society, London, Special Publications*, *380*, 15–37, doi:https://doi.org/10.1144/SP380.2.

Erhan, D., Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio (2010), Why does unsupervised pre-training help deep learning?, *Journal of Machine Learning Research*, *11*, 625–660.

Fernández, J., A. Pepe, M. P. Poland, and F. Sigmundsson (2017), Volcano Geodesy: Recent developments and future challenges, *Journal of Volcanology and Geothermal Research*, *344*, 1 – 12, doi:https://doi.org/10.1016/j.jvolgeores.2017.08.006.

Gaddam, S. R., V. V. Phoha, and K. S. Balagani (2007), K-Means+ID3: A novel method for supervised anomaly detection by cascading k-means clustering and ID3 decision tree learning methods, *IEEE Transactions on Knowledge and Data Engineering*, *19*(3), 345–354, doi:https://doi.org/10.1109/TKDE.2007.44.

Global Volcanism Program (2013), *Volcanoes of the World, v. 4.7.0*, Smithsonian Institution, doi:https://doi.org/10.5479/si.GVP.VOTW4-2013.

Goldstein, R. M., H. A. Zebker, and C. L. Werner (1988), Satellite radar interferometry: Two-dimensional phase unwrapping, *Radio science*, *23*(4), 713–720, doi:https://doi.org/10.1029/RS023i004p00713.

Goodfellow, I., Y. Bengio, and A. Courville (2016), *Deep Learning*, MIT Press, http://www.deeplearningbook.org.

He, K., X. Zhang, S. Ren, and J. Sun (2016), Deep residual learning for image recognition, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, doi:https://doi.org/10.1109/CVPR.2016.90.

Jia, Y., E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell (2014), Caffe: Convolutional architecture for fast feature embedding, in *Proceedings of the 22Nd ACM International Conference on Multimedia*, pp. 675–678.

Jolivet, R., P. S. Agram, N. Y. Lin, M. Simons, M.-P. Doin, G. Peltzer, and Z. Li (2014), Improving insar geodesy using global atmospheric models, *Journal of Geophysical Research: Solid Earth*, *119*(3), 2324–2341.

Kotsiantis, S. B. (2007), Supervised machine learning: A review of classification techniques, in *Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, pp. 3–24, IOS Press.

Krizhevsky, A., I. Sutskever, and G. E. Hinton (2012), Imagenet classification with deep convolutional neural networks, in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, pp. 1097–1105, Curran Associates Inc., USA.

Lary, D. J., A. H. Alavi, A. H. Gandomi, and A. L. Walker (2016), Machine learning in geosciences and remote sensing, *Geoscience Frontiers*, *7*(1), 3 – 10, doi:https://doi.org/10.1016/j.gsf.2015.07.003, special Issue: Progress of Machine Learning in Geosciences.

LeCun, Y., Y. Bengio, and G. Hinton (2015), Deep learning, *Nature*, *521*, 436–444, doi:https://doi.org/10.1038/nature14539.

Li, C., J. Wang, L. Wang, L. Hu, and P. Gong (2014), Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery, *Remote Sensing*, *6*(2), 964–983, doi:https://doi.org/10.3390/rs6020964.

Li, Z., J.-P. Muller, P. Cross, and E. J. Fielding (2005), Interferometric synthetic aperture radar (insar) atmospheric correction: Gps, moderate resolution imaging spectroradiometer (modis), and insar integration, *Journal of Geophysical Research: Solid Earth*, *110*(B3).

Lloyd, R., J. Biggs, M. Wilks, A. Nowacki, J. M. Kendall, A. Ayele, E. Lewi, and H. Eysteinsson (2018), Evidence for cross rift structural controls on deformation and seismicity at a continental rift caldera, *Earth and Planetary Science Letters*, *487*, 190–200, doi:https://doi.org/10.1016/j.epsl.2018.01.037.

Loughlin, S. C., R. S. J. Sparks, S. K. Brown, S. F. Jenkins, and C. Vye-Brown (2015), *Global volcanic hazards and risk*, Cambridge University Press.

Matsugu, M., K. Mori, Y. Mitari, and Y. Kaneda (2003), Subject independent facial expression recognition with robust face detection using a convolutional neural network, *Neural Networks*, *16*(5-6), 555–559, doi:https://doi.org/10.1016/S0893-6080(03)00115-1.

Mei, E. T. W., F. Lavigne, A. Picquout, E. De Bélizal, D. Brunstein, D. Grancher, J. Sartohadi, N. Cholik, and C. Vidal (2013), Lessons learned from the 2010 evacuations at merapi volcano, *Journal of volcanology and geothermal research*, *261*, 348–365.

Parker, A. L., J. Biggs, R. J.Walters, S. K.Ebmeier, T. J.Wright, N. A.Teanby, and Z. Lu (2015a), Systematic assessment of atmospheric uncertainties for insar data at volcanic arcs using large-scale atmospheric models: Application to the Cascade volcanoes, United States, *Remote Sensing of Environment*, *170*, 102–114, doi:https://doi.org/10.1016/j.rse.2015.09.003.

Parker, A. L., J. Biggs, R. J. Walters, S. K. Ebmeier, T. J. Wright, N. A. Teanby, and Z. Lu (2015b), Systematic assessment of atmospheric uncertainties for insar data at volcanic arcs using large-scale atmospheric models: Application to the cascade volcanoes, united states, *Remote Sensing of Environment*, *170*, 102 – 114, doi:https://doi.org/10.1016/j.rse.2015.09.003.

Pinel, V., M. P. Poland, and A. Hooper (2014), Volcanology: lessons learned from synthetic aperture radar imagery, *Journal of Volcanology and Geothermal Research*, *289*, 81–113.

Pritchard, M., and M. Simons (2004a), An insarâĂŘbased survey of volcanic deformation in the central andes, *Geochemistry, Geophysics, Geosystems*, *5*(2), 42, doi:https://doi.org/10.1029/2003GC000610.

Pritchard, M. E., and M. Simons (2004b), An insar-based survey of volcanic deformation in the southern andes, *Geophysical research letters*, *31*(15).

Pritchard, M. E., and M. Simons (2004c), Surveying volcanic arcs with satellite radar interferometry: The central andes, kamchatka, and beyond, *GSA Today*, *14*(8), 4–11.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986), Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1, chap. Learning Internal Representations by Error Propagation, pp. 318–362, MIT Press, Cambridge, MA, USA.

Samuel, A. L. (1959), Some studies in machine learning using the game of checkers, *IBM Journal of Research and Development*, *3*(3), 210–229, doi:https://doi.org/10.1147/rd.33.0210.

Simonyan, K., and A. Zisserman (2014), Very deep convolutional networks for large-scale image recognition, *arXiv:1409.1556*, pp. 1–11.

Singer, B. S., N. L. Andersen, H. L. Mével1, K. L. Feigl, C. DeMets, B. Tikoff, C. H. Thurber, B. R. Jicha, C. Cardona, L. Córdova, F. Gil, M. J. Unsworth, G. Williams-Jones, C. Miller, J. Fierstein, W. Hildreth, and J. Vazquez (2014), Dynamics of a large, restless, rhyolitic magma system at Laguna del Maule, southern Andes, Chile, *GSA TODAY*, *24*(12), 4–10, doi:https://doi.org/10.1130/GSATG216A.1.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov (2014), Dropout: A simple way to prevent neural networks from overfitting, *Journal of Machine Learning Research*, *15*, 1929–1958.

Szegedy, C., V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna (2016), Rethinking the inception architecture for computer vision, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, doi:https://doi.org/10.1109/CVPR.2016.308.

Tien Bui, D., T. A. Tuan, N.-D. Hoang, N. Q. Thanh, D. B. Nguyen, N. Van Liem, and B. Pradhan (2017), Spatial prediction of rainfall-induced landslides for the lao cai area (vietnam) using a hybrid intelligent approach of least squares support vector machines inference model and artificial bee colony optimization, *Landslides*, *14*(2), 447–458, doi:https://doi.org/10.1007/s10346-016-0711-9.

Totaro, C., B. Orecchio, D. Presti, S. Scolaro, and G. Neri (2016), Seismogenic stress field estimation in the calabrian arc region (south italy) from a bayesian approach, *Geophysical Research Letters*, *43*(17), 8960–8969, doi:https://doi.org/10.1002/2016GL070107.

Wei, Z., G. Si, C. Bin, and C. Kai (2013), The application of decision tree to intensity change classification of tropical cyclones in western north pacific, *Geophysical Research Letters*, *40*(9), 1883–1887, doi:https://doi.org/10.1002/grl.50280.

Xu, W., E. Rivalta, and X. Li (2017), Magmatic architecture within a rift segment: Articulate axial magma storage at Erta Ale volcano, Ethiopia, *Earth and Planetary Science Letters*, *476*, 79–86, doi:http://doi.org/10.1016/j.epsl.2017.07.051.

Yilmaz, I. (2010), Comparison of landslide susceptibility mapping methodologies for koyulhisar, turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine, *Environmental Earth Sciences*, *61*(4), 821–836, doi:https://doi.org/10.1007/s12665-009-0394-9.

Yu, C., Z. Li, and N. T. Penna (2018), Interferometric synthetic aperture radar atmospheric correction using a gps-based iterative tropospheric decomposition model, *Remote Sensing of Environment*, *204*, 109–121, doi:https://doi.org/10.1016/j.rse.2017.10.038.

Zanero, S., and S. M. Savaresi (2004), Unsupervised learning techniques for an intrusion detection system, in *Proceedings of the ACM Symposium on Applied Computing*, pp. 412–419, ACM, doi:https://doi.org/10.1145/967900.967988.