Hotspotter: A Generalizable Pipeline for Automated Detection of Subtle Volcanic Thermal Features in Satellite Images

Aditya Mohan[1], Andie Gomez-Patron[2], Matthew Pritchard[2], Hannah Kerner[1]

[1]School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85283, USA, {amohan62,hkerner}@asu.edu

[2]Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14850, USA

Abstract

Geologists seek to understand the relationship between volcanic unrest and eruptions by identifying subtle Volcanic Thermal Features (VTFs) in high-resolution satellite imagery. This analysis requires the careful curation of large databases of relevant volcanic thermal information. However, volcanic unrest is characterized by highly subtle thermal anomalies. Manual identification on a global scale is highly labor- and time-intensive. We propose Hotspotter: an end-to-end system to automatically detect subtle volcanic thermal anomalies in satellite images and derive relevant thermal statistics. Previous solutions for automated VTF detection have limited data size and geographic diversity. To accommodate an unprecedentedly large and diverse volcanic dataset, we propose an automated pipeline combining unsupervised anomaly detection with supervised classification to filter anomalous regions. Hotspotter gives 90% anomaly detection accuracy and robust generalization to new volcanoes. Our automated approach can accelerate scientists' search for VTFs to help identify relevant thermal precursors and enable more precise forecasts of global volcanic eruptions.

Introduction

Volcanic eruptions are hazardous phenomena that negatively impact human health and livelihoods. Over 800 million people live within 100 km of an active volcano and are at risk of a potential eruption (Brown, Auker, and Sparks 2015). Predicting volcanic eruptions is a topic of long-standing interest among volcanologists. Eruptions are usually preceded by subtle Volcanic Thermal Features (VTFs) characterized by abrupt background temperature fluctuations (Barberi et al. 1984), as is depicted in Figure 1. These pre-eruption unrests are usually identified by a combination of ground sensors (Winson et al. 2014). However, in regions where ground sensors are not viable, unrest may be identified in high-resolution Thermal InfraRed (TIR) data acquired by satellite sensors. For example, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) sensor is widely used to collect TIR data owing to its accessibility and high TIR spatial resolution (NASA 2023).

In an attempt to develop a new volcanic database that is populated with subtle VTFs, Reath et al. (2019) curated the ASTER Volcanic Thermal Output Database (AVTOD) by manually analyzing volcanoes that were active in the Holocene age (11,700 years ago to present day). In this manual process, investigators identified volcanic regions of interest with geological characteristics indicative of a potential volcano. These characterizations included regions lacking non-volcanic thermal features and those with a composition similar to the material surrounding known volcanic thermal features. After identifying such a region, the analysis involved probing for pixels within that region that were at least 2 Kelvin warmer than the average temperature of a surrounding 10x10 pixel area. These anomalous pixels were classified as thermal unrest and cataloged with related thermal data.

Gomez-Patron et al. (2023) built on the AVTOD method and created a time series of each VTF on a volcano for over 200 volcanoes. Carrying out this process for multi-year high-resolution satellite imagery is a tedious process that requires substantial time and effort. This calls for a simplified automated approach, that serves two purposes: to assist volcanologists in compiling an ASTER VTF time series database and to provide a means of verifying manually collected data for analysts.

To our knowledge, there is no existing solution for thermal anomaly detection using a global volcanic database. However, there have been smaller-scale initiatives focused on detecting subtle volcanic thermal anomalies. Existing solutions face several challenges, including: (i) inadequate data size and diversity, (ii) opaque selection processes for appropriate algorithms, (iii) reliance on smaller scenes that may overlook volcanic unrest, and (iv) limited output detail, often restricted to binary pixel-wise information.

Instead of a fully supervised pipeline which would require large amounts of labeled VTF examples, we approached the problem using unsupervised anomaly detection techniques. We used Domain-Agnostic Outlier Ranking Algorithms (DORA) (Kerner et al. 2022) to compare different anomaly detection algorithms. We found that the Local Reed-Xiaoli algorithm (LRX) (Molero et al. 2013a) most accurately identified VTFs in our database. LRX generates continuous pixel-wise anomaly scores instead of simple binary masks. This helps analysts pinpoint anomalous regions around volcanic summits more precisely than less granular methods like image-level classification or binary masks.

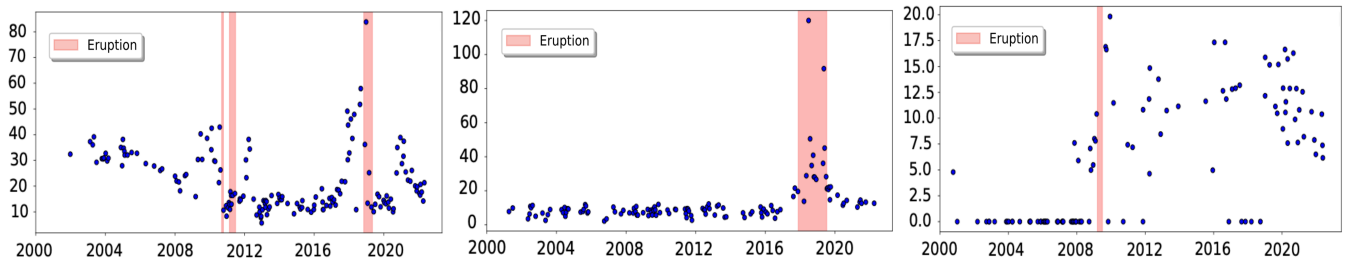


Figure 1: Examples of volcanoes, from left to right, in (i) Chile, (ii) Indonesia, and the (iii) United States with diverse temperature patterns around eruptions. Here, the x-axis represents the year, and the y-axis represents temperature in degrees Kelvin (K).

LRX outputs anomaly scores for each pixel but does not provide a hard classification of whether an image contains an anomaly or not. Images could be flagged as anomalous using a threshold on the anomaly scores in the image, but determining a single global threshold is difficult due to the diverse temperature ranges and varying degrees of anomalies across different volcanic zones (Corradino et al. 2022). Some anomalies detected by LRX may also be "false positives" that are anomalies in the data but not the VTF anomalies desired by analysts (e.g., clouds). To address this, we implemented a custom Convolutional Neural Network (CNN) (LeCun, Bengio, and Hinton 2015) to reduce false positives. The CNN processes the anomaly score image output by LRX and classifies whether the image contains a true anomaly or a false positive. The output from this two-stage pipeline (LRX followed by CNN) answers the question: "Does this ASTER image contain a volcanic thermal anomaly?"

Volcanologists require additional thermal information about anomalies beyond a coarse binary classification of an anomaly's presence (Reath et al. 2019). For images flagged as anomalous by our pipeline, we combined the LRX anomaly image output with the original ASTER image to extract key thermal statistics about the detected VTFs including maximum temperature, maximum temperature above background, mean background temperature, and standard deviation. This mirrors the metrics recorded by volcanologists in our reference expert-curated VTF database.

Related Work

Recent advancements in automated monitoring systems of volcanic unrest can be characterized by their data acquisition methods, size and types of data used, and classification approaches.

The MOUNTS system uses a CNN to detect ground deformations in Sentinel-1 Synthetic Aperture Radar (SAR) interferograms (Valade et al. 2019). Compared to prior methods, MOUNTS achieved higher accuracy and true positive and true negative rates (Anantrasirichai et al. 2018). However, the global applicability of this system is limited by the difficulty of obtaining SAR interferograms and non-uniform processing of input data globally.

Corradino et al. (2022) trained a model on optical Sentinel-2 features to predict binary masks for anomalous

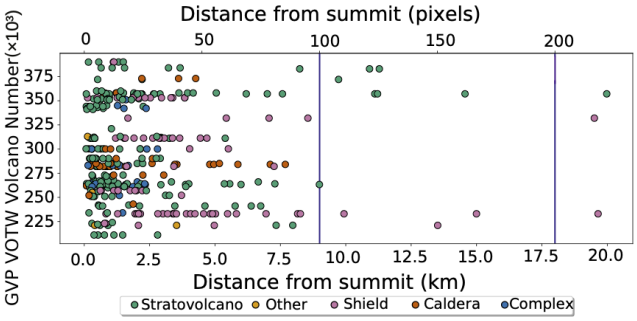


Figure 2: Plot of distance from VTF centroids recorded in our database to the volcanic summit. 96.3% of VTFs are within 100 pixels of the summit.

regions. The study experimented with variations of the random forest algorithm, claiming ensembles of decision trees typically yield highly generalizable performance. However, decision trees are limited in their ability to exploit spatial correlations in images, as they assume pixel-wise features are independent of each other. Moreover, the model was trained using data from only 5 volcanoes, which restricts its ability to capture complex topographic and structural differences across global volcanic zones. Massimetti et al. (2020) analyzed Sentinel-2 multispectral data using a dataset from a larger, but still limited, dataset of 8 volcanoes.

Corradino et al. (2023) used a U-Net segmentation model to detect VTFs in ASTER thermal images. However, the proposed method has significant limitations for global generalizability and deployment due to a lack of diverse data and a small sample size. Corradino et al. (2023) trained a U-Net on 1500 48×48-pixel images from 5 VTF occurrences. The model predicted a binary mask, trained using labels manually annotated by experts. Given the small data size and limited diversity, a complex architecture like U-Net may overfit this small dataset. In addition, the patch size of 48×48 pixels (4×4 km) around volcanic summits is not sufficient to identify unrest indicators. Figure 2 shows that many thermal features are more than 4 km away from the volcano summit.

Previous methods for volcanic thermal feature detection have been limited to a small set of volcanoes and had relatively small sample sizes, resulting in models that are unlikely to generalize to the diverse temperature patterns in dif-

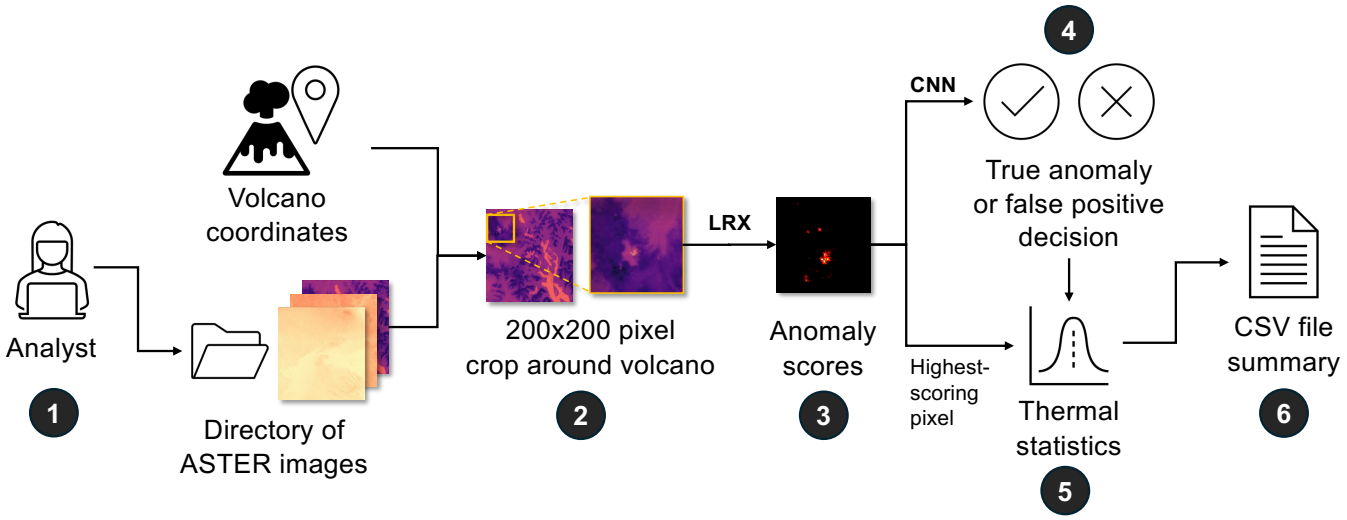


Figure 3: Hotspotter pipeline for automated detection of volcanic thermal features (VTFs). Step 1: Analyst specifies a path to a directory containing ASTER thermal images. Step 2: 200 × 200-pixel patch is cropped from each ASTER image around the matching volcano coordinates. Step 3: LRX computes pixel-wise anomaly scores. Step 4: CNN classifies whether the anomaly score image contains a true VTF or a false positive. Step 5: Thermal statistics are computed using the pixel with the highest anomaly score, for images classified as true anomalies. Step 6: CSV file of VTF detections and thermal statistics is saved.

ferent volcanic zones globally. Figure 1 illustrates how volcanic regions worldwide exhibit distinct patterns and temperatures in the times surrounding eruptions. This highlights the need for a comprehensive global volcanic dataset that captures the full spectrum of temperature profiles. Considering the limitations identified in recent research within this field, our objective was to reduce reliance on labeled data for models detecting VTFs and prioritize high generalizability across diverse volcanic zones.

Data

Thermal Satellite Data

We used thermal satellite images from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) sensor. ASTER offers the highest spatial resolution in the thermal infrared (TIR) range among openly accessible TIR satellite data. These images were manually filtered to discard images containing clouds within 100-200 km of the volcano summit. We used the atmospherically corrected AST_08 product (kinetic surface temperature), which has a resolution of 90 m/pixel, for our dataset (Reath et al. 2019). The ASTER images were mostly 700×830 pixels with some minor variations. These GeoTIFFs contain thermal observations as well as metadata such as acquisition date, which is used to populate our database. Our dataset included a time series of thermal images (GeoTIFFs) from 2000 to 2022 for 129 global volcanoes.

Volcanic Thermal Features

We used the database from Gomez-Patron et al. (2023) to validate our results and assign VTF labels to each ASTER image. This database contains thermal anomaly and other relevant thermal statistics for each volcano as well as the date the measurement was recorded. This database extended the previous AVTOD methodology (Reath et al. 2019) for analyzing volcanoes by identifying individual features on each volcano.

Approach

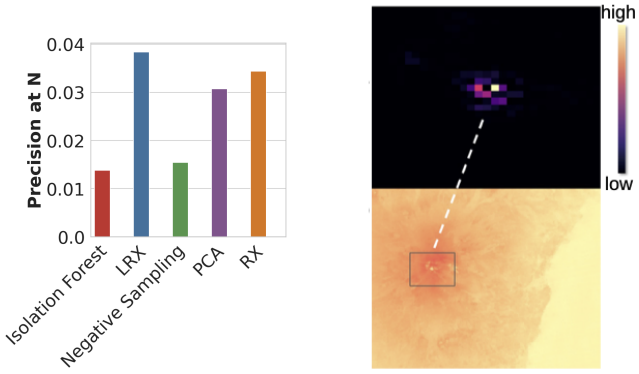
Figure 3 illustrates the pipeline of our proposed approach. An analyst starts by pointing the pipeline to a folder containing ASTER images. Hotspotter determines whether (and where) an image contains a VTF then calculates thermal statistics surrounding the VTF. These results and metadata are then stored in a CSV file.

Preprocessing

The VTFs in ASTER images can be as small as a single pixel (90×90 m). To reduce the search space and possible false positive detections far away from the volcano while preserving important unrest indicators, we cropped each image to a limited region around the volcano summit. For each volcano, we computed the distance from each VTF to the summit using coordinates from the GVP Volcanoes of the World database (Global Volcanism Program and Venzke 2024). We found that 96.3% of the VTFs were captured within 100 pixels (9 km) of the summit, with only a few outliers beyond this range. Thus we determined that a 200×200 pixel crop would provide a compromise between feature inclusivity and image processing efficiency.

Unsupervised Anomaly Detection using LRX

After extracting the 200×200 pixel patch around the volcanic summit, we apply an unsupervised algorithm to



(a) Performance for anomaly detection algorithms in DORA (Kerner et al. 2022), measured by Precision at N.

(b) LRX scores (top) for VTF in example volcano (Etna) and the corresponding thermal ASTER image (bottom).

Figure 4: Comparison of anomaly detection algorithms evaluated using DORA (a) and example output from Local RX (LRX), the best performing algorithm (b).

identify anomalous pixels in the image. To select an appropriate algorithm for this specific task, we used the Domain-agnostic Outlier Ranking Algorithm (DORA) library (Kerner et al. 2022). DORA (Kerner et al. 2022) is a versatile API that evaluates multiple outlier detection algorithms in a variety of domains. DORA contains diverse anomaly scoring methods that employ a variety of factors to score anomalies: Reconstruction Error (Principal Component Analysis), Distance (Reed-Xiaoli, Local Reed-Xiaoli detectors), Sparsity (Isolation Forest), and Likelihood (Negative Sampling). An important metric provided by DORA is "Precision at N", which calculates the fraction of N selections (samples with highest anomaly scores) made by the algorithm that are actually known outliers (determined using our label database) (Campos et al. 2016).

We applied DORA to 10 randomly chosen volcanoes and evaluated each algorithm using Precision at N, setting N to the number of known anomalous pixels in each image. We found that the Local Reed-Xiaoli (LRX) detector had the highest Precision at N (Figure 4a). Qualitatively, we found that LRX localized VTF outliers well (Figure 4b). Other algorithms produced poor precision scores or failed to identify clear outliers in visualizations.

LRX computes local means and covariances using pixels between inner and outer windows surrounding the pixel being scored. LRX scores each pixel using the Mahalanobis distance between the pixel and outer window statistics, using the inner window to mask the immediate area around the pixel being scored. We found that a 3×3 inner window and 5×5 outer window gave good results since VTFs tend to be very small. Prior work also used LRX to detect small anomalies in hyperspectral images (Molero et al. 2013b).

LRX narrowed down the search space for anomalies to a few pixels. However, LRX uses only the local context around a pixel to score anomalies and does not consider the larger context within an image. Some anomalies detected by LRX may also be "false positives" that are anomalies in the data but not the VTF anomalies desired by analysts (e.g., clouds). We used a CNN to classify whether an LRX detection is a true VTF anomaly or a false positive, as described in the next section.

Supervised Anomaly Classification using CNN

To remove false positives and incorporate global context, we implemented a Convolutional Neural Network (CNN) that processes the anomaly image output from LRX. CNNs capture complex relationships within an entire image through successive convolutions. We experimented with various configurations of CNNs with varying depth, width, and activation functions using cross-validation. We found that a shallow 3-layer CNN performed best, eliminating the need for deeper networks prone to overfitting and increased computation.

We trained the model using the binary cross-entropy loss function, traditionally used for binary classification problems. We applied batch normalization after each convolutional layer. Batch normalization aids anomaly detection by clustering normal samples and pushing anomalies away from the distribution (Li et al. 2023). It also helps in regularizing the model, which can reduce the likelihood of overfitting to irrelevant features or noise, a common issue in anomaly detection tasks where the anomalies can be very subtle and sparse. The max-pooling layers reduce the spatial dimensions, which not only helps in reducing the computational load but also ensures that the network focuses on the most salient features, increasing its ability to detect anomalies.

We balanced the dataset to ensure an equal number of positive and negative labels. Initially, the dataset contained a majority of positive samples due to analysts' tendency to investigate regions prone to thermal anomalies (which are correctly identified and considered true positive). To deal with this imbalance, we augmented the negative class by cropping non-VTF regions from ASTER images. Any thermal variations in these regions would result from other non-VTF phenomena.

Deriving Thermal Statistics

Once the CNN eliminates false positives and determines that the image is anomalous, we extract thermal statistics from the highest-scoring pixel in the corresponding LRX image. We compute the following thermal features: (i) maximum anomalous temperature, (ii) maximum temperature above the background, (iii) mean background temperature, and (iv) standard deviation.

DORA records the highest-scoring LRX pixel along with the image coordinates of that pixel. We then mapped this coordinate back to the original GeoTIFF image to determine the maximum temperature. To capture the background thermal statistics, we analyzed a 10×10 pixel patch surrounding the highest-scoring LRX pixel, following AVTOD's methodology. The maximum temperature above background was calculated by subtracting the mean of this 10×10 patch from the highest-scoring LRX pixel value.

Model	Train-Test Split	Acc.	F1
LRX+CNN (ours)	Random	**0.903**	**0.901**
CNN	Random	0.831	0.882
Fixed Threshold	Random	0.631	0.770
LRX+CNN (ours)	Non-overlapping	**0.889**	**0.884**
CNN	Non-overlapping	0.637	0.752

Table 1: Performance of our approach using unsupervised LRX for detecting candidate anomalies followed by a supervised CNN for identifying volcanic anomalies. We compared this approach to using only the supervised CNN without LRX or using a fixed threshold on pixel temperatures.

Results

Our performance evaluation was guided by three key questions: (1) How well did our pipeline detect VTF anomalies? (2) Did the inferred thermal statistics match those computed in the manual database? (3) Did our model generalize well to unseen volcanoes?

Anomaly Detection

Table 1 summarizes the VTF detection experiment results. Our labeled dataset comprises 7,968 volcanic data points from 129 globally distributed volcanoes. We implemented an 80/10/10% random sample split for training, testing, and validation, with balanced positive and negative samples. We trained the CNN for 50 epochs on the LRX score images. Our LRX+CNN method achieved 90.3% test accuracy and 88.4% F1 score.

Ablation experiment To assess the necessity of LRX as an intermediate step, we conducted an ablation experiment in which we trained the CNN directly on thermal images rather than the LRX output. This is comparable to previous methods that used a CNN approach (such as MOUNTS (Valade et al. 2019)), however, we could not compare directly to previous methods due to differences in datasets and satellite data sources. The CNN-only method gave 81% test accuracy, approximately 7% lower than training on LRX scores.

Spatial generalization We performed an experiment to evaluate the model's performance across diverse topographic and thermal conditions that were not seen during training. We created a new train/test split ensuring each volcano was part of only one split. Despite being a more challenging task to generalize to completely new volcanoes, our LRX+CNN approach still achieved an impressive 88.9% accuracy, significantly higher than the CNN-only accuracy of 63.7%.

Baseline We implemented a simple thresholding method as a baseline for our proposed model's performance. This mimics the manual process analysts currently use to find candidate VTFs. We classified a pixel as anomalous if its temperature exceeds the mean background temperature by 2 Kelvin. This technique achieved a much lower accuracy of 63%. This shows that our LRX+CNN pipeline performs significantly better than the thresholding approach used by analysts today. Moreover, LRX scores provide deeper insights that can help us derive background thermal statistics and identify subtle precursors.

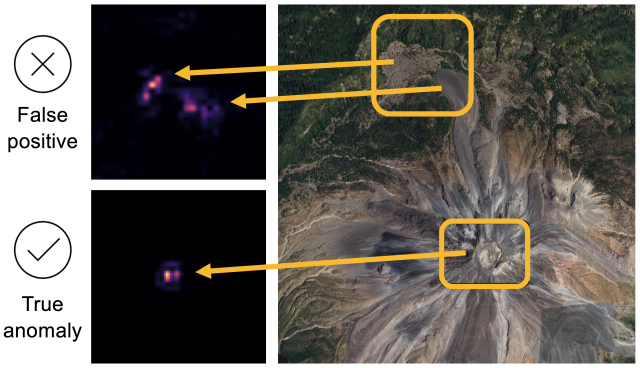


Figure 5: LRX anomaly scores for a false positive anomaly (top left) and true VTF anomaly (bottom left) detected at the Colima volcano. LRX identified both regions as anomalies based on their thermal values and the CNN correctly filtered the patches to record only the true anomaly in the database.

Colima: A qualitative case study To demonstrate our algorithm's ability to identify subtle anomalies in the presence of false positives, we analyzed two distinct snapshots from Volcán de Colima, one of the most active volcanoes in Mexico and North America (Spica 2017). We examined ASTER TIR images from two specific dates: March 5, 2003, containing a manually identified volcanic anomaly, and March 9, 2022, labeled as non-anomalous. Figure 5 shows the LRX outputs for these two images.

The LRX output for the non-anomalous image assigned high anomaly scores to non-volcanic thermal features in the upper part of the image. These features likely correspond to an old volcanic lava flow. Even when inactive, lava can often exhibit elevated temperatures in thermal imagery due to residual heat retention. This is an example of a local thermal anomaly in the data that is not the specific VTF thermal anomaly that analysts seek to catalog. Despite this ambiguity, our CNN correctly classified both images according to their respective labels. This illustrates how the CNN effectively learned to distinguish true volcanic anomalies from other thermal features that might be misinterpreted as volcanic thermal hotspots by human observers or more naive algorithmic approaches.

Thermal statistics

To evaluate the accuracy of the thermal statistics derived from Hotspotter, we calculated the Mean Absolute Error (MAE) between the ground truth and inferred value for each statistic. We chose MAE since it has an interpretable unit of temperature in Kelvin (K), same as the data. For unseen volcanoes, our inferred statistics had an MAE of 3.23 K for the maximum temperature pixel, 4.27 K for the mean background temperature, 11.48 K for the maximum temperature above the background, and 5.16 K for the standard devia-

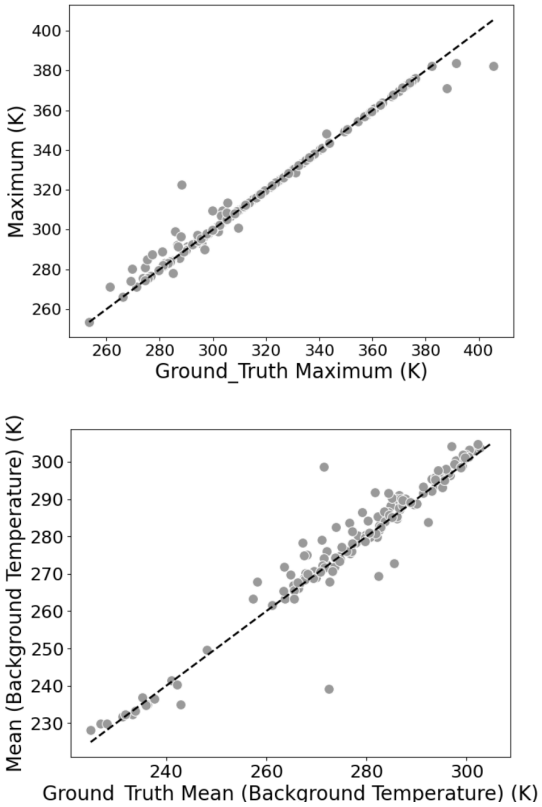


Figure 6: Scatter plots illustrating the relationship between ground truth and predicted values for maximum temperature (top) and mean background temperature (bottom). The black dashed diagonal line represents the ideal match where ground truth equals predicted values.

tion. Figure 6 shows a scatter plot of the ground-truth and inferred values for maximum temperature and mean background temperature. These results show a strong correlation between the ground-truth and Hotspotter's inferred values.

Sensitivity to train/test split The experiments in Table 1 were performed with a single train/test split and random seed. To evaluate the sensitivity of Hotspotter's performance to a specific train/test split and CNN initialization (controlled by the random seed), we created 24 random train/test splits with different random seeds and repeated the experiment from row 1 of Table 1. Figure 7 shows box-plots of the results for each metric (accuracy, F1 score, and MAE of thermal statistics). These results show slightly higher mean accuracy and F1 scores than Table 1. They also show that some performance variation could be expected depending on the random seed and resulting train/test split. In the deployed Hotspotter application, using all of the available data for training could help reduce such variability.

Pathway to Deployment

Hotspotter is publicly available on GitHub (https://github.com/kerner-lab/VTFDetection) and can be run via the command line. Given input and output directories, Hotspotter generates LRX anomaly scores for each image, VTF classifications for each image, and thermal statistics for detected VTFs. In the future, Hotspotter can become the primary method for populating VTF time series databases with unrest indicators and thermal statistics.

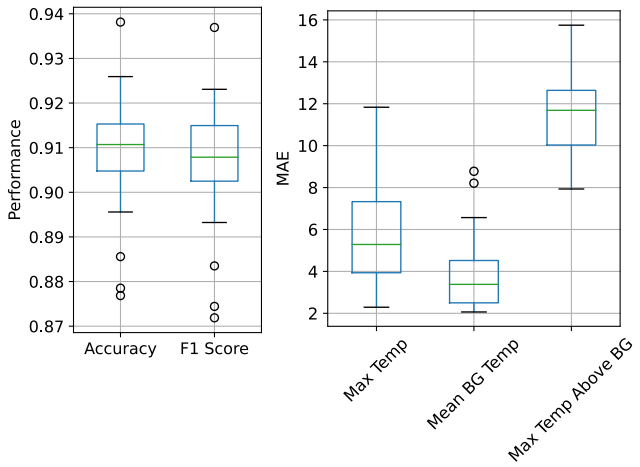


Figure 7: Hotspotter performance across 24 train/test splits with different random seeds.

Conclusion

Volcanic eruptions have taught us to be wary of important, yet subtle precursors that could easily escape human observation. To better understand thermal precursors and develop more precise forecasts of volcanic eruptions, scientists need global databases cataloging the occurrences and properties of volcanic thermal features. Cataloging time series of VTFs for every volcano in multi-year high-resolution satellite imagery is a tedious manual process that requires substantial time and effort today. An automated pipeline would significantly accelerate the scale and speed at which volcanologists can identify subtle thermal features.

Our proposed Hotspotter pipeline for automated VTF detection is the first to use a comprehensive global VTF time series database. Hotspotter accurately detects VTFs in ASTER thermal satellite images. Thermal statistics of the detected VTFs correlate strongly with ground-truth values from the reference VTF database. Our experiments show that our model generalizes well across diverse volcanic zones and new volcanoes not seen during training. In addition to VTF classifications and thermal statistics, Hotspotter's intermediate LRX pixel-wise scores enable additional exploration of anomalous features. Hotspotter will accelerate scientists' search for VTFs and help identify relevant thermal precursors that will enable more precise forecasts of global volcanic eruptions.

Acknowledgements

This research was supported by a grant from the NASA Earth Systems and Interior program, award #80NSSC21K0842.

References

Anantrasirichai, N.; Biggs, J.; Albino, F.; Hill, P.; and Bull, D. 2018. Application of Machine Learning to Classification of Volcanic Deformation in Routinely Generated InSAR Data. *Journal of Geophysical Research: Solid Earth*.

Barberi, F.; Corrado, G.; Innocenti, F.; and Luongo, G. 1984. Phlegraean Fields 1982–1984: Brief chronicle of a volcano emergency in a densely populated area. *Bulletin Volcanologique*, 47(2): 175–185.

Brown, S. K.; Auker, M.; and Sparks, R. 2015. Populations around Holocene volcanoes and development of a Population Exposure Index. *Cambridge University Press*, 223–232.

Campos, G. O.; Zimek, A.; Sander, J.; Campello, R. J. G. B.; Micenková, B.; Schubert, E.; Assent, I.; and Houle, M. E. 2016. On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study. *Data Mining and Knowledge Discovery*, 30(4): 891–927.

Corradino, C.; Amato, E.; Torrisi, F.; and Del Negro, C. 2022. Data-Driven Random Forest Models for Detecting Volcanic Hot Spots in Sentinel-2 MSI Images. *Remote Sensing*, 14(17): 4370.

Corradino, C.; Ramsey, M. S.; Pailot-Bonnetat, S.; Harris, A. J.; and Negro, C. D. 2023. Detection of Subtle Thermal Anomalies: Deep Learning Applied to the ASTER Global Volcano Dataset. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 1–15.

Global Volcanism Program; and Venzke, E. 2024. Volcanoes of the World (v. 5.2.2; 22 Aug 2024). Database.

Gomez-Patron, A.; Way, L.; Crothers, C.; Downes, A.; Zapata, L.; Duato, R.; Jackson, L.; Nicastro, N.; Pritchard, M. E.; and Reath, K. 2023. Evaluating multi-decade time series of volcanic thermal features for potential eruption precursors using satellite ASTER data. In *IAVCEI 2023 Scientific Assembly*. Rotorua, New Zealand. Abstract 6B-1119.

Kerner, H. R.; Rebbapragada, U.; Wagstaff, K. L.; Lu, S.; Dubayah, B.; Huff, E.; Lee, J.; Raman, V.; and Kulshrestha, S. 2022. Domain-Agnostic Outlier Ranking Algorithms—A Configurable Pipeline for Facilitating Outlier Detection in Scientific Datasets. *Frontiers in astronomy and space sciences*, 9.

LeCun, Y.; Bengio, Y.; and Hinton, G. 2015. Deep learning. *nature*, 521(7553): 436.

Li, A.; Qiu, C.; Kloft, M.; Smyth, P.; Rudolph, M.; and Mandt, S. 2023. Zero-Shot Anomaly Detection via Batch Normalization. *arXiv*.

Massimetti, F.; Coppola, D.; Laiolo, M.; Valade, S.; Cigolini, C.; and Ripepe, M. 2020. Volcanic Hot-Spot Detection Using SENTINEL-2: A Comparison with MODIS–MIROVA Thermal Data Series. *Remote Sensing*, 12(5): 820.

Molero, J. M.; Garzon, E. M.; Garcia, I.; and Plaza, A. 2013a. Analysis and Optimizations of Global and Local Versions of the RX Algorithm for Anomaly Detection in Hyperspectral Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2): 801–814.

Molero, J. M.; Garzon, E. M.; Garcia, I.; and Plaza, A. 2013b. Analysis and Optimizations of Global and Local Versions of the RX Algorithm for Anomaly Detection in Hyperspectral Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2): 801–814.

NASA. 2023. Terra Instruments: ASTER. https://terra.nasa.gov/about/terra-instruments/aster. Accessed: 2024-06-24.

Reath, K.; Pritchard, M.; Moruzzi, S.; Alcott, A.; Coppola, D.; and Pieri, D. 2019. The AVTOD (ASTER Volcanic Thermal Output Database) Latin America archive. *Journal of Volcanology and Geothermal Research*, 376: 62–74.

Spica, P. M. L. D., Zack. 2017. Anatomy of the Colima volcano magmatic system, Mexico. *Earth and Planetary Science Letters*, 459: 1–13.

Valade, S.; Ley, A.; Massimetti, F.; D'Hondt, O.; Laiolo, M.; Coppola, D.; Loibl, D.; Hellwich, O.; and Walter, T. R. 2019. Towards Global Volcano Monitoring Using Multisensor Sentinel Missions and Artificial Intelligence: The MOUNTS Monitoring System. *Remote Sensing*, 11(13): 1528.

Winson, A. E. G.; Costa, F.; Newhall, C. G.; and Woo, G. 2014. An analysis of the issuance of volcanic alert levels during volcanic crises. *Journal of Applied Volcanology*, 3(1).