JGR Solid Earth

RESEARCH ARTICLE
10.1029/2024JB029194

Detection and Characterization of Seismic and Acoustic Signals at Pavlof Volcano, Alaska, Using Deep Learning

Darren Tan[1], David Fee[1], Alex Witsil[2], Társilo Girona[1], Matthew Haney[3], Aaron Wech[3], Chris Waythomas[3], and Taryn Lopez[1]

[1]Geophysical Institute, Alaska Volcano Observatory, University of Alaska Fairbanks, Fairbanks, AK, USA, [2]Applied Research Associates, Raleigh, NC, USA, [3]U.S. Geological Survey, Volcano Science Center, Alaska Volcano Observatory, Anchorage, AK, USA

Key Points:

- We develop a pair of convolutional neural networks that detect and classify volcano seismic and acoustic signals
- We apply our models to Pavlof Volcano eruptions (2007, 2013, 2014, 2016, and 2021–2022) and derive volcano seismoacoustic timelines
- The seismoacoustic timelines reveal shifts in unrest regimes linked to explosions and effusive activity

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Tan,
ptan@alaska.edu

Citation:

Tan, D., Fee, D., Witsil, A., Girona, T., Haney, M., Wech, A., et al. (2024). Detection and characterization of seismic and acoustic signals at Pavlof Volcano, Alaska, using deep learning. *Journal of Geophysical Research: Solid Earth*, *129*, e2024JB029194. https://doi.org/10.1029/2024JB029194

Received 27 MAR 2024
Accepted 11 JUN 2024

Author Contributions:

Conceptualization: Darren Tan, David Fee
Data curation: Taryn Lopez
Formal analysis: Darren Tan, David Fee, Alex Witsil, Társilo Girona, Matthew Haney, Aaron Wech, Chris Waythomas, Taryn Lopez
Investigation: Darren Tan, David Fee
Methodology: Darren Tan, David Fee, Alex Witsil, Társilo Girona
Supervision: David Fee
Visualization: Darren Tan
Writing – original draft: Darren Tan, David Fee
Writing – review & editing: Darren Tan, David Fee, Alex Witsil, Társilo Girona,

Abstract Volcanic tremor is a semi-continuous seismic and/or acoustic signal that occurs at time scales ranging from seconds to years, with variable amplitudes and spectral features. Tremor sources have often been related to fluid movement and degassing processes, and are recognized as a potential geophysical precursor and co-eruptive geophysical signal. Eruption forecasting and monitoring efforts need a fast, robust method to automatically detect, characterize, and catalog volcanic tremor. Here we develop VOlcano Infrasound and Seismic Spectrogram Network (VOISS-Net), a pair of convolutional neural networks (one for seismic, one for acoustic) that can detect tremor in near real-time and classify it according to its spectral signature. Specifically, we construct an extensive data set of labeled seismic and low-frequency acoustic (infrasound) spectrograms from the 2021–2022 eruption of Pavlof Volcano, Alaska, and use it to train VOISS-Net to differentiate between different tremor types, explosions, earthquakes and noise. We use VOISS-Net to classify continuous data from past Pavlof Volcano eruptions (2007, 2013, 2014, 2016, and 2021–2022). VOISS-Net achieves an 81.2% and 90.0% accuracy on the seismic and infrasound test sets respectively, and successfully characterizes tremor sequences for each eruption. By comparing the derived seismoacoustic timelines of each eruption with the corresponding eruption chronologies compiled by the Alaska Volcano Observatory, our model identifies changes in tremor regimes that coincide with observed volcanic activity. VOISS-Net can aid tremor-related monitoring and research by making consistent tremor catalogs more accessible.

Plain Language Summary Volcanic tremor is a persistent vibration of the ground, atmosphere, or both that can occur before and during volcanic eruptions. Despite its importance in volcano monitoring and eruption forecasting, volcano observatories do not have a reliable way of automatically detecting and identifying tremor due to the variable intensities and frequencies at which it occurs. In order to accomplish this, we develop and test a pair of machine learning models that classify spectrograms (i.e., images representing a signal's frequency content over time) from seismic and low-frequency acoustic data. The models are trained on manually labeled images derived from the recent 2021–2022 eruption of Pavlof Volcano, Alaska, which demonstrated substantial signal diversity (e.g., different tremor types, earthquakes, explosions and noise). Our models achieve 81.2% and 90.0% accuracy on the seismic and low-frequency acoustic test sets respectively, and perform well when applied to data recorded from past Pavlof Volcano eruptions. In addition, transitions in tremor sequences identified from our analysis generally coincide with shifts in eruptive patterns from Pavlof Volcano. Our tools can help volcano observatories systematically monitor tremor, and advance tremor research by making catalogs of their occurrences more consistent and accessible.

1. Introduction

Volcanic tremor is a persistent but often intractable seismic and/or acoustic signal that is regularly observed at active volcanoes, often preceding or accompanying eruptive activity (Chouet & Matoza, 2013; Chouet et al., 1987; Konstantinou & Schlindwein, 2003; Koyanagi et al., 1987). It is distinguished from other volcanic signals primarily by its consistent or gliding spectral peaks and the highly variable time scales of its occurrence, which range from minutes to years. Volcano seismic tremor has been hypothesized to be generated by complex interactions within volcanic systems, which can constitute the elastic resonance of magma conduits (Neuberg et al., 2000) and fluid-filled cracks (Chouet, 1985, 1996), pressure oscillations driven by porous gas flow (Girona et al., 2019), magma wagging (Jellinek & Bercovici, 2011), hydrothermal boiling (Leet, 1988), non-linear fluid-structure interaction (Julian, 1994), and the superposition of frequently repeating stick-slip earthquakes

Matthew Haney, Aaron Wech, Chris Waythomas, Taryn Lopez

(Dmitrieva et al., 2013; Hotovec et al., 2013; Powell & Neuberg, 2003). A separate class of volcanic tremor, appropriately termed "eruption tremor" (Ichihara, 2016), has seismic and acoustic components and directly relates to the extended signal observed during sustained volcanic paroxysms (McNutt & Nishimura, 2008; Scandone & Malone, 1985). Its seismic component has been modeled as the reactionary force of ejected material (Haney et al., 2018; Prejean & Brodsky, 2011) or a combination of turbulence and particle impacts (Fee, Haney, et al., 2017; Gestrich et al., 2020; McNutt & Nishimura, 2008) among other models, while its acoustic component has been modeled using various source radiation patterns (Woulff & McGetchin, 1976) in numerous studies (e.g., Fee, Izbekov, et al., 2017; Yamada et al., 2017). Other sources of volcano acoustic tremor have demonstrated a range of spectral patterns, where their source mechanisms are dependent on the style of eruptive activity (Fee & Matoza, 2013). For example, harmonic infrasonic tremor generated by degassing bursts within Halema'uma'u, the primary crater at Kīlauea, Hawaii, has been modeled as the result of cavity oscillations and resonances (Fee et al., 2010), while the infrasonic tremor that coincided with effusive activity at Kilauea's Pu'u'ō'ō crater and its adjacent lava tube were attributed to a mixture of bubble cloud oscillations and interactions between gas jets and their solid boundaries (Matoza et al., 2010). Other proposed mechanisms inferred from observations at other volcanoes include gravity-driven bubble column dynamics (Ripepe et al., 2010), the continuous outbursting of gas bubbles within the upper conduit (Ripepe et al., 1996), or successive explosions at a volcanic conduit's fragmentation surface (Ichihara, 2016). Many of these processes are inferred based on the spectral characteristics of tremor and how they change over time, which requires a method for consistently cataloging tremor.

In addition to the visual inspection of spectrograms, several methods have been proposed to automatically detect volcano seismic tremor, such as envelope cross-correlation (Waythomas et al., 2014; Wech & Creager, 2008), pitch-based methods (Roman, 2017), tracking statistical change points in spectra (Picard, 1985; Reiss et al., 2023), backprojection using an assumed Rayleigh wave velocity (Haney, 2014), and amplitude source location assuming isotropic S-wave radiation (Kumagai et al., 2015). Tracking seismic network coherence (Seydoux et al., 2016; Soubestre et al., 2018) has also shown considerable recent success in identifying tremor at several volcanic settings (Journeau et al., 2022; Maher et al., 2023). Acoustic tremor, on the other hand, has been primarily detected and monitored using array processing techniques (Bishop et al., 2020; Matoza et al., 2010), cross-correlation functions on locally deployed acoustic sensors (Barrière et al., 2023), or statistical tests on successive spectral windows (Luo et al., 2023). Ripepe et al. (2007) also demonstrated the detection of acoustic tremor from Stromboli Volcano, Italy, using multi-channel semblance analyses (Ripepe & Marchetti, 2002), while Johnson and Palma (2015) used a combination of cross-correlation lag times and semblance analyses on array data to study infrasonic tremor from lahars at Volcán Villarrica, Chile. While these methods have all shown promise in detecting volcanic tremor, they are either tuned to detect specific types of tremor (e.g., harmonic or broadband) or are limited to well-monitored volcanoes with numerous seismic and/or acoustic sensors. Furthermore, many of these techniques have not been rigorously tested at numerous volcanic settings, or might be challenging to apply in near real-time for monitoring purposes.

Volcano observatories and scientists worldwide still primarily detect and analyze volcanic tremor using spectrograms due to its highly variable amplitudes, durations, and often emergent onsets. However, this is a labor intensive task which can be affected by human inconsistencies in identifying tremor. This issue is further compounded in the event of volcanic unrest or eruption where information flux is high, a situation also encountered in earthquake seismology when a seismic swarm or a large mainshock-aftershock sequence occurs (Mousavi & Beroza, 2023; Zali et al., 2021). Motivated by the need to automate the spectrogram scanning process, we propose a method that leverages analyst spectrogram-monitoring expertise to detect and characterize volcanic tremor by their spectral properties in both seismic and infrasound (low-frequency acoustic, <20 Hz) data. Specifically, we use Convolutional Neural Networks (CNNs) (LeCun et al., 1998) to classify successive overlapping spectrogram windows in order to interpret tremor timelines across various time scales. CNNs, which are commonly employed in computer vision (Khan et al., 2018), are also being applied to seismic data. For example, they are used to identify seismic phase arrivals (Lapins et al., 2021; Mousavi et al., 2020; Zhu & Beroza, 2019), classify volcano-seismic signals (Ferreira et al., 2023; Titos et al., 2018), and differentiate anthropogenic explosions from earthquakes (Kong et al., 2022; Linville et al., 2019). Although CNNs have been used on audible acoustics (Hershey et al., 2017), their application to infrasound, or specifically volcanic infrasound, has been limited. Prior work in seismoacoustics has also mostly focused on applying machine learning algorithms to classify cataloged events (Ferreira et al., 2023; Kong et al., 2022; Linville et al., 2019; Titos et al., 2018) rather than on continuous data. The application of our models to continuous data opens up the

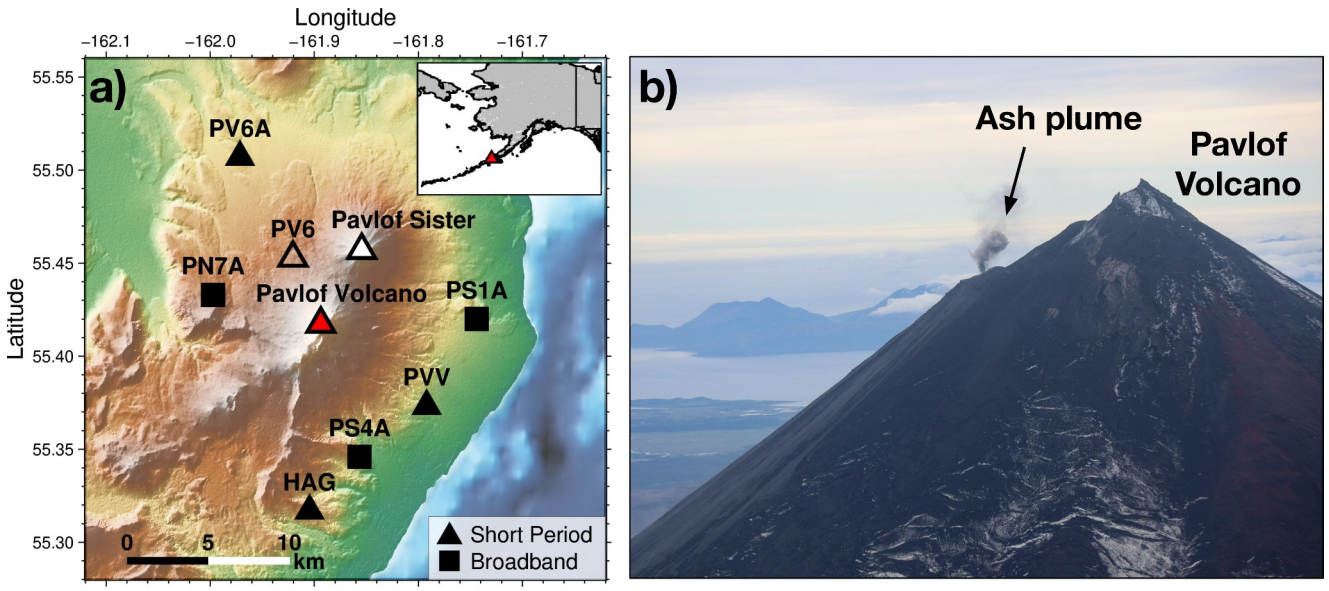

Figure 1. (a) Station map of the Alaska Volcano Observatory seismic and infrasound network at Pavlof Volcano. All seismic stations have a colocated infrasound sensor. Note that station PV6 was destroyed during the 2016 eruption, and is shown for reference because it was operational during the 2007, 2013, and 2014 eruptions. The inset map shows the location of Pavlof Volcano along the Alaska Peninsula. (b) Helicopter-based view of a small ash burst from Pavlof Volcano's active vent on 26 August 2022. The vents for the 2021–2022 eruption were located on the upper southeastern flank of the volcano relative to the summit. Photo by Valerie Wasser, University of Alaska Fairbanks.

possibility of detecting volcanic tremor in near real-time in addition to characterizing pre-existing data. As spectrograms are inherently images representing amplitude in frequency-time bins, a well-trained CNN should be able to discern patterns in frequency-time space and distinguish transient signals (earthquakes, explosions or both) from prolonged signals (tremor) with distinct spectral properties. While we acknowledge that other tremor classification works have found success in using unsupervised learning to establish volcano-seismic regimes (Duque et al., 2020; Langer et al., 2011; Steinke et al., 2023; Unglert et al., 2016; Watson, 2020; Zali et al., 2024), we choose to explore the supervised learning approach of CNNs, where a comprehensive and expert-identified set of predetermined classes are used for model training. Restricting the output classifications to a set of pre-determined classes avoids ambiguities in class/cluster interpretations and allows for quantitative assessments for its performance by using a labeled "ground truth" test set.

Although supervised machine learning techniques like CNNs benefit from unambiguous interpretations, the need to compile a robust, large, and consistently labeled data set remains a key challenge. Here, we leverage seismic and acoustic signals detected from the 2021–2022 eruption of Pavlof Volcano, Alaska—a 1.5 years long eruption that exhibited diverse tremor signals and numerous discrete explosions. Pavlof Volcano is located on the Alaska Peninsula, 55 km northeast of Cold Bay, Alaska (Figure 1a inset), and is one of the most historically active volcanoes in the Aleutian arc (www.avo.alaska.edu). Volcanic tremor is characteristic of seismic unrest at Pavlof Volcano, which contributes toward the difficulty of forecasting and detecting its eruptions (Cameron et al., 2018; Fee, Haney, et al., 2017; McNutt, 1987). Pavlof Volcano typically shows little to no detectable edifice deformation (Lu et al., 2014) or pre-eruptive seismicity (Pesicek et al., 2018). Although capable of producing sudden, subplinian style eruptions such as its preceding eruption in 2016 (Fee, Haney, et al., 2017; Gestrich et al., 2020), the 1.5 years long eruption that spanned 5 August 2021–17 December 2022 was comparatively more passive and was characterized by intermittent explosions, lava effusion, and ash emissions <3.6 km above sea level.

The volcanic phenomena observed during the 2021–2022 eruption of Pavlof Volcano were regularly documented by the Alaska Volcano Observatory (AVO) from satellite, web-camera, and pilot observations. The eruption also experienced minimal geophysical network interruptions. The eruption generated a variety of seismic and infrasonic tremor signals that underwent transitions in intensity and spectral character as the volcano progressed from background activity, to seismic unrest, and into its active eruption phases. As such, the 2021–2022 eruption presents a unique opportunity to compile labeled seismic and infrasonic signal examples to train our model—the

VOlcano Infrasound and Seismic Spectrogram Network (VOISS-Net), and compare its tremor classification timelines with time-referenced optical and remote sensing observations. We also demonstrate the transferability of our VOISS-Net seismic model to previous historical eruptions, despite changes in instrumentation and variations in eruption style and duration. The manuscript is organized as follows: Section 2 introduces the local geophysical network at Pavlof Volcano, and provides a chronological summary of the 2021–2022 eruption of Pavlof Volcano. Section 3 describes the process of labeling both seismic and infrasound spectrograms, training the VOISS-Net models, and implementing the models on continuous data. Section 4 discusses the VOISS-Net classification results after applying the model to the 2021–2022 eruption, as well as past eruptions in 2007, 2013, 2014, and 2016. Lastly, the conclusions and implications of this work are summarized in Section 5.

2. Data and Observations

2.1. Seismic and Infrasound Data

AVO operates a local seismic network of six seismometers on Pavlof Volcano, each with a colocated infrasound sensor (Figure 1a). At the time of the 2021–2022 eruption, three of the seismometers (PN7A, PS1A, PS4A) were broadband sensors (Nanometrics Trillium Compact, 120 s low frequency corner), while the other three (PV6A, PVV, HAG) were short-period sensors (Sercel L-22, 0.5 s low-frequency corner). A short-period seismometer at PV6 was part of the local network, until it was destroyed during the 2016 eruption. All infrasound sensors are Chaparral 64-UHP sensors with a 50 s low-frequency corner that were installed after the 2016 eruption. Data are sampled at 50 Hz on Nanometrics Centaur digitizers. We choose to exclude data from the more distant station HAG, due to a poor signal-to-noise ratio and persistent electronic noise that overlaps with the observed tremor frequency band in the seismic records. The operational stations considered here are ∼7–12 km from the summit. We analyze only the vertical component of the seismic data. We remove the instrument response for all data using a frequency-domain deconvolution with a cosine-tapered bandpass filter between 0.005 and 25 Hz in ObsPy (Beyreuther et al., 2010). All times are reported in UTC.

Prior studies of the 2013 and 2014 Pavlof Volcano eruptions relied on local seismic data to detect tremor (Waythomas et al., 2014; Wech & Creager, 2008) and quantify real-time seismic amplitude (RSAM) (Endo & Murray, 1991) as an indication of eruption intensity (Waythomas et al., 2014, 2017). The correlation of eruption tremor amplitude and ash plume height during the 2016 eruption was investigated by Fee, Haney, et al. (2017) and Haney et al. (2018). Studies of Pavlof Volcano eruptions before 2000 had relied on a short-period instrument network that predates the current instruments, but was successful in detecting long-period seismicity, explosions, and tremor (McNutt, 1986). The short-period seismic data were also used to document seismic event rates for comparisons with tidal and tectonic observations (McNutt, 1987). The 2021–2022 eruption is unique as it was the first Pavlof Volcano eruption to be monitored with a local broadband seismic and infrasound network. In addition, the 2021–2022 eruption was the longest in historical times, lasting about 493 days (1 year, 4 months), while many of its prior monitored eruptions only lasted several weeks to a few months (Waythomas et al., 2014).

2.2. The 2021–2022 Eruption of Pavlof Volcano

The 2021–2022 eruption of Pavlof Volcano lasted from 5 August 2021 to 17 December 2022. However, prior to eruption onset as defined by AVO, retrospective analysis shows that Pavlof Volcano exhibited months of elevated unrest, a behavior that is uncommon due to the absence of detectable precursors in many of its preceding eruptions (Pesicek et al., 2018).

Background activity at Pavlof Volcano from January to May 2021 was characterized by AVO as occasional short, low-energy bursts of seismic tremor and long-period earthquakes (Orr et al., 2024). On 11 May 2021, gliding harmonic tremor was first noted during duty checks alongside several moderately deep long-period earthquakes located 15–20 km below sea level. Low confidence satellite SO_2 detections were also noted by AVO later that month, but the lack of elevated surface temperatures and high amplitude seismicity led AVO to keep the Aviation Color Code at GREEN, which is indicative of background activity. However, on 9 July 2021, harmonic tremor was more energetic, infrasonic tremor was noted on the local infrasound network for the first time since unrest started, and vapor emissions were observed in web-camera images of the summit. This prompted AVO to raise the Aviation Color Code and Volcano Alert Level to YELLOW/ADVISORY, which is indicative of elevated

volcanic unrest above known background levels. Tremor bursts, lasting for tens of seconds to minutes, continued through the rest of the month, and on 5 August 2021, minor ash emissions and accompanying explosion signals were observed in webcam images and geophysical data. The explosive ash emissions prompted AVO to raise the Aviation Color Code and Volcano Alert Level to ORANGE/WATCH, officially marking the start of the eruption as documented by AVO. A M8.2 earthquake that occurred on 29 July 2021, south of Chignik, Alaska (255 km east of Pavlof Volcano) (Liu et al., 2022; Ye et al., 2022), could have been a possible trigger for the 5 August 2021 eruption. However, retrospective analysis of seismic data revealed a small ground-coupled airwave from Pavlof Volcano on 28 July 2021, which was likely an early explosion associated with the eventual eruptive activity.

The active eruption phase of the 2021–2022 eruption was characterized by intermittent explosions that produced diffuse ash clouds that drifted in the vicinity of the volcano up to 10 km laterally from the summit before dissipating. Low-amplitude explosion signals were identified in both local seismic data and regional infrasound arrays, and Reverse-Time Migration (RTM) analysis (Fee et al., 2021) on 5 August 2021 using local infrasound data suggested an explosion source location on the upper southeast flank of the edifice. The three previous eruptions in 2013, 2014, and 2016 were from vents on the volcano's summit, while 2007 eruptive activity was from a vent on the upper southeast flank. Synthetic Aperture Radar (SAR) images acquired on 7 August and 18 August 2021 confirmed the upper southeast flank location of the 2021 vent and showed that its crater was widening and deepening as low-level explosive activity persisted. The active crater migrated slightly east by the end of November as explosions and ash emissions continued. Moderately to strongly elevated surface temperatures were also regularly noted by AVO after 8 November 2021. Although web-camera views of the volcano were largely obscured by clouds, the increase in thermal unrest was inferred to be a result of active lava fountaining and related fountain-fed lava flows on the upper southeast flank. Incandescent lava was eventually confirmed in a WorldView-3 Shortwave Infrared (SWIR) band image obtained on 11 November 2021.

As the eruption progressed into 2022, low-level explosions and ash emissions continued to occur, albeit less frequently, while thermal output remained high. Nearly continuous sequences of broadband tremor also persisted throughout most of 2022, although this was punctuated by occasional periods of frequent explosive activity. Interestingly, retrospective RTM analysis of the explosion signals indicated that some of the explosions originated from the summit crater, which was last active during the 2016 eruption. Ash and ballistic ejecta were observed around the crater in satellite imagery, thus confirming this inference. The last explosion signal from the 2021–2022 eruption was detected on 7 December 2022. Seismic tremor gradually waned, and observations of elevated surface temperatures and incandescent lava ceased. AVO lowered the Aviation Color Code and Volcano Alert Level to YELLOW/ADVISORY on 17 December 2022. While at YELLOW/ADVISORY, seismicity remained above background levels with intermittent brief episodes of seismic tremor, but explosions had ceased. Weakly elevated surface temperatures and minor steaming from the active southeast flank crater were observed intermittently in both satellite and web-camera images, which was consistent with cooling lava on the surface. As seismicity decreased to background levels, AVO lowered the Aviation Color Code and Volcano Alert Level of Pavlof Volcano to GREEN/NORMAL on 19 January 2023. The volcano continued to show occasional weak long-period earthquake signals and tremor until March 2023, but no further volcanic activity was noted.

3. Methods

3.1. Spectrogram Classes and Labeling

The 2021–2022 eruption of Pavlof Volcano exhibited a variety of tremor types over a period of unrest and eruption lasting at least 493 days. In order to begin constructing a comprehensive labeled data set to train our models, we first visually scan the seismic and infrasound data to determine a representative set of classes and a time period appropriate for our labeling procedure. This time period should capture the observed tremor varieties in both seismic and infrasound data and be focused near the eruption onset, such that its applicability to unseen data from later unrest can be appropriately examined. We plot hour-long vertical seismic and infrasound spectrograms starting from the retrospectively determined onset of seismic unrest in May 2021, and manually examine 5 months of data to determine our spectrogram classes. All spectrograms are computed using a 10 s Hann window with a 90% overlap over the frequency limits of 0.5–10 Hz, in which our target signals appear most prominently. This frequency band was selected to reduce the impact of lower frequency microbarom signals and wind noise and capture the dominant tremor band.

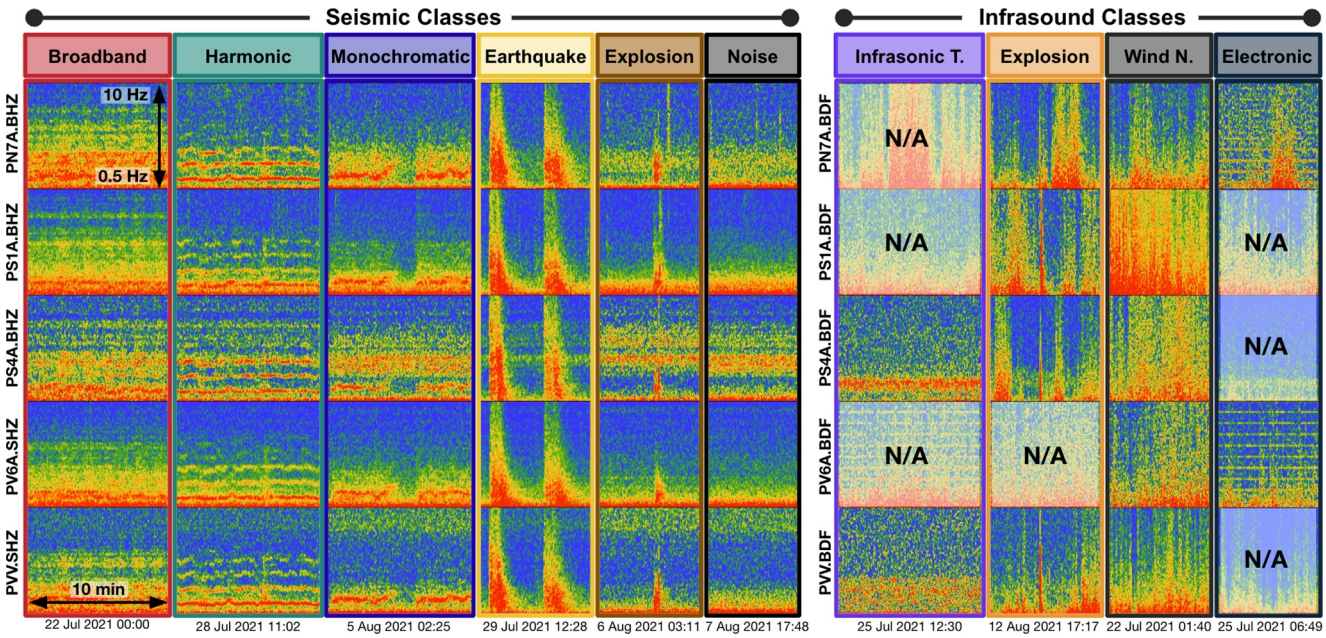

Figure 2. Selected labeling classes for both data types, presented as seismic and acoustic spectrograms calculated from the 5 stations used in our analysis. Class type examples are separated by data type, and are presented as columns of appended spectrograms from the different stations of the Pavlof Volcano network. Each station spectrogram spans 0.5–10 Hz on the y-axis, and are scaled similarly by time. The classes are determined from the variety of signals observed during the early months of the 2021–2022 eruption, and the examples shown above are obtained from our labeling period spanning 22 July 2021–22 September 2021. The precise start time of each example shown is annotated under each column. Note that the infrasound classes are often not consistent across all stations, due to wind or electronic noise masking the signal of interest or unfavorable propagation effects. Infrasound station spectrograms that do not adequately represent the class type suggested in the column header are annotated as not applicable ("N/A").

AVO duty seismologists noted diverse tremor spectral signatures and numerous volcanic explosions evident from their ground-coupled airwaves on spectrograms (Fee et al., 2016; Smith et al., 2016) during the early months of this eruption. The large mainshock-aftershock sequence associated with the M8.2 Chignik earthquake on 29 July 2021 (Liu et al., 2022; Ye et al., 2022) was also recorded in the spectrograms as high energy, broadband, and long duration seismic arrivals over several days during this time period. These observations, coupled with periods of quiescence, provided us with the opportunity to label unique tremor types against background noise and transient events such as explosions and both local and regional earthquakes. We label a comprehensive 2 month period from 22 July 2021 to 22 September 2021, using the following class types and methodology which are also illustrated in Figure 2. Note that the following class definitions are specific to this work, and that there is variability in how volcano seismic and acoustic signals are defined in relevant literature (Matoza & Roman, 2022; Wassermann, 2012).

Seismic classes and definitions:

1. *Broadband tremor*: long-duration (minutes) seismic signals with a broad (>1 Hz) frequency peak below 5 Hz.
2. *Harmonic tremor*: long-duration signals (minutes) with multiple narrow-band (<1 Hz) overtones with energies comparable to that of the fundamental tone.
3. *Monochromatic tremor*: long-duration (minutes) signals with a narrow-band fundamental tone and no energetic overtones.
4. *Earthquake*: transient signals (seconds) produced by local and regional seismic events. Local earthquakes typically show an energetic broadband onset with a lingering low-frequency coda. Distinguishable P- and S-wave packets are also used to verify distal regional earthquakes.
5. *Explosion*: transient signals (seconds) which are reliably determined to be volcanic explosions either via (a) ground-coupled airwaves evident in seismic data, or (b) retrospective RTM analysis (Fee et al., 2021) of infrasound data with a source location near the active vent. Explosions generally have a low frequency (<5 Hz) seismic onset followed by a higher frequency, typically broadband, ground-coupled airwave.
6. *Noise*: catch-all class for any non-systematic variations in seismic spectral signature and instances of quiescence where no notable signal of any kind is observed.

Infrasound classes and definitions:

1. *Infrasonic tremor*: long-duration (minutes) acoustic signals that exhibit a stable peak frequency over time. Although infrasonic tremor has been observed to show harmonic behavior at other volcanoes (Cannata et al., 2013; Fee et al., 2010; Watson et al., 2020), the homogeneity of acoustic tremor seen at Pavlof Volcano during the 2021–2022 eruption (and the sparsity of different infrasonic tremor types) leads us to condense all observed infrasonic tremor into one class.
2. *Explosion*: shares the same definition as the similarly named Explosion seismic class.
3. *Wind noise*: catch-all class for any non-systematic variations in acoustic spectral signature and instances of quiescence (which is interpreted as low wind). Wind noise often dominates the infrasound records at Pavlof Volcano due to persistently high winds and the exposed locations of the sensors.
4. *Electronic noise*: any persistent, non-physical spectra typically consisting of sharp, invariable frequency peaks at integer hertz values.

We label the above classes on seismic and infrasound spectrograms spanning 22 July to 22 September 2021 using the open-source software Label Studio (Tkachenko et al., 2020). Specifically, we produce and upload hour-long spectrograms (appended vertically by station) onto Label Studio, and use the bounding box labeling tool to draw boxes around each class observed on the spectrograms. We also overlay regional earthquake origin times queried from the Advanced National Seismic System (ANSS) Comprehensive Earthquake Catalog (ComCat) (U.S. Geological Survey, 2017), RTM-derived explosion times (Fee et al., 2021), and hourly averaged spectra to guide our labeling effort, illustrated in a screenshot in Figure S1 of the Supporting Information S1. The spectrogram axis labels are removed during labeling, which allows the exported pixel bounds to be easily referenced for the station and class associations and their corresponding start and end times. Instances where the associated class is ambiguous, such as during periods of high noise or superimposed classes, are not labeled.

After the pixel bounds of all labeling boxes are exported and the temporal bounds of all class observations are derived, we divide each hour-long, station-specific spectrogram into non-overlapping 4-min spectrogram slices. We then determine the percentage of each class observed in the spectrogram slice. We assign the final label to each 4-min spectrogram slice based on the most prominent class observed within it, which corresponds to the class that has the highest number of time samples associated with it. Note that transient signals such as earthquakes and explosions are prioritized over tremor and noise as long as 10% of the spectrogram time samples (i.e., 24 s) are associated with these transient events. For the other classes (tremor and noise), we require a minimum of half the time samples being meaningfully labeled prior to granting each slice its label. Slices with no or insufficient labeled time samples are discarded. Following this convention, we obtain a total of 6235 Broadband Tremor, 886 Harmonic Tremor, 4536 Monochromatic Tremor, 9762 Earthquake, 846 Explosion, and 64262 Noise seismic spectrogram slices, and 5964 Infrasonic Tremor, 452 Explosion, 59847 Wind Noise, and 12214 Electronic Noise infrasound spectrogram slices. A bar graph showing the class distributions is included in Figure S2 of the Supporting Information S1.

3.2. Model Architecture and Training

The proposed VOISS-Net CNN architecture (LeCun et al., 1998) contains three consecutive convolutional and max-pooling layer pairs, a flattening operation, and two dense layers that lead up to a softmax classification layer (Figure 3). Convolutional layers apply convolutional kernels to extract features by multiplying them with localized regions of input data, while max pooling layers downsample the output by retaining maximum values within these localized regions. Dense layers then perform a linear transformation on the derived features, which are then put through an activation function to obtain even higher level features used for classification. We use a separate network for each data type due to their differing spectrogram input units and the number of output classes used (6 for seismic vs. 4 for infrasound). Each spectrogram slice that is fed into the CNNs is a 94×240 matrix with a power value in each pixel, calculated relative to 1 m/s for seismic and 20 μPa for infrasound.

As illustrated on Figure 3, the three convolutional layers use 32, 64, and 128 filters respectively, each with a 3×3 kernel size and a 1×1 kernel stride. The max-pooling layer that follows each convolutional layer uses a 3×3 pooling window and a 3×3 window stride. Consequentially, the sequence of convolutional and max-pooling operations transform the 94×240 spectrogram into an intermediate output tensor of size $3 \times 8 \times 128$. This intermediate output is then flattened into its 3072 elements and fed into the two dense (also known as fully

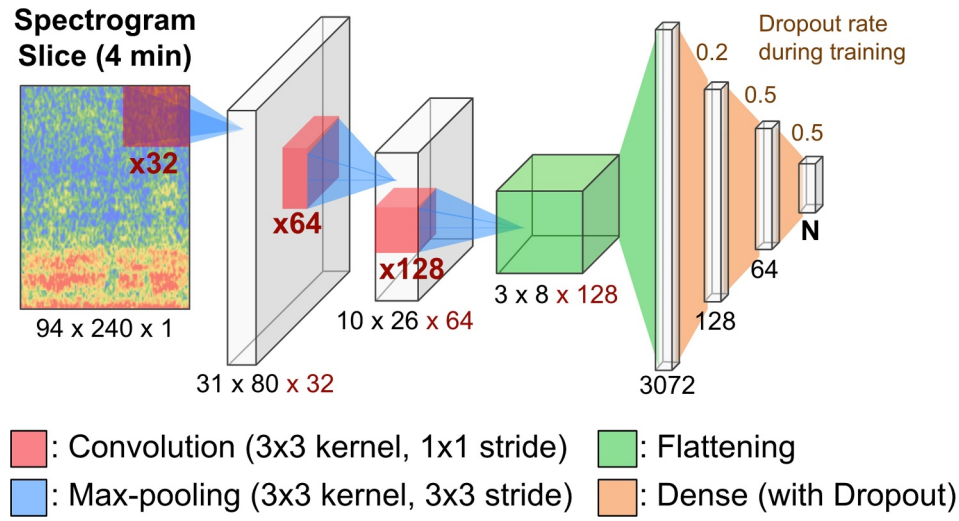

Figure 3. VOISS-Net Convolutional Neural Network (CNN) architecture used for each data type. Label Studio bounding box labels are used to assign classes to non-overlapping, 4-min spectrogram slices, which are used as inputs to the CNN model. The example shown is from the Broadband Tremor seismic class. Dropout layers are only applied during training time, and are not used during testing and implementation. The length of the output layer *N* depends on the data type, as the seismic and infrasound models contain 6 and 4 output classes respectively.

connected) layers with 128 and 64 nodes respectively. During training, we implement dropout layers (Srivastava et al., 2014) after the flattening operation, with rates of 0.2 for the flattened layer, and 0.5 after each of the two dense layers. The penultimate dense layer is then connected to the classification neurons with softmax activation functions (Goodfellow et al., 2016) to evaluate the probabilities of each seismic or infrasound class within each spectrogram slice. Our selected model architecture results in seismic and infrasound models that contain 494,662 and 494,532 total trainable parameters respectively.

The VOISS-Net CNN architecture layer sequence follows a similar structure from earlier studies that have applied CNNs to seismic spectrograms (Ferreira et al., 2023; Kong et al., 2022; Linville et al., 2019; Mousavi et al., 2019; Rouet-Leduc et al., 2020). We note that Mousavi et al. (2019), Linville et al. (2019), and Kong et al. (2022) utilized three component seismic data by combining horizontal component (either North and East or rotated Radial and Transverse) spectrograms with the vertical seismic spectrogram prior to forward feeding. Although three component seismic data exists for the 2021–2022 Pavlof Volcano eruption, our intention to transfer our network to older Pavlof Volcano data that predates three-component instrumentation or other volcano seismic networks leads us to focus on vertical component spectrograms only. Ferreira et al. (2023) utilized a unique approach of concatenating convolutional outputs from multiple stations prior to classifying the type of volcano-seismic signal (i.e., early integration). Here, we choose to develop a station-specific model that can be applied independently to each station, and devise a method to consolidate multi-station output probabilities (i.e., late integration), which will be discussed later. Doing so grants us the flexibility to apply our model on seismic and acoustic networks of different sizes or when stations are temporarily unavailable. An additional characteristic of our classification procedure that distinguishes it from other studies is the inclusion of a noise class for each data type. Although the addition of a noise class introduces the risk of slower model convergence, it allows our model to differentiate signal against noise when applied to continuous, successive spectrogram windows.

In order to train and evaluate our model, we first split our labeled data set into training, validation and test data sets for all classes across both seismic and infrasound data types. The training data set is used to update model weights, the validation data set is used to assess model performance during training and prevent overfitting, and the test data set is used to provide a final, unbiased assessment of the trained model. We note that our labeled data set is unbalanced with many more Noise, Earthquake, and Broadband Tremor classes than Explosion and Harmonic Tremor classes (Figures S2a and S2b in Supporting Information S1). We therefore take steps to ensure that our training and validation data sets are balanced across the classes to speed up model convergence, reduce model biases and overfitting (Buda et al., 2018), and retain robustness in our reported evaluation metrics such as classification accuracy. We start by constructing the test and validation data sets by randomly sampling across the

seismic and infrasound classes. The number of random samples for each data type is determined by the smallest class count for each data set, which is the Explosion class (846 in the seismic data set and 452 in the infrasound data set). By taking sample quantities equivalent to 20% of those numbers (169 and 90, respectively), we obtain balanced validation and test data sets that consist of 1014 class examples for seismic and 360 class examples for infrasound. As the remaining samples are inevitably unbalanced, we introduce a novel data augmentation strategy to balance the class numbers in the training data set. Data augmentation, which is the process of artificially expanding a data set by making modifications or transformations to existing data samples, is typically employed for CNNs by recoloring, stretching, zooming or translating the subject of interest in labeled images (Shorten & Khoshgoftaar, 2019). For seismic and infrasound spectrograms, stretching or zooming into an image with clear geophysical meaning might be undesirable, and pitch-adjusting augmentation strategies from audio engineering applications (McFee et al., 2015) are inappropriate due to the lack of harmonics in some signals. While the use of synthetic signals (Witsil et al., 2022) or the addition of synthetic noise (Zhao et al., 2022) has been used to augment labeled geophysical data, we choose not to employ those strategies in order to avoid integrating human-engineered biases into our labeled data set. Instead, we use the surplus of noise examples in both data types to augment our data set. Augmented spectrogram slices are generated by summing, in an element-wise manner, 65% and 35% of the spectrogram amplitudes from a random training and noise spectrogram slice respectively (Figure S2c in Supporting Information S1). Data augmentation is carried out for the sparser classes until their class counts match that of the most abundant tremor class. This ultimately provides us with a training set of size 35,382 for seismic (5897 per class) and size 23,136 for infrasound (5784 per class).

We set up our model architecture using Tensorflow (Abadi et al., 2016) and train our models using the Categorical Cross Entropy loss function (Goodfellow et al., 2016), the Adam optimizer (Kingma & Ba, 2014), and a learning rate of 0.0005. Other learning rates did not produce significantly faster convergence or improved results. Batches of 100 spectrogram slices are standardized by each pixel using running mean and variance values that are updated per training batch with a momentum of 0.90 (Yang et al., 2022), before being normalized by the minimum and maximum pixel values per spectrogram slice. We additionally implement an early stopping checkpoint, which tracks decreases in validation loss during training to reduce overfitting. When the validation loss does not decrease over 20 epochs during training, we terminate training and save the best-performing CNN iteration, as well as the final mean and variance matrices calculated during training. The corresponding learning curves of the final VOISS-Net models are shown in Figure S3 of the Supporting Information S1 and the confusion matrices of both the seismic and infrasound test sets are shown in Figure 4. Overall, the seismic CNN converged in 19 epochs to achieve a test set accuracy of 81.2%, while the infrasound CNN converged in 15 epochs to achieve a test set accuracy of 90.0%. Both VOISS-Net models were trained in less than 3 hr using a standard 8 CPU computer and less than 10 min using a single GPU card. Both seismic and infrasound VOISS-Net learning curves show a smooth decrease in validation loss (and a smooth increase in test set accuracy) until the point of the early stop, where validation loss starts to take on an erratic increasing trend. As the validation loss and training loss curves deviate from one another only after the early stop, we are confident that model overfitting is minimal.

3.3. Model Results and Validation

The confusion matrices in Figure 4 highlight the strong performance demonstrated by both the VOISS-Net seismic and infrasound models, as both demonstrate high class-specific accuracies as shown by the values down the diagonal. With the exception of transient signals (Earthquakes and Explosions), which are arguably difficult to differentiate even for dedicated models and analysts, the other class-specific accuracies exceed 80%, demonstrating the ability of the models to distinguish tremor and noise. Admittedly, the off-diagonal values within each matrix cannot be ignored and can provide important insight into model performance. For instance, we find that our VOISS-Net seismic model confuses the different seismic tremor types, where 11% of the Broadband Tremor examples are misclassified as Monochromatic Tremor, and 12% of the Harmonic Tremor examples are misclassified as Broadband Tremor. Our seismic model also struggles to differentiate Earthquakes and Explosions, where 15% of the Earthquake examples are misclassified as Explosions, and 12% of the Explosion examples are misclassified as Earthquakes. For the VOISS-Net infrasound model, the Explosion class is most often mistaken as Wind Noise, with 12% of the Explosion examples misclassified as such. In spite of these model imperfections, we still find that our model is useful and that its performance is realistic, as the type of

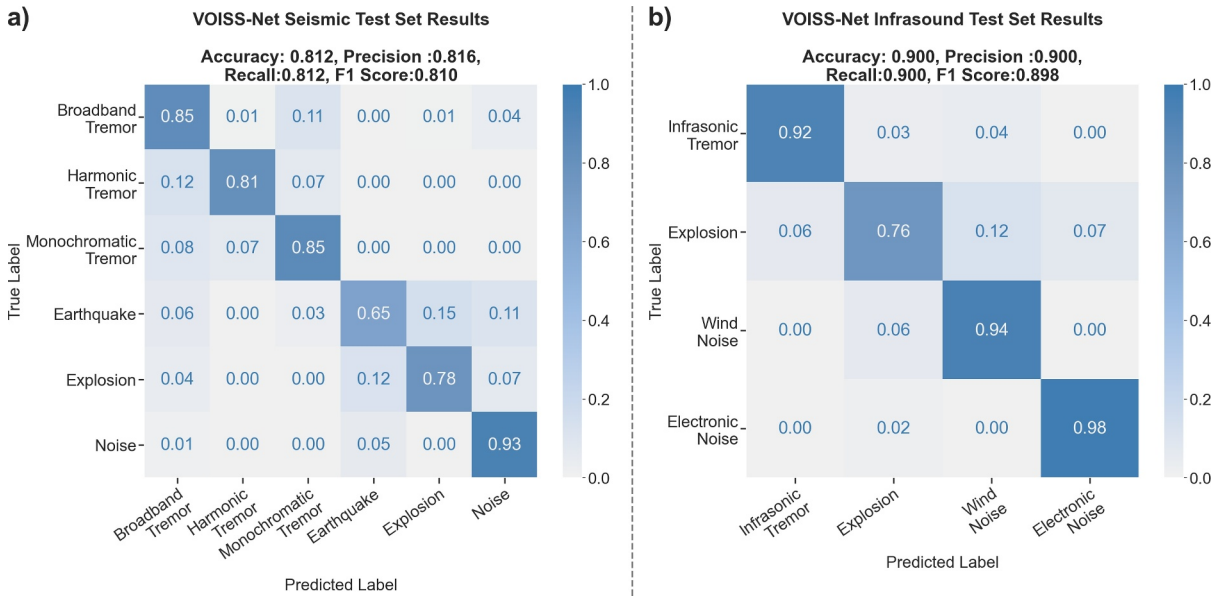

Figure 4. Test set confusion matrices for both (a) seismic and (b) infrasound VOISS-Net models. Rows and columns correspond to true and predicted labels, respectively, and each confusion matrix value C_{ij} corresponds to the ratio of the number of observations predicted as label j out of the number of observations known to have label i. The ratio of true positive classifications per class is reflected down the diagonal. Standard summary statistics such as accuracy, precision, recall and F1 score across all classes are shown above each confusion matrix. The VOISS-Net models perform well on the seismic and infrasound test sets, achieving an accuracy of 81.2% and 90.0% respectively. Individual class performance appear satisfactory as well, with high class-specific accuracies down each diagonal. However, variability in the class-specific accuracies exist, with transient signals such as the Earthquake and Explosion classes showing a slightly poorer performance.

classification mistakes made by our model are mistakes that could be made by a human analyst as well. For example, differentiating between the seismic tremor types can at times be challenging even for an analyst when the width of the spectral peak is ambiguously broad or narrow or if the presence of a higher frequency overtone is subtle. Our spectrogram computation window and segment length for each VOISS-Net model (10 s and 4 min, respectively) were also chosen to facilitate classification of longer duration tremor instead of transient signals such as Explosions and Earthquakes, which make up a significant proportion of the misclassified examples. Overall, the summary statistics of each model, such as their accuracy and F1 score, are still reasonably high, despite our classifier using only single component data and incorporating more classes than those from earlier spectrogram classification studies (Kong et al., 2022; Linville et al., 2019; Mousavi et al., 2019). Lastly, it should be noted that the confusion matrices in Figure 4 correspond to station-generic models. When implemented on a network of stations, the VOISS-Net models are applied on each station independently before a higher quality network result is derived, a process which will be described later.

To robustly evaluate the performance of our model, we iterate our train-validation-test split, data augmentation, and model training process for each data type using 50 different randomization seeds. Doing so allows us obtain a statistically significant distribution of test set accuracies (Figure S4 in Supporting Information S1) to ensure that our chosen model's performance falls within an expected performance range. It also eliminates the possibility that our chosen models demonstrate a skewed result due to a lucky (or unlucky) draw of training, validation, or test set examples. Our selected VOISS-Net seismic model, which achieved an accuracy of 81.2% on its test set, compares well with the mean accuracy of the 50 retrained seismic models on their test sets, which is 82.1% $\pm$ 3.1% (2σ). The selected VOISS-Net infrasound model, which achieved an accuracy of 90.0% on its test set, also compares well with the mean of the 50 retrained infrasound models, which is 90.2% $\pm$ 3.8% (2σ).

After finalizing our VOISS-Net models, we further investigate and validate our models using Gradient-weighted Class Activation Mapping (Grad-CAM) (Selvaraju et al., 2017). Grad-CAM, which has also been applied to the analysis of earthquake and explosion spectrograms by Kong et al. (2022), extracts the gradients of the output class relative to the final convolutional layer in order to determine areas of the input image that contribute most to the model's output. This approach is driven by the idea that the final convolutional layer contains the highest level features while retaining spatial coherence relative to the input image. When these gradient maps are resized and

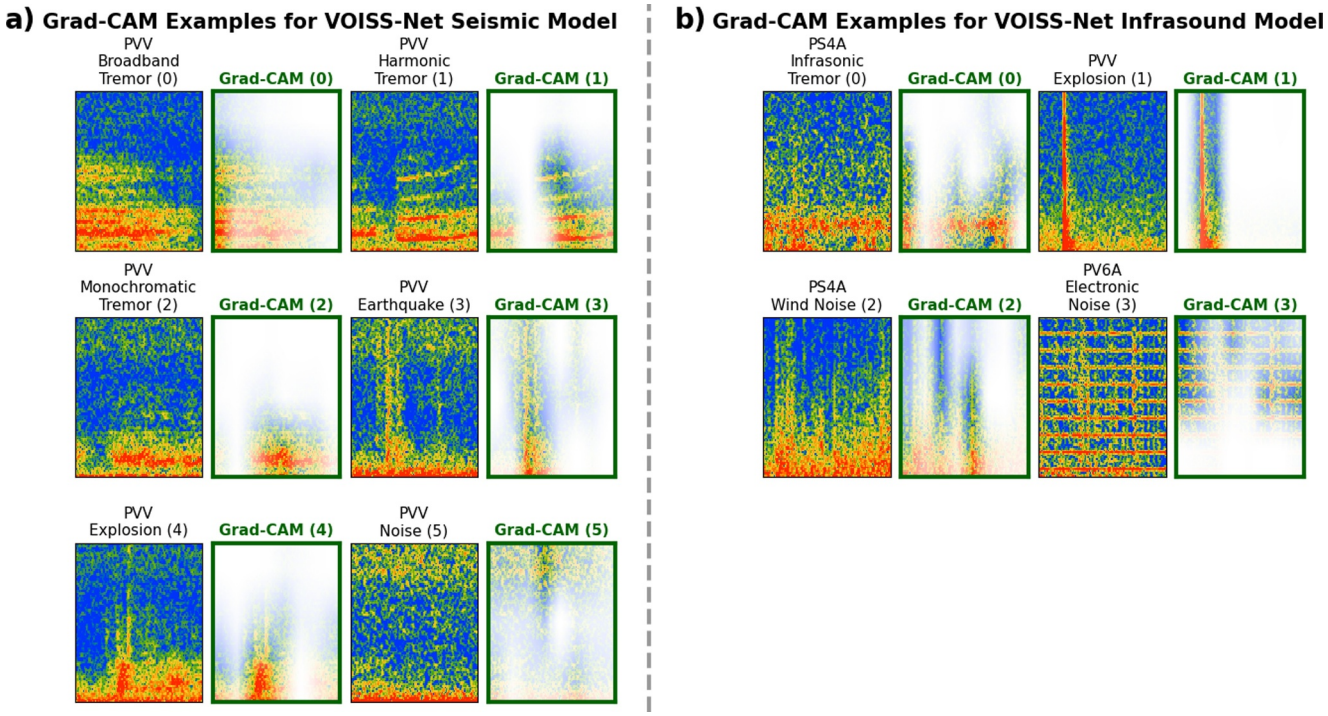

Figure 5. Gradient-weighted Class Activation Mapping (Grad-CAM) results for selected VOISS-Net (a) seismic and (b) infrasound test set examples. Squared normalized Grad-CAM weights are mapped as opacity values for each spectrogram slice, where opaqueness is coarsely indicative of pixel importance. Stations PVV and PS4A are chosen as examples. In each example, Grad-CAM successfully highlights the pixels surrounding each signal of interest, except for the Noise classes which are more random. Note that the pixels which Grad-CAM highlights changes with each input 4 min spectrogram segment, even if segments belong to the same class.

extrapolated to fit the input dimensions, they can be used as "importance" heatmaps to understand the relative contribution of each frequency-time bin of the input spectrogram to the model, albeit coarsely. Figure 5 illustrates the Grad-CAM results for both models by mapping the normalized Grad-CAM weights as opacity values on top of each chosen class example. Across all signal classes (Tremor, Earthquakes, and Explosion), Grad-CAM generally highlights the area surrounding the signals of interest, and noticeably captures patterns such as the wider frequency range of Broadband Tremor and Harmonic Tremor, narrow-band characteristics of Monochromatic Tremor and Infrasonic Tremor, and the transient nature of Earthquakes and Explosions. Interestingly, in the Explosion example from the seismic model, Grad-CAM focuses on the lower frequency seismic onset and the broadband ground-coupled airwave that succeeds it, both of which are guiding features during our manual labeling exercise. However, less can be said about the Grad-CAM results of the different Noise examples, which appear more random. Here Grad-CAM seemingly captures the chaotic nature of Wind Noise and the irregular breaks in artificial spectral peaks in the Electronic Noise example.

These validation steps give us confidence that the application of our model to the seismic and infrasound data associated with the 2021–2022 eruption of Pavlof Volcano will be fruitful. Unlike during labeling, where spectrograms are cut into non-overlapping 4-min segments, during model implementation we apply each VOISS-Net model on 4-min spectrogram segments with a 75% overlap (1 min time step), where the output class of each time step corresponds to the ±2-min window around it. We also discard any spectrogram segments that contain data gaps, which we identify using a low spectral amplitude-based exception. As the seismic and infrasound VOISS-Net models are applied on each station independently, we obtain softmax probabilities across each set of classes per station for each time step. We identify the class with the highest probability on each station for each time step to obtain station-specific class timelines. We also calculate the average class probabilities across the contributing stations for each time step to determine network-averaged probabilities and a subsequent network-informed class "vote" (Witsil et al., 2022). The network-averaged class probability for each network result, which we term P_{norm}, is included as an accompanying confidence metric.

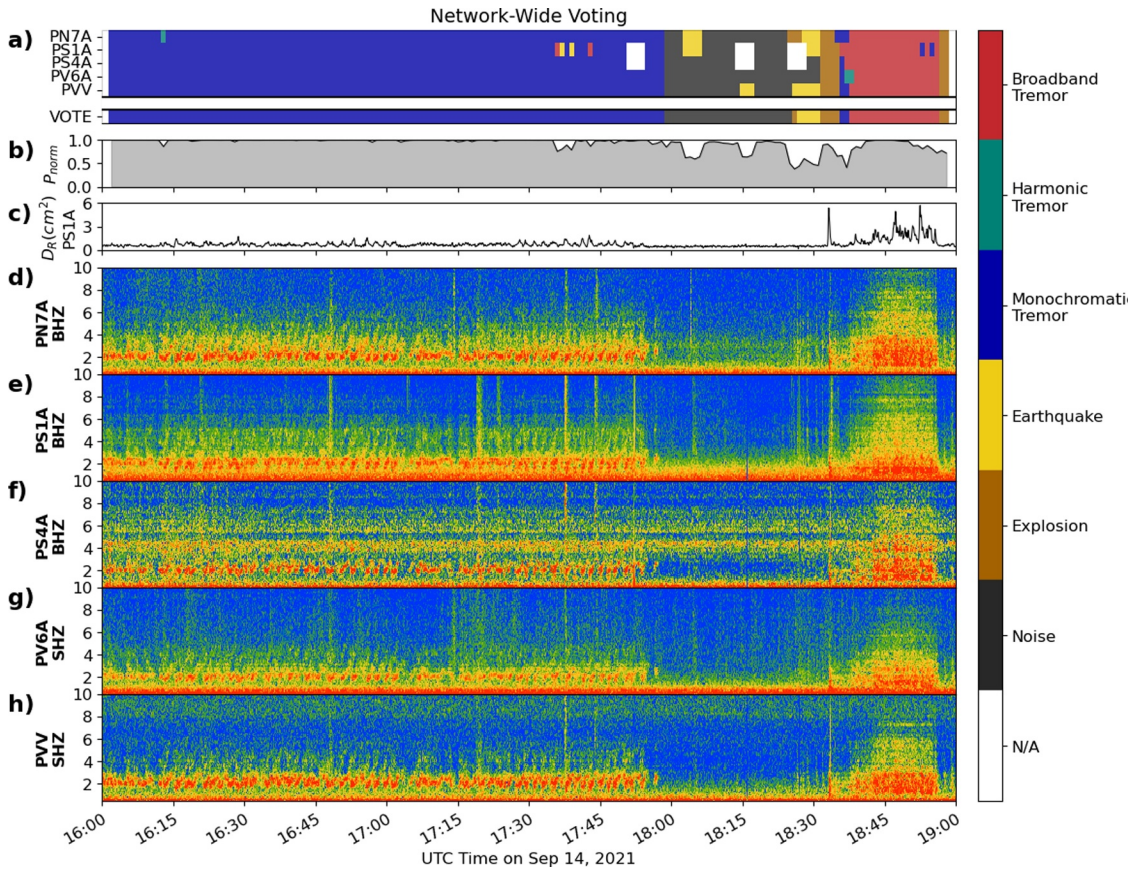

Figure 6. Example VOISS-Net classifications for a 3 hr seismic sequence from 14 September 2021. (a) Single station and network vote classifications, with the class denoted by the color bar on the right. (b) Network vote probability averaged from individual station probabilities. (c) Reduced displacement for a reference station (PS1A). (d–h) Spectrograms for each of the seismic stations considered. The sequence demonstrates a monochromatic tremor episode from 16:00 to 17:55, which abruptly shuts off into a period of quiescence. Small earthquakes start to occur around 18:25, which leads into an 18:35 explosion signal that is followed by broadband tremor until 18:55. The VOISS-Net seismic model does well in classifying the sequence of observed signals, with the network voting result returning a more reliable result.

4. Results and Discussion

4.1. VOISS-Net Application

We apply the VOISS-Net models to data between 1 January 2021 and 1 March 2023. However, prior to discussing the VOISS-Net results for the entire 2021–2022 eruption, we present the implementation of our models on two 3-hr periods from the eruption that contain clear, diverse spectrogram signals. These examples allow us to visualize each model's performance on a short timescale, and elucidate the network-wide voting scheme used to consolidate the results from the different stations into a single network timeline.

Figures 6 and 7 show VOISS-Net classifications, network-wide probabilities, and spectrograms from a 3 hr seismic sequence on 14 September 2021 and a 3 hr infrasound sequence on 6 August 2021. Both figures illustrate the results from our VOISS-Net models, which can run for an hours-long 50 Hz data sequence in well under a minute on an 8 CPU computer, assuming that a manageable number of stations (<10) are used. Each figure presents the spectrograms from all input stations on the lower panels (Figures 6d–6h and 7c–7g), and the VOISS-Net seismic or infrasound output on the upper panels (Figures 6a, 6b, 7a, and 7b). As described in the previous section, we consolidate station-specific VOISS-Net outputs into a network result, which is shown in the network voting ("VOTE") row in Figures 6a and 7a. The corresponding network average class probabilities are plotted as a curve in the P_{norm} panel (Figures 6b and 7b), to provide an indication of model confidence. Lastly, for seismic timelines, we additionally compute and plot the reduced displacement D_R (Aki & Koyanagi, 1981) using a single reference station (PS1A) and assuming that tremor is primarily composed of surface waves (Figure 6c). Reduced

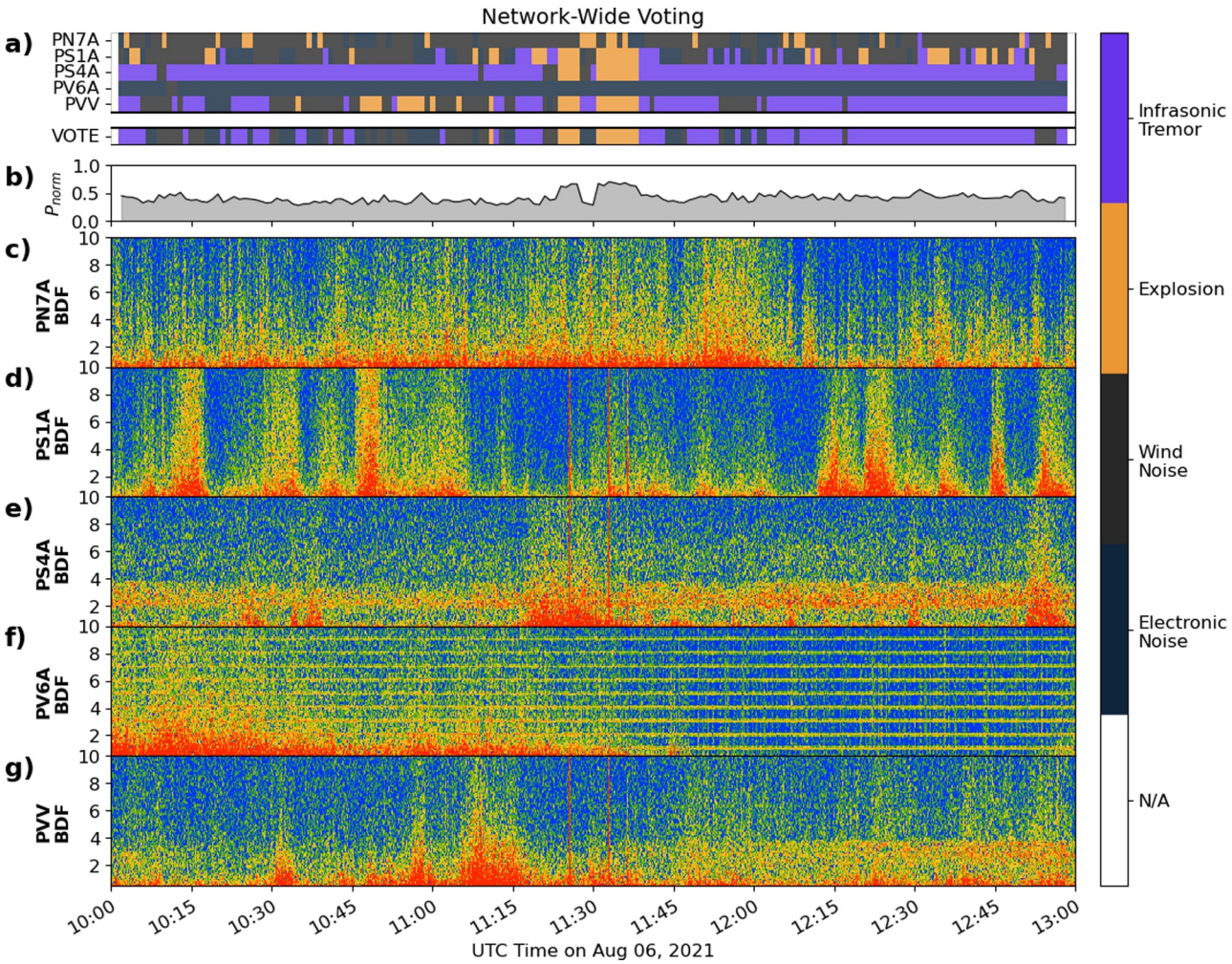

Figure 7. Individual station VOISS-Net classifications and network voting result for a 3 hr infrasound sequence from 6 August 2021. The layout is very similar to (Figure 6): (a) Single station and network vote classifications, with the class denoted by the colorbar on the right. (b) Network vote probability averaged from individual station probabilities. (c–g) Spectrograms for each of the infrasound stations considered. The sequence contains substantial wind noise across all stations and electronic noise on PV6A. Three explosions occur in the middle of sequence at 11:25, 11:32, and 11:36, and are clearly visible on PS1A, PS4A, and PVV spectrograms, but less clearly seen on PN7A. Infrasonic tremor is prominent between ∼2–4 Hz on PS4A throughout the entire sequence, but is intermittent on PVV until emerging more consistently starting around 11:30. The VOISS-Net infrasound model does well in picking out the explosions and infrasonic tremor on individual stations, but inter-station variability (owing to noise and presumably infrasound propagation effects) makes network voting complicated.

displacement provides a measure of seismic amplitude over time, which is useful for interpreting the classification and seismicity trends (Haney et al., 2018; McNutt & Nishimura, 2008).

There are clear trends in seismicity apparent in the spectrograms captured by the VOISS-Net classifications (Figure 6). The sequence contains an ∼2 hr period of monochromatic seismic tremor centered around 2 Hz from 16:00 to 17:55. This monochromatic tremor eventually transitions into 30 min of quiescence, before discrete microseismicity and a clear explosion signal occur, the latter of which is apparent by their high D_R value. The explosion signal is followed by a period of increasingly broadband tremor. Overall, VOISS-Net does well in classifying each portion of this seismic sequence on each independent station, although there are some minor misclassifications scattered throughout the sequence. For example, the period of quiescence preceding the explosion signal has numerous stations returning a transient signal class despite there being no energetic signals, and the period of broadband tremor contains several instances of misclassified tremor types. We attribute these misclassifications to weak high frequency spikes in the data, inter-class ambiguity (e.g., monochromatic vs. broadband tremor), and the challenge of differentiating earthquakes and explosions given our relatively broad analysis window of 4 min and spectrogram computation window of 10 s. Nevertheless, the network voting result

provides a near-perfect description of the seismic sequence that agrees with our own visual examination; it overrules station-specific misclassifications and appropriately assigns high probabilities where misclassifications are minimal.

In comparison, the infrasound sequence classified by VOISS-Net in Figure 7 is more challenging to interpret. Due to significant wind and electronic noise, or infrasound propagation effects (e.g., the influence of source directionality, topography, and wind), infrasonic tremor is typically not seen coherently across the Pavlof Volcano network. In this particular sequence, the infrasonic tremor is centered around 2–4 Hz and is most clearly visible on PS4A, and on the later half of the time sequence on PVV. Recall that the active vents of the 2021–2022 Pavlof Volcano eruption were located on the upper southeastern flank of the volcano, and that stations PS4A and PVV are on the south-southeast flank of the volcano (Figure 1). There are also 3 discrete explosions visible at 11:25, 11:32, and 11:36, which are observed clearly on the south-side stations PS1A, PS4A, and PVV, and less clearly on the north side station PN7A. Station PV6A, located on the north side of the volcano, is not useful in this selected time sequence due to persistent electronic noise apparent as horizontal "streaks" in the spectrogram. When the VOISS-Net results are scrutinized, we notice that the 3 explosions are appropriately identified on the aforementioned stations (Figure 7a), although several explosion misclassifications exist at other time steps. Infrasonic tremor is also appropriately identified on PS4A and PVV whenever they are prominent, but occasional false positives in other time steps exist. In this case, the susceptibility of the model to misclassifications and the absence of common signals observed across the network proves to be a challenge for our network voting strategy. The network voting results and their corresponding probabilities are influenced by noisy stations, which introduces interpretive challenges. Nevertheless, the voting results still provide a reliable description of infrasonic signals observed in the 3-hr sequence. The shortcomings of our model and network voting procedure can be mitigated by using a larger labeled data set for training, or by using a smaller subset of stations with higher signal-noise ratio to constrain the infrasound timeline.

Overall, we find that our VOISS-Net seismic and infrasound timelines replicate what is generally observed in the spectrograms, and provide useful classification information that aids the interpretation of the seismic and infrasonic character of the 2021–2022 Pavlof Volcano eruption. In particular, the network timelines provide a comprehensive description of seismic and infrasonic signals observed, and are generally resistant to single-station outliers or misclassifications. In the following subsections, we expand VOISS-Net's application, first to the years bounding the 2021–2022 Pavlof Volcano eruption, then to Pavlof Volcano's eruptions over the past 20 years.

4.2. Application to the 2021–2022 Pavlof Volcano Eruption

VOISS-Net successfully classifies the general trends in Pavlof Volcano's seismic and infrasound data for the 2021–2022 eruption. Figure 8 shows the multidisciplinary timeline spanning 1 January 2021–1 March 2023, which bounds the 5 August 2021–17 December 2022 eruption. It displays: our seismic and infrasonic VOISS-Net timelines (c, g), the 2-year spectrograms (b, f), the AVO Aviation Color Code (a, which is a qualitative proxy for the level of volcanic unrest), the median-filtered PS1A reduced displacement calculations (d), radiative power (e, in black) (Saunders-Shultz et al., 2024), SO_2 emissions determined by AVO (e, in red) (Lopez et al., 2021), and an explosion catalog comparison between VOISS-Net and one independently derived using the infrasound network and the Reverse Time Migration (RTM) method of Fee et al. (2021) (h). Focusing on the period of background activity from January 2021 to May 2021, the VOISS-Net seismic model accurately returns infrequent seismic tremor classifications which agree well with AVO interpretations of intermittent tremor during this period. However, it is apparent that the seismic model also regularly returns explosion misclassifications long before the eruption onset in August 2021. Upon visual inspection, we find that these are mostly misclassified long-period earthquake activity (Figure S5 in Supporting Information S1); there is no long-period earthquake class in our seismic model because long-period earthquakes were not observed in the time period scanned to shortlist our seismic labels. The VOISS-Net infrasound model also returned little to no infrasonic tremor during this time period as well, and infrasound explosion classifications seem to follow a background rate of misclassifications, which we attribute to noise spikes in the data.

As volcanic unrest picked up in May 2021, the VOISS-Net seismic model reliably returns monochromatic tremor and occasionally harmonic tremor classifications. As unrest continued into July 2021, more instances of broadband tremor were identified, which coincide with AVO's transition to Aviation Color Code and Volcano Alert Level YELLOW/ADVISORY. The corresponding reduced displacement values computed from PS1A

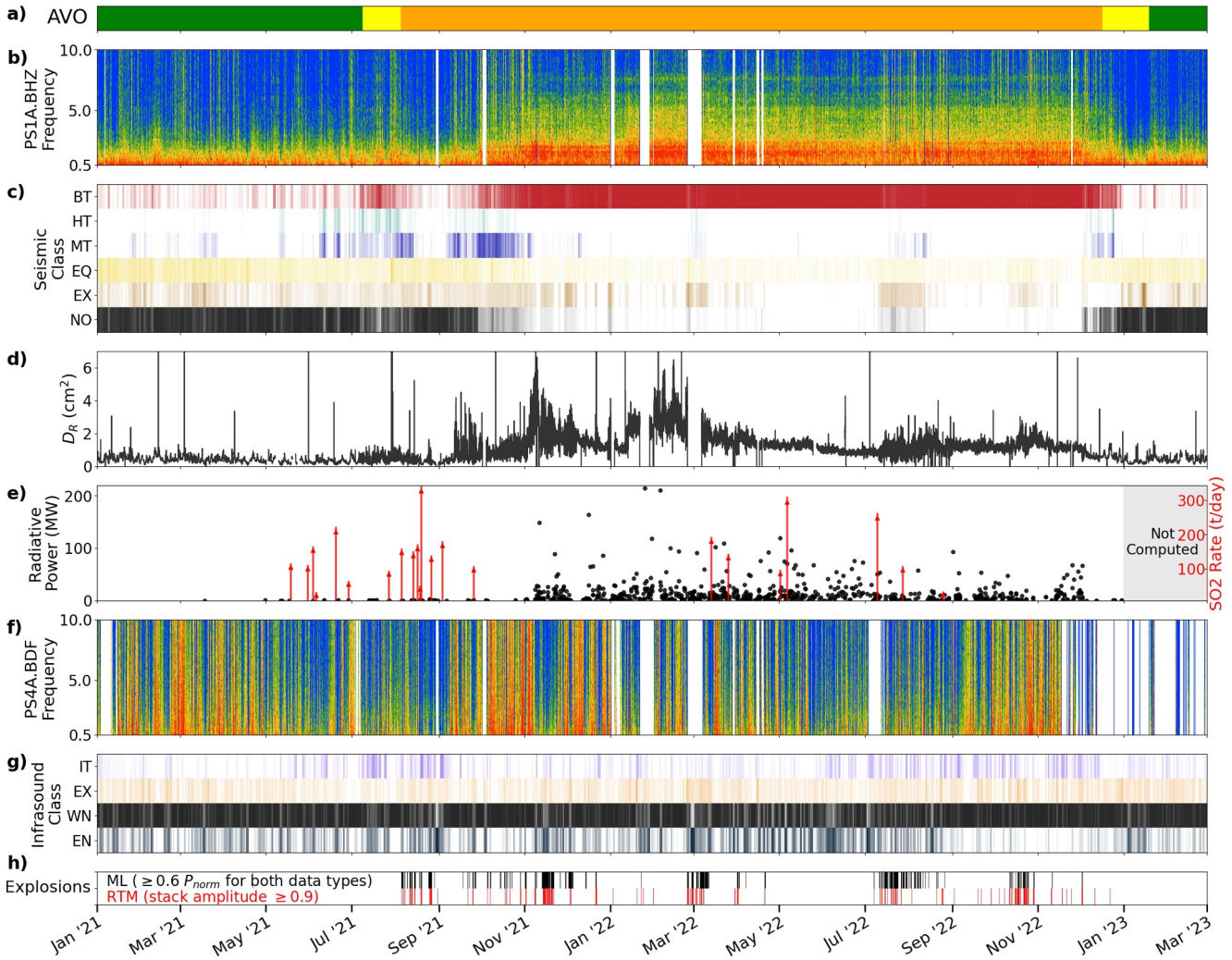

Figure 8. Seismic and infrasound spectra and their corresponding VOISS-Net classifications from 1 January 2021 to 1 March 2023, compared to Alaska Volcano Observatory's (AVO's) unrest timeline and selected multidisciplinary metrics. The panels represent (a) the AVO Aviation Color Code, (b) the seismic spectrogram from PS1A, (c) the VOISS-Net seismic class timeline, (d) the median-filtered D_R calculated from PS1A seismic data, (e) radiative power (black) (Saunders-Shultz et al., 2024) and SO_2 emission rates (red) (Lopez et al., 2021), (f) the infrasound spectrogram from PS4A, (g) the VOISS-Net infrasound class timeline, and (h) the explosion timeline using VOISS-Net results (explosion $P_{\text{norm}} \geq 0.6$ for both seismic and infrasound network classifications) and independent RTM analysis Fee et al. (2021). Note that the VOISS-Net timelines in (c) and (g) are binned by day and separated into class-specific rows; the opacity value of each time bin represents the ratio of each class' occurrence relative to the total number of time steps per day. Individual class names are abbreviated for presentation purposes.

seismic data also started showing significant fluctuations in July 2021, when broadband tremor pulses produced reduced displacement values ranging from 0 to 1 cm^2 (Figure 8d). Similarly, the infrasound model recorded an increasing number of infrasonic tremor classifications beginning in May 2021, with rates peaking around late July 2021. Although there were seemingly numerous explosion misclassifications on the independent seismic and infrasound model timelines, our amalgamated explosion timeline (Figure 8h) accurately reflects the lack of real explosions during this time period. Fifteen satellite SO_2 detections from the TROPOspheric Monitoring Instrument (TROPOMI) sensor also occurred during this pre-eruptive unrest period between 18 May and 20 August 2021 (Figure 8e). SO_2 emission rates during this period ranged from <100 tonnes/day (t/d) to a maximum of 330 t/d on 19 August 2021. Retrospective analyses using a machine learning based hotspot detection algorithm applied to Visible Infrared Imaging Radiometer Suite (VIIRS) satellite data (Saunders-Shultz et al., 2024) also revealed subtle thermal anomalies during this time period (black dots in Figure 8e). Radiative power values associated with high confidence (nighttime only, with probabilities exceeding 0.75) detections mostly stayed well under 1 MW this time period, but showed an increasing rate of detections from May through August, with the first value exceeding 1 MW on 12 August 2021. AVO remote sensing duty analysts also reported weakly elevated

surface temperatures between June and August 2021. In general, the high number of satellite SO_2 detections coincided with the increase in tremor diversity that was identified by our VOISS-Net analysis, followed by increased observations of elevated surface temperatures. Together, these derivative indicators of unrest indicate that Pavlof Volcano was likely transitioning from unrest to eruption between May 2021 and August 2021. Although the source depth of our cataloged tremor is not constrained, the increasing amplitude and broadband character of the tremor signals obtained from VOISS-Net and our D_R calculations may be inferred as active fluid (i.e., magma, supercritical fluids, and/or gas) migration within the shallow magmatic system. This hypothesis is also supported by the increasing amount of infrasonic tremor identified by VOISS-Net and the increasing number of satellite SO_2 and hotspot detections.

Shifting into the active eruption period, the VOISS-Net seismic model identifies a period of steady and energetic broadband tremor that dominated the eruptive period from October 2021 to December 2022 (Figure 8c). The onset of explosive activity in August 2021 (which coincides with AVO's transition to Aviation Color Code and Volcano Alert Level ORANGE/WATCH) and its progression until December 2021 was also evident in the VOISS-Net explosion catalog (Figure 8h). In particular, when both seismic and infrasound VOISS-Net explosion classifications are integrated, periods of misclassification (arising from long-period earthquake swarms and infrasound noise spikes) are filtered out, and explosion-rich sequences, such as those in February, July, and October 2022, are highlighted. Reduced displacement values track the amplitude of seismic broadband tremor well, although values are occasionally contaminated by storm noise which cause amplitude peaks as observed in late November 2021 (Figure 8d). The reduced displacement values peaked at approximately 5 cm^2 during the 2021–2022 eruption, which is higher than the peak value of the 2007 eruption (2.5 cm^2, Haney et al., 2014), but lower than the peak values of the 2013 eruption (8 cm^2), the 2014 eruption (8 cm^2), and the 2016 eruption (17 cm^2, Haney et al., 2018). Twelve SO_2 detections were made during the August 2021 to November 2022 phase of activity with emission rates ranging from ~100 to 296 t/d. Neither the SO_2 detection frequency nor emission rates showed clear relations to the observed tremor types or amplitudes, nor with the eruptive state of the volcano. This is likely because these relatively low Pavlof Volcano emission rates are similar to the TROPOMI detection limit under optimal conditions (Theys et al., 2019). Ultraviolet satellite-based SO_2 monitoring is also diminished during the winter due to limited ultraviolet radiation at the relatively high latitude of Pavlof Volcano (Lopez et al., 2020), limiting satellite SO_2 detection capabilities. However, the lack of persistent or high quantity SO_2 emission rates, especially during the first month of the eruption when UV levels are still high, likely indicates that the lava erupted during parts of the 2021–2022 eruption was probably degassed, or that the amounts of magma erupted are too small to result in notable changes in the detected SO_2 emission rates—at least at the temporal resolution of this data set (maximum of one satellite detection per day). On the contrary, the radiative power values correlate well with observed periods of lava production (January–February 2022 and May–June 2022), where anomalously high values >50 MW were recorded. These time periods of high thermal output are inversely correlated with periods of explosive activity (November 2021, February–March 2022, July–August 2022) as determined by the VOISS-Net analysis. This can be interpreted as transitions between volcanic eruption regimes: from an effusion-dominated, lava fountaining regime to a more explosive regime characterized by small ash clouds and tephra fallout. As the eruption waned in December 2022, the VOISS-Net seismic model accurately characterizes the declining prominence of broadband seismic tremor and transitions to narrowband harmonic and monochromatic tremor types. The noise class also becomes more prevalent, and eventually becomes dominant, indicating that the volcano had returned to near-background seismic levels. By December 2022, the reduced displacement values also decline to pre-eruptive levels of about 0.5 cm^2. Remote sensing observations of ash clouds and elevated surface temperatures are absent during all of December 2022. The infrasonic timeline shows little to no infrasonic tremor, further indicating a return to quiescence. The VOISS-Net explosion timeline correctly indicates an absence of explosions in December 2022.

4.3. Comparing VOISS-Net Classifications With Eruption Observations

Here we further evaluate explosive and effusion periods and their associated seismic signals and select two 6-hr seismic sequences with clear web-camera views for comparison. Figures 9 and 10 show two 6-hr seismic sequences from 4 December 2021 and 28 May 2022, respectively. Each figure follows the VOISS-Net plotting fashion described in Figure 6, but both plots have been modified to include web-camera and satellite views. The corresponding infrasound VOISS-Net analyses are omitted here because they are dominated by wind and electronic noise in those time periods. Although explosions in the 4 December 2021 sequence were validated by

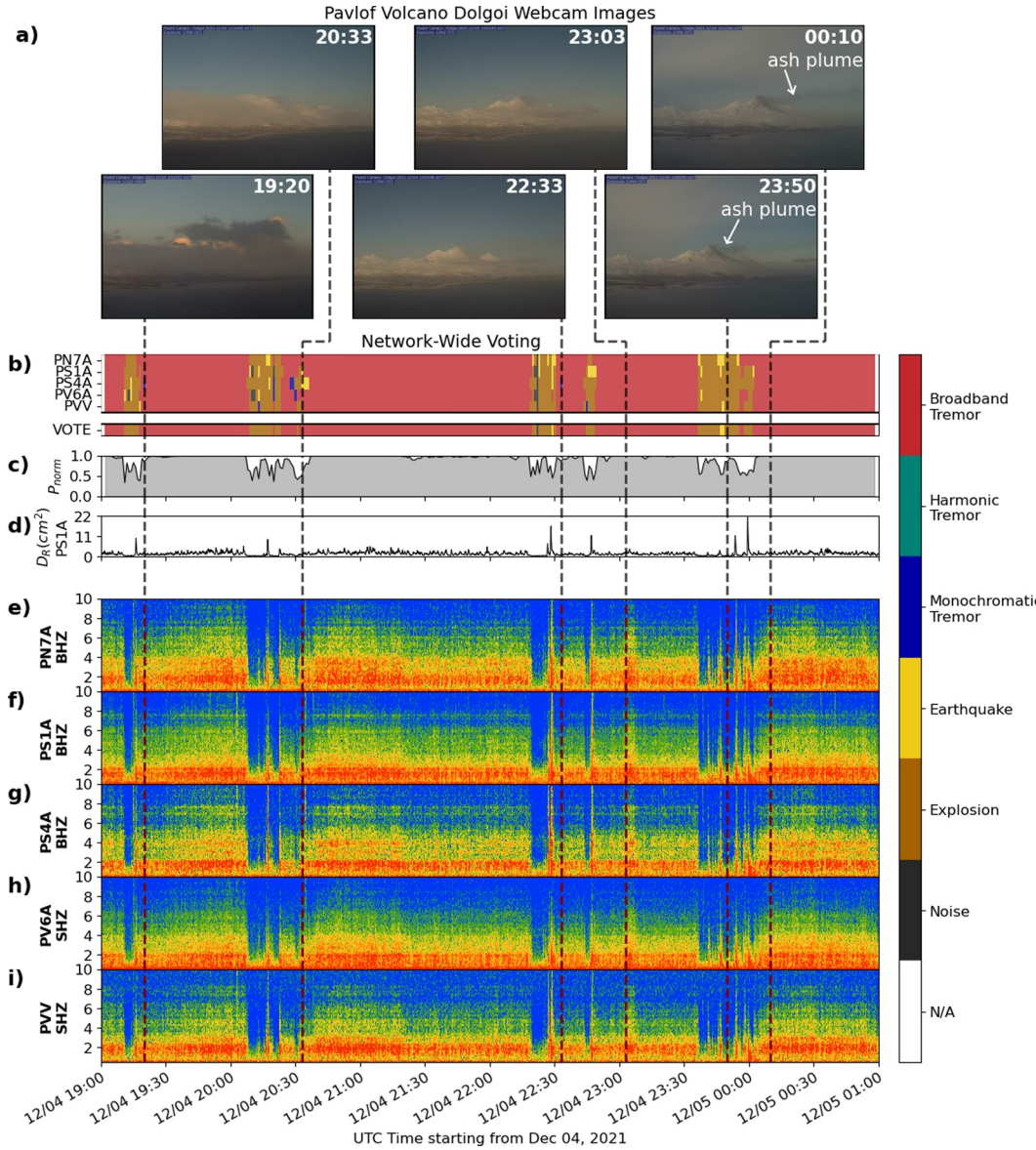

Figure 9. VOISS-Net seismic classifications for a 6-hr sequence from 4 December 2021, 19:00 UTC, to 5 December 2021, 01:00 UTC, compared to selected clear views from an Alaska Volcano Observatory web camera on Dolgoi Island, about 30 km south of Pavlof Volcano (55.149621, −161.8647; 453 m above sea level). The width of the image is about 20 km. The view direction from the Dolgoi Island web camera is to the north-northeast. The spectrograms and VOISS-Net results show numerous explosions and persistent broadband tremor. Web-camera images of ash clouds are broadly correlative with the explosions indicated on the seismic timeline.

infrasound classifications, no infrasonic tremor was detected in either time period that can contribute to the discussion of unrest regimes.

Figure 9 shows an interesting explosion sequence nested within a period of persistent broadband tremor on 4 December 2021. The seismic broadband tremor exhibited by Pavlof Volcano sometimes pauses for minutes to hours prior to explosive activity. This pattern of pauses preceding explosion signals in tremor is clearly depicted in the sequence in Figure 9, where VOISS-Net switches between the broadband tremor class and explosion class over time scales ranging from tens of minutes to hours. The web-camera on Dolgoi Island ~30 km south of Pavlof Volcano captured several clear images of the volcano during this time period. However, as images are only transmitted once every 30 min, the web-camera was unable to capture the onset of each explosion signal. Nevertheless, it is apparent that most or all clear web-camera images immediately succeeding a VOISS-Net explosion class show a weak plume either growing or dissipating. This is clearest on the 23:50 web-camera

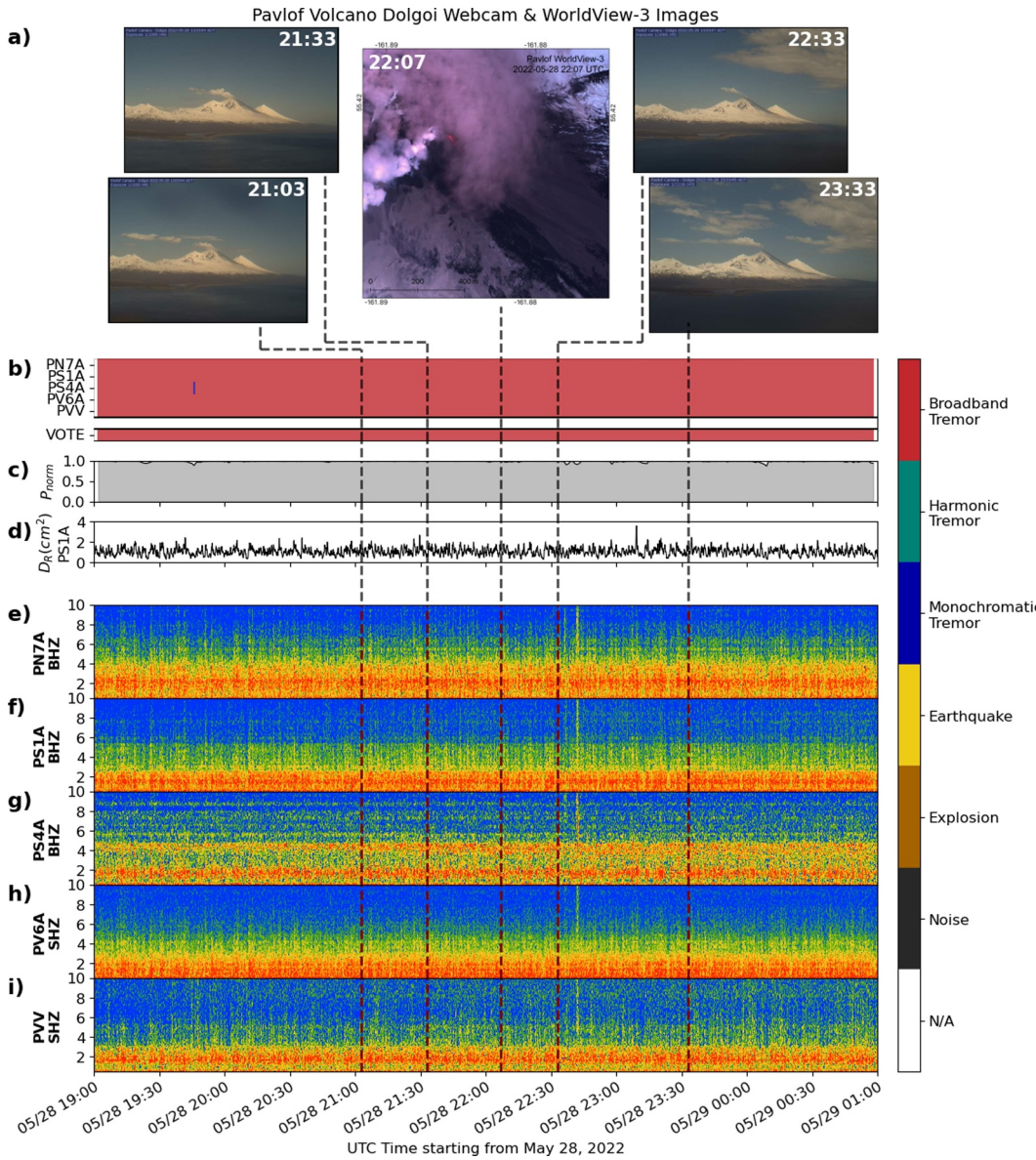

Figure 10. VOISS-Net seismic classifications for a 6-hr sequence from 28 May 2022, 19:00 UTC, to 29 May 2022, 01:00 UTC, compared with selected clear views from the Dolgoi Island web-camera and a coincident WorldView-3 satellite image (© 2023, Maxar, USG Plus). The timeline demonstrates a continuous sequence of broadband tremor. The web-camera views show steam emissions and the red area in satellite image is a fountain-fed lava flow.

image in Figure 9a, which shows an ash plume at a time where many explosion signals are observed between 23:40 and 00:05.

In contrast, Figure 10 shows a seemingly uneventful seismic sequence of continuous broadband seismic tremor from 28 May 2022. Lava fountaining was occurring during this time period, and vapor plumes were emanating from the vent area. The vapor plumes were likely the result of lava-snow interaction near the vent or were derived from degassing magma, lava flows, or both. A WorldView-3 image acquired at 22:07 (Figure 10a) shows bright red linear features that are channelized lava flows branching close to the vent. Web-camera images taken around this time period illustrate the coincidence of lava fountaining and vapor emissions. This sequence highlights how volcanic activity at Pavlof Volcano might not be fully represented by seismic data streams alone. Although our VOISS-Net output successfully identifies explosion sequences (as shown in Figure 9), it fails to highlight fluctuations in the intensity of lava fountaining, be it in terms of tremor type or in D_R values. Episodic lava fountaining is commonly observed during historical eruptions of Pavlof Volcano, and this leads to the formation of clastogenic

lava flows which are characteristic of many of its past eruptions (Waythomas et al., 2014, 2017). Lava fountaining results in the accumulation of spatter, which forms unstable mounds that eventually collapse under gravity to form hot granular mass flows that are capable of melting significant quantities of snow and ice. This process generates lahars that in the past have traveled as far as the Bering Sea (40 km) on the north side of the volcano.

In summary, VOISS-Net is successful in deriving machine learning-based tremor catalogs that track seismic and infrasonic signatures of unrest during the 2021–2022 Pavlof Volcano eruption. The network vote timelines derived from the VOISS-Net seismic and infrasound models provide a reliable and comprehensive description of the seismoacoustic signals produced by Pavlof Volcano over time, and when combined, provides a high confidence explosion timeline that compares well with the explosion catalogs independently derived using RTM (Fee et al., 2021). When compared with multidisciplinary metrics and observations as shown in Figure 8, the VOISS-Net timelines allow for an easy visual comparison which permits inferences on volcanic activity, such as periods dominated by explosions and periods dominated by effusion which show stable broadband tremor and elevated thermal output.

4.4. Application to Pavlof Volcano Eruptions in 2007, 2013, 2014, and 2016

The good performance of VOISS-Net on the 2021–2022 eruption motivates us to explore the applicability of our model to previous historical eruptions of Pavlof Volcano where suitable data exist. Pavlof Volcano's past eruptions investigated here began in August 2007, May 2013, May 2014, November 2014, and March 2016. The May 2014 eruption is omitted from our analysis because the seismic data contain significant electronic noise at stations PN7A and PV6, and a major seismic data outage occurred on all stations except PN7A during most of the eruptive period. These previous eruptions predate the installation of broadband seismic instruments at PN7A, PS1A, and PS4A, as well as the colocated infrasound sensors at each site. In addition, station PV6A did not exist prior to July 2016—in its place on the northwest side of the volcano was station PV6, which was closer to the volcano (4.5 km from the volcano's summit, Figure 1). PV6 was destroyed by pyroclastic flows during the 2016 eruption. For the eruptions in 2007, 2013, 2014, and 2016, we implement the VOISS-Net seismic model independently on stations that provided high quality data during the time of each eruption, over an analysis window from the eruption onset until background activity was reestablished. Specifically, the analysis of the 2007 eruption utilized stations PN7A, PS1A, PV6, and PVV, while the analyses of the 2013, 2014, and 2016 eruptions utilized stations PS1A, PS4A, and PVV. We summarize each eruption's seismic timeline analysis using the same network-wide voting procedure implemented for the 2021–2022 eruption.

Figure 11 shows the VOISS-Net seismic timelines for the selected past Pavlof Volcano eruptions alongside the AVO Aviation Color Codes assigned to the volcano in response to escalating unrest or observations of eruptive activity. Times are aligned by eruption onset. Across the four eruptions, our VOISS-Net seismic model successfully describes the tremor sequences of each eruption, and highlights the relative prevalence of each seismic class over time. Notably, it reveals periods of broadband tremor, transitions into quiescence or different seismic tremor regimes, and the duration of each tremor regime. Unlike the 2021–2022 eruption, the past eruptions of Pavlof Volcano shown in Figure 11 showed little to no pre-eruption unrest. Other than the 2007 eruption, which spent one day at Aviation Color Code and Volcano Alert Level YELLOW/ADVISORY from precursory seismicity, the eruptions in 2013, 2014, and 2016 began with no recognized precursory unrest and were first detected after eruptions were underway and observed by passing pilots or AVO duty remote sensing scientists. Each of these eruptions began with, or were closely followed by, the rapid onset of broadband tremor that was not apparent prior to the recognized eruptive activity. This is apparent in Figures 11a–11d as swaths of high broadband tremor occurrence rates occur almost immediately after eruption onset (denoted by the black dashed line). The 2013 eruption represents a slight departure from this behavior; AVO raised the Aviation Color Code and Volcano Alert Level from GREEN/NORMAL to ORANGE/WATCH after detecting a strongly elevated thermal anomaly at the summit of the volcano, which prompted seismologists to scrutinize seismic data and identify weak tremor signals. Figure 11 also shows that the seismic unrest associated with the different eruptions can exhibit a broad range of timescales; broadband tremor lasted for more than a year in the 2021–2022 eruption, for weeks during the 2007 eruption, but only a few days at a time in the 2016 eruption.

We also see that the 2007 eruption (Figure 11a) shows a seismic unrest profile that is similar to seismic unrest profile of the 2021–2022 eruption (Figure 8c), albeit over different timescales. Both eruptions start with a hybrid of mostly broadband and monochromatic tremor, followed by a longer period of nearly continuous broadband

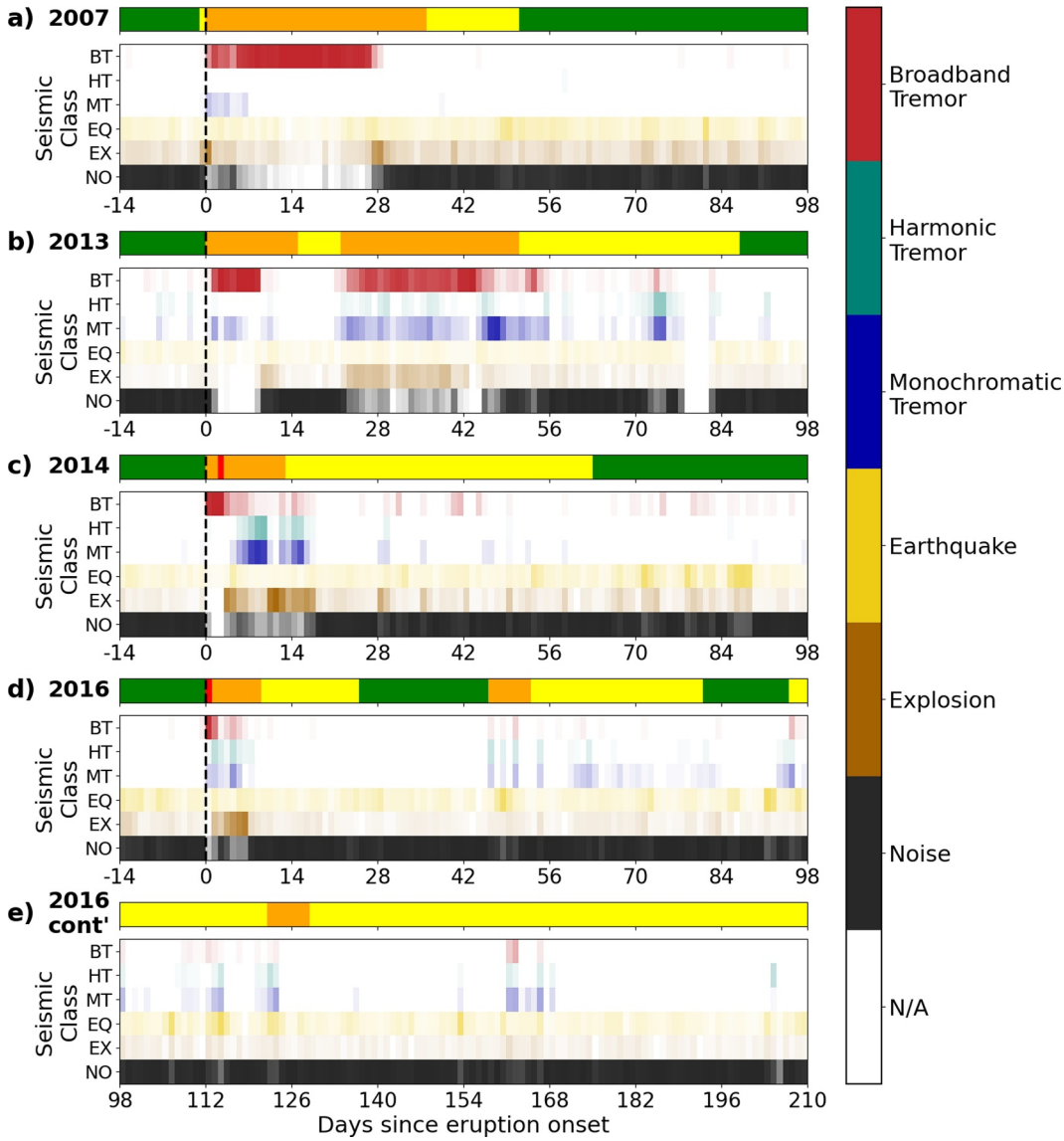

Figure 11. Summary of VOISS-Net seismic model results for past Pavlof Volcano eruptions. (a) August 2007, (b) May 2013, (c) November 2014, and (d) March 2016. Panel (e) is a continuation of the March 2016 eruption shown in (d), due to the longer duration of that eruption. For each eruption, we present Alaska Volcano Observatory's (AVO's) Aviation Color Codes for Pavlof Volcano alongside their VOISS-Net seismic timelines, which are binned by day similar to Figure 8c. Seismic class names are abbreviated for presentation purposes. Note that all VOISS-Net timelines are obtained using the seismic model trained on data from the Pavlof Volcano 2021–2022 eruption. Time axes are scaled similarly for all eruptions, and are aligned by eruption onset (black dashed line), which we take as AVO's Aviation Color Code change to ORANGE or RED for each eruption. Each plotted sequence starts 2 weeks prior to the eruption onset in order to illustrate the increase in seismic unrest from background.

tremor. Both eruptions also wane and conclude relatively quickly, without showing any resurgence in tremor activity. In contrast, the 2013, 2014, and 2016 eruptions show more complex behavior, in which tremor diversity is prominent at different times throughout their active eruption periods. For example, both the 2013 and 2016 eruptions show resurgence of tremor after days to weeks of quiescence. Both the 2013 and 2014 eruptions also show a period of monochromatic tremor that is mostly independent of broadband tremor (see Figure 11b Day 45, and Figure 11c Day 8), which is not observed during the 2007 and 2016 eruptions. Interestingly, the 2013, 2014, and 2016 eruptions, which show complex tremor activity, share a vent location on the summit crater, while the 2007 and 2021–2022 eruptions, which show predominantly broadband tremor, share a vent location on the southeast flank. These observations demonstrate that Pavlof Volcano eruptions do not follow a straightforward or predictable eruptive behavior but rather seismic signatures can vary from eruption to eruption, or from vent to vent.

VOISS-Net's ability to detect broadband tremor with a high degree of certainty allows it to help characterize the onset of Pavlof Volcano's eruptive activity. Although not shown in the day-long timescales on Figure 11, our model can determine the onset of broadband tremor down to the nearest minute (e.g., Figure 6). VOISS-Net's capability to help detect eruption onsets is exemplified in the analysis of the seismic unrest profile of the 2016 eruption, which had no obvious precursory seismicity (Fee, Haney, et al., 2017). Figure S6 in Supporting Information S1 shows the VOISS-Net seismic timeline in the 3 hr bounding the start of the 2016 eruption. VOISS-Net accurately indicates the onset of broadband tremor at 23:52 in the network vote, and even highlights a very subtle instance of precursory tremor on PVV at 23:43 that was not observed at other stations. Retrospective analysis by Fee, Haney, et al. (2017) suggested a tremor onset time of 23:55, 3 min later than the time determined by the VOISS-Net network classifications. We note that VOISS-Net did not find any earlier precursory seismic signal prior to this subtle instance of tremor (Figure 11d), which disagrees with an independent analysis using normalized Displacement Seismic Amplitude Ratio (nDSAR) that suggested otherwise (Ardid et al., 2022).

4.5. Implications and Potential Model Improvements

We have demonstrated that VOISS-Net is able to summarize seismoacoustic timelines at Pavlof Volcano over various timescales, and among different eruptions in a consistent and comparable manner. Recall that VOISS-Net can be run both retroactively, and also automatically in near real-time. The ability to automatically identify and characterize tremor (and hence, help constrain eruption onset) in a timely fashion would be a significant monitoring advancement for AVO since current checks involve analysts scanning spectrograms manually every 12 hr, or in shorter intervals during eruptions. VOISS-Net's utility in distilling large quantities of data into a classification timeline can reduce the burden of seismic analysts while providing time-sensitive information on volcanic behavior including tremor and explosions. The near real-time identification of tremor also presents new avenues for integrating tremor-related metrics in physics-based data assimilation strategies in volcano monitoring (e.g., Albright et al., 2019; Gregg et al., 2022).

We acknowledge that our VOISS-Net models have room for improvement. The incorporation of a dedicated long-period seismicity class in the seismic model could mitigate explosion misclassifications, and the expansion of infrasonic tremor types in the infrasound model would provide additional utility, especially if it were to be applied to a different volcano with a different infrasonic signature. Infrasonic tremor from Pavlof Volcano recorded on the local network stations has thus far been weak with limited signal-to-noise, so training VOISS-Net with higher signal-to-noise spectrograms would likely improve the model and its utility. Expanding the labeled data set for the sparse classes, such as harmonic tremor and explosions, would also contribute to a more robust validation and test data set that would improve model evaluation. Adding labeled spectrograms from different volcanoes, and from data recorded on more proximal or distal stations, could also enable future iterations of the model to generalize to different volcanic settings and network geometries. A more rigorous test on hyperparameters, such as spectrogram segment lengths and frequency limits, could also be conducted to further evaluate related sensitivities and optimize VOISS-Net. As with most machine learning models, a more extensive and comprehensive labeled data set will improve the model and open up the possibility of adding model complexity should it be required. For example, the current VOISS-Net models are limited to selecting one class per time step and handles superimposed classes by prioritizing the transient classes. A larger labeled data set could test image segmentation strategies, which can identify different classes occurring at different frequencies within the same spectrogram segment at the same time. Further iterations of VOISS-Net that utilize three component seismic data (as it becomes more widespread at volcanoes) or polarization spectrograms (Haney et al., 2020) could further improve model performance.

5. Conclusion

In this study, we successfully trained and demonstrated the proficiency of VOISS-Net, a pair of seismic and infrasound CNNs that offer a fast, robust and automated method of classifying volcano seismic and infrasound data to derive high resolution seismoacoustic timelines at Pavlof Volcano. These seismoacoustic timelines can be combined with multidisciplinary observations to infer and interpret source processes driving the various volcano monitoring signatures produced during volcanic unrest and eruption. The VOISS-Net models are applicable to volcano monitoring and geophysical research and are suitable for the construction of first-order tremor catalogs.

We apply VOISS-net to eruptions of Pavlof Volcano that occurred in 2007, 2013, 2014, 2016, and 2021–2022, and use it to identify volcanic unrest (as manifested in seismic and infrasound data) and shifts in unrest and

eruption regimes. Using VOISS-Net, we derive a machine learning based tremor catalog from 2 years of seismic and infrasound data associated with the 2021–2022 eruption of Pavlof Volcano. Network-wide voting with this data set helps to more confidently classify the seismic and infrasound signals produced during eruptions of the volcano. The VOISS-Net-derived timelines identify periods of substantial tremor diversity prior to and after months of stable broadband tremor, and episodes of explosive activity nested within the prolonged broadband tremor period. We additionally compare seismic spectrograms and their corresponding VOISS-Net timelines with web-camera and satellite images of explosions and lava fountaining. This comparison reveals that although explosions can be uniquely identified in geophysical data, the same cannot be said about the intensity of lava fountaining and steaming. In addition, the VOISS-Net seismic model generalizes well when applied on past Pavlof Volcano eruptions (2007, 2013, 2014, and 2016) despite there being changes in instrumentation. In our retrospective analyses, we successfully summarize the tremor sequences in each of the past eruptions in a consistent manner, and identify the similarities and differences of the seismic signatures across the four eruptive periods. Although Pavlof Volcano produces a characteristic tremor signal observed during its recent eruptions, the volcano does not always follow the same patterns of unrest or eruptive behavior.

Although VOISS-Net was trained and applied to Pavlof Volcano, we have preliminarily applied it with success on other volcanoes in Alaska. However, seismic and infrasound data with characteristics that exist outside our labeled data set are generally poorly classified by VOISS-Net, so caution should be exercised in interpreting the classifications in the model's current iteration. Future work will involve expanding our training data set to include additional signal variability, incorporating a specific long-period earthquake class to the seismic model, adjusting spectrogram parameters to better capture short-duration signals, and generalizing the model to eruptions from other volcanoes. We envision that VOISS-Net, or future improved iterations of it, will be used as an operational tool for AVO.

Data Availability Statement

The facilities of EarthScope Consortium were used for access to waveforms and related metadata used in this study. Waveform data for all Pavlof Volcano eruptions were downloaded through the EarthScope Consortium Web Services (https://service.iris.edu/) from AV network (Alaska Volcano Observatory/USGS, 1988). Volcano color codes and corresponding observations of volcanic activity were made by the Alaska Volcano Observatory (https://avo.alaska.edu/). Relevant code, VOISS-Net models, and labeled data sets can be accessed via our archived GitHub repository release on Zenodo (Tan & Fee, 2024).

Acknowledgments

We thank Zahra Zali, the Associate Editor, and an anonymous reviewer for their thorough reviews and insightful comments that improved this work. We thank Pablo Saunders-Shultz for providing the radiative power data set for the Pavlof Volcano 2021–2022 eruption, and all PREEVENTS collaborators and AVO for insightful discussions and helpful suggestions. This work was funded by the NSF PREEVENTS Grant 1855126. This material is based upon work supported by the U.S. Geological Survey under Cooperative Agreement No. G21AC10384. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. These services are funded through the Seismological Facility for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Agreement EAR-1724509.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous distributed systems. CoRR, abs/1603, 04467.

Aki, K., & Koyanagi, R. (1981). Deep volcanic tremor and magma ascent mechanism under Kilauea, Hawaii. *Journal of Geophysical Research*, *86*(B8), 7095–7109. https://doi.org/10.1029/jb086ib08p07095

Alaska Volcano Observatory/USGS. (1988). Alaska volcano observatory. *International Federation of Digital Seismograph Networks*. https://doi.org/10.7914/SN/AV

Albright, J. A., Gregg, P. M., Lu, Z., & Freymueller, J. T. (2019). Hindcasting magma reservoir stability preceding the 2008 eruption of Okmok, Alaska. *Geophysical Research Letters*, *46*(15), 8801–8808. https://doi.org/10.1029/2019gl083395

Ardid, A., Dempsey, D., Caudron, C., & Cronin, S. (2022). Seismic precursors to the Whakaari 2019 phreatic eruption are transferable to other eruptions and volcanoes. *Nature Communications*, *13*(1), 2002. https://doi.org/10.1038/s41467-022-29681-y

Barrière, J., Oth, A., d'Oreye, N., Subira, J., Smittarello, D., Brenot, H., et al. (2023). Local infrasound monitoring of lava eruptions at Nyiragongo Volcano (DR Congo) using urban and near-source stations. *Geophysical Research Letters*, *50*(18), e2023GL104664. https://doi.org/10.1029/2023gl104664

Beyreuther, M., Barsch, R., Krischer, L., Megies, T., Behr, Y., & Wassermann, J. (2010). ObsPy: A Python toolbox for seismology. *Seismological Research Letters*, *81*(3), 530–533. https://doi.org/10.1785/gssrl.81.3.530

Bishop, J. W., Fee, D., & Szuberla, C. A. L. (2020). Improved infrasound array processing with robust estimators. *Geophysical Journal International*, *221*(3), 2058–2074. https://doi.org/10.1093/gji/ggaa110

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, *106*, 249–259. https://doi.org/10.1016/j.neunet.2018.07.011

Cameron, C. E., Prejean, S. G., Coombs, M. L., Wallace, K. L., Power, J. A., & Roman, D. C. (2018). Alaska volcano observatory alert and forecasting timeliness: 1989–2017. *Frontiers in Earth Science*, *6*, 86. https://doi.org/10.3389/feart.2018.00086

Cannata, A., Di Grazia, G., Aliotta, M., Cassisi, C., Montalto, P., & Patanè, D. (2013). Monitoring seismo-volcanic and infrasonic signals at volcanoes: Mt. Etna case study. *Pure and Applied Geophysics*, *170*(11), 1751–1771. https://doi.org/10.1007/s00024-012-0634-x

Chouet, B. A. (1985). Excitation of a buried magmatic pipe: A seismic source model for volcanic tremor. *Journal of Geophysical Research*, *90*(B2), 1881–1893. https://doi.org/10.1029/jb090ib02p01881

Chouet, B. A. (1996). Long-period volcano seismicity: Its source and use in eruption forecasting. *Nature*, *380*(6572), 309–316. https://doi.org/10.1038/380309a0

Chouet, B. A., Koyanagi, R. Y., & Aki, K. (1987). Origin of volcanic tremor in Hawaii. *U. S. Geological Survey Professional Paper*, *1350*, 1259–1280.

Chouet, B. A., & Matoza, R. S. (2013). A multi-decadal view of seismic methods for detecting precursors of magma movement and eruption. *Journal of Volcanology and Geothermal Research*, *252*, 108–175. https://doi.org/10.1016/j.jvolgeores.2012.11.013

Dmitrieva, K., Hotovec-Ellis, A. J., Prejean, S., & Dunham, E. M. (2013). Frictional-faulting model for harmonic tremor before Redoubt Volcano eruptions. *Nature Geoscience*, *6*(8), 652–656. https://doi.org/10.1038/ngeo1879

Duque, A., Gonzalez, K., Perez, N., Benitez, D., Grijalva, F., Lara-Cueva, R., & Ruiz, M. (2020). Exploring the unsupervised classification of seismic events of Cotopaxi volcano. *Journal of Volcanology and Geothermal Research*, *403*, 107009. https://doi.org/10.1016/j.jvolgeores.2020.107009

Endo, E. T., & Murray, T. (1991). Real-time seismic amplitude measurement (RSAM): A volcano monitoring and prediction tool. *Bulletin of Volcanology*, *53*(7), 533–545. https://doi.org/10.1007/bf00298154

Fee, D., Garcés, M., Patrick, M., Chouet, B., Dawson, P., & Swanson, D. (2010). Infrasonic harmonic tremor and degassing bursts from Halema'uma'u Crater, Kilauea Volcano, Hawaii. *Journal of Geophysical Research*, *115*(B11), B11316. https://doi.org/10.1029/2010jb007642

Fee, D., Haney, M., Matoza, R., Szuberla, C., Lyons, J., & Waythomas, C. (2016). Seismic envelope-based detection and location of ground-coupled airwaves from volcanoes in Alaska. *Bulletin of the Seismological Society of America*, *106*(3), 1024–1035. https://doi.org/10.1785/0120150244

Fee, D., Haney, M. M., Matoza, R. S., Van Eaton, A. R., Cervelli, P., Schneider, D. J., & Iezzi, A. M. (2017). Volcanic tremor and plume height hysteresis from Pavlof Volcano, Alaska. *Science*, *355*(6320), 45–48. https://doi.org/10.1126/science.aah6108

Fee, D., Izbekov, P., Kim, K., Yokoo, A., Lopez, T., Prata, F., et al. (2017). Eruption mass estimation using infrasound waveform inversion and ash and gas measurements: Evaluation at Sakurajima Volcano, Japan. *Earth and Planetary Science Letters*, *480*, 42–52. https://doi.org/10.1016/j.epsl.2017.09.043

Fee, D., & Matoza, R. S. (2013). An overview of volcano infrasound: From Hawaiian to plinian, local to global. *Journal of Volcanology and Geothermal Research*, *249*, 123–139. https://doi.org/10.1016/j.jvolgeores.2012.09.002

Fee, D., Toney, L., Kim, K., Sanderson, R. W., Iezzi, A. M., Matoza, R. S., et al. (2021). Local explosion detection and infrasound localization by reverse time migration using 3-D finite-difference wave propagation. *Frontiers in Earth Science*, *9*, 44. https://doi.org/10.3389/feart.2021.620813

Ferreira, A., Curilem, M., Gomez, W., & Rios, R. (2023). Deep learning and multi-station classification of volcano-seismic events of the Nevados del Chillán volcanic complex (Chile). *Neural Computing & Applications*, *35*(35), 1–18. https://doi.org/10.1007/s00521-023-08994-z

Gestrich, J. E., Fee, D., Tsai, V. C., Haney, M. M., & Van Eaton, A. R. (2020). A physical model for volcanic eruption tremor. *Journal of Geophysical Research: Solid Earth*, *125*(10), 1–28. https://doi.org/10.1029/2019JB018980

Girona, T., Caudron, C., & Huber, C. (2019). Origin of shallow volcanic tremor: The dynamics of gas pockets trapped beneath thin permeable media. *Journal of Geophysical Research: Solid Earth*, *124*(5), 4831–4861. https://doi.org/10.1029/2019JB017482

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Gregg, P. M., Zhan, Y., Amelung, F., Geist, D., Mothes, P., Koric, S., & Yunjun, Z. (2022). Forecasting mechanical failure and the 26 June 2018 eruption of Sierra Negra volcano, Galápagos, Ecuador. *Science Advances*, *8*(22), eabm4261. https://doi.org/10.1126/sciadv.abm4261

Haney, M. M. (2014). Backprojection of volcanic tremor. *Geophysical Research Letters*, *41*(6), 1923–1928. https://doi.org/10.1002/2013gl058836

Haney, M. M., Fee, D., McKee, K. F., Lyons, J. J., Matoza, R. S., Wech, A. G., et al. (2020). Co-eruptive tremor from Bogoslof volcano: Seismic wavefield composition at regional distances. *Bulletin of Volcanology*, *82*(2), 18. https://doi.org/10.1007/s00445-019-1347-0

Haney, M. M., Hotovec-Ellis, A. J., Bennington, N. L., De Angelis, S., & Thurber, C. (2014). Tracking changes in volcanic systems with seismic interferometry.

Haney, M. M., Matoza, R. S., Fee, D., & Aldridge, D. F. (2018). Seismic equivalents of volcanic jet scaling laws and multipoles in acoustics. *Geophysical Journal International*, *213*(1), 623–636. https://doi.org/10.1093/gji/ggx554

Hershey, S., Chaudhuri, S., Ellis, D. P., Gemmeke, J. F., Jansen, A., Moore, R. C., et al. (2017). CNN architectures for large-scale audio classification. In *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 131–135). IEEE.

Hotovec, A. J., Prejean, S. G., Vidale, J. E., & Gomberg, J. (2013). Strongly gliding harmonic tremor during the 2009 eruption of Redoubt Volcano. *Journal of Volcanology and Geothermal Research*, *259*, 89–99. https://doi.org/10.1016/j.jvolgeores.2012.01.001

Ichihara, M. (2016). Seismic and infrasonic eruption tremors and their relation to magma discharge rate: A case study for sub-Plinian events in the 2011 eruption of Shinmoe-dake, Japan. *Journal of Geophysical Research: Solid Earth*, *121*(10), 7101–7118. https://doi.org/10.1002/2016jb013246

Jellinek, A. M., & Bercovici, D. (2011). Seismic tremors and magma wagging during explosive volcanism. *Nature*, *470*(7335), 522–525. https://doi.org/10.1038/nature09828

Johnson, J. B., & Palma, J. L. (2015). Lahar infrasound associated with Volcán Villarrica's 3 March 2015 eruption. *Geophysical Research Letters*, *42*(15), 6324–6331. https://doi.org/10.1002/2015gl065024

Journeau, C., Shapiro, N. M., Seydoux, L., Soubestre, J., Koulakov, I. Y., Jakovlev, A. V., et al. (2022). Seismic tremor reveals active trans-crustal magmatic system beneath Kamchatka volcanoes. *Science Advances*, *8*(5), eabj1571. https://doi.org/10.1126/sciadv.abj1571

Julian, B. R. (1994). Volcanic tremor: Nonlinear excitation by fluid flow. *Journal of Geophysical Research*, *99*(B6), 11859–11877. https://doi.org/10.1029/93jb03129

Khan, S., Rahmani, H., Shah, S. A. A., Bennamoun, M., Medioni, G., & Dickinson, S. (2018). A guide to convolutional neural networks for computer vision. (Vol. *8*, pp. 1–207). Springer. https://doi.org/10.2200/s00822ed1v01y201712cov015

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kong, Q., Wang, R., Walter, W. R., Pyle, M., Koper, K., & Schmandt, B. (2022). Combining deep learning with physics based features in explosion-earthquake discrimination. *Geophysical Research Letters*, *49*(13), e2022GL098645. https://doi.org/10.1029/2022gl098645

Konstantinou, K. I., & Schlindwein, V. (2003). Nature, wavefield properties and source mechanism of volcanic tremor: A review. *Journal of Volcanology and Geothermal Research*, *119*(1–4), 161–187. https://doi.org/10.1016/s0377-0273(02)00311-6

Koyanagi, R. Y., Chouet, B. A., & Aki, K. (1987). Origin of volcanic tremor in Hawaii. *U. S. Geological Survey Professional Paper*, *1350*(2), 1221–1257.

Kumagai, H., Mothes, P., Ruiz, M., & Maeda, Y. (2015). An approach to source characterization of tremor signals associated with eruptions and lahars. *Earth Planets and Space*, *67*, 1–12. https://doi.org/10.1186/s40623-015-0349-1

Langer, H., Falsaperla, S., Messina, A., Spampinato, S., & Behncke, B. (2011). Detecting imminent eruptive activity at Mt Etna, Italy, in 2007–2008 through pattern classification of volcanic tremor data. *Journal of Volcanology and Geothermal Research*, *200*(1–2), 1–17. https://doi.org/10.1016/j.jvolgeores.2010.11.019

Lapins, S., Goitom, B., Kendall, J.-M., Werner, M. J., Cashman, K. V., & Hammond, J. O. (2021). A little data goes a long way: Automating seismic phase arrival picking at Nabro volcano with transfer learning. *Journal of Geophysical Research: Solid Earth*, *126*(7), e2021JB021910. https://doi.org/10.1029/2021jb021910

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Leet, R. C. (1988). Saturated and subcooled hydrothermal boiling in groundwater flow channels as a source of harmonic tremor. *Journal of Geophysical Research*, *93*(B5), 4835–4849. https://doi.org/10.1029/jb093ib05p04835

Linville, L., Pankow, K., & Draelos, T. (2019). Deep learning models augment analyst decisions for event discrimination. *Geophysical Research Letters*, *46*(7), 3643–3651. https://doi.org/10.1029/2018gl081119

Liu, C., Lay, T., & Xiong, X. (2022). The 29 July 2021 MW 8.2 Chignik, Alaska Peninsula earthquake rupture inferred from seismic and geodetic observations: Re-rupture of the western 2/3 of the 1938 rupture zone. *Geophysical Research Letters*, *49*(4), e2021GL096004. https://doi.org/10.1029/2021gl096004

Lopez, T., Brewster, I., Kushner, S., Wasser, V. K., Saunders-Shultz, P., Randall, M., et al. (2021). Using TROPOMI satellite SO_2 observations to inform near-real-time volcano monitoring and eruption warnings by the Alaska Volcano Observatory. In *AGU fall meeting abstracts* (Vol. 2021, p. V33C-10).

Lopez, T., Clarisse, L., Schwaiger, H., Van Eaton, A., Loewen, M., Fee, D., et al. (2020). Constraints on eruption processes and event masses for the 2016–2017 eruption of Bogoslof volcano, Alaska, through evaluation of IASI satellite SO_2 masses and complementary datasets. *Bulletin of Volcanology*, *82*, 1–17. https://doi.org/10.1007/s00445-019-1348-z

Lu, Z., Dzurisin, D., Lu, Z., & Dzurisin, D. (2014). *InSAR imaging of Aleutian volcanoes*. Springer.

Luo, Y., Taisne, B., Lythgoe, K. H., & Tan, C. T. (2023). *SuPreMeChiF-a new approach to detect subtle changes in continuous monitoring data, with a case study of COVID-19 impact in Singapore through seismic and infrasound recordings*. Authorea Preprints.

Maher, S. P., Dawson, P., Hotovec-Ellis, A., Thelen, W. A., Jolly, A., Bennington, N., et al. (2023). Characterizing and locating seismic tremor during the 2022 eruption of Mauna Loa volcano, Hawai'i, with network covariance. *The Seismic Record*, *3*(3), 228–238. https://doi.org/10.1785/0320230020

Matoza, R. S., Fee, D., & Garcés, M. A. (2010). Infrasonic tremor wavefield of the Pu'u 'Ō'ō crater complex and lava tube system, Hawaii, in April 2007. *Journal of Geophysical Research*, *115*(B12), B12312. https://doi.org/10.1029/2009JB007192

Matoza, R. S., & Roman, D. C. (2022). One hundred years of advances in volcano seismology and acoustics. *Bulletin of Volcanology*, *84*(9), 86. https://doi.org/10.1007/s00445-022-01586-0

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015). Librosa: Audio and music signal analysis in python. In *Proceedings of the 14th Python in science conference* (Vol. 8, pp. 18–25).

McNutt, S. R. (1986). Observations and analysis of B-type earthquakes, explosions, and volcanic tremor at Pavlof Volcano, Alaska. *Bulletin of the Seismological Society of America*, *76*(1), 153–175.

McNutt, S. R. (1987). Volcanic tremor at Pavlof Volcano, Alaska, October 1973–April 1986. *Pure and Applied Geophysics*, *125*(6), 1051–1077. https://doi.org/10.1007/bf00879368

McNutt, S. R., & Nishimura, T. (2008). Volcanic tremor during eruptions: Temporal characteristics, scaling and constraints on conduit size and processes. *Journal of Volcanology and Geothermal Research*, *178*(1), 10–18. https://doi.org/10.1016/j.jvolgeores.2008.03.010

Mousavi, S. M., & Beroza, G. C. (2023). Machine learning in earthquake seismology. *Annual Review of Earth and Planetary Sciences*, *51*(1), 105–129. https://doi.org/10.1146/annurev-earth-071822-100323

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nature Communications*, *11*(1), 1–12. https://doi.org/10.1038/s41467-020-17591-w

Mousavi, S. M., Zhu, W., Sheng, Y., & Beroza, G. C. (2019). CRED: A deep residual network of convolutional and recurrent units for earthquake signal detection. *Scientific Reports*, *9*(1), 10267. https://doi.org/10.1038/s41598-019-45748-1

Neuberg, J., Luckett, R., Baptie, B., & Olsen, K. (2000). Models of tremor and low-frequency earthquake swarms on Montserrat. *Journal of Volcanology and Geothermal Research*, *101*(1–2), 83–104. https://doi.org/10.1016/s0377-0273(00)00169-4

Orr, T. R., Dietterich, H. R., Fee, D., Girona, T., Grapenthin, R., Haney, M. M., & Waythomas, C. F. (2024). 2021 volcanic activity in Alaska and the Commonwealth of the Northern Mariana Islands—Summary of events and response of the Alaska Volcano Observatory (Report No. 2024-5014). https://doi.org/10.3133/sir20245014

Pesicek, J. D., Wellik, J. J., Prejean, S. G., & Ogburn, S. E. (2018). Prevalence of seismic rate anomalies preceding volcanic eruptions in Alaska. *Frontiers in Earth Science*, *6*. https://doi.org/10.3389/feart.2018.00100

Picard, D. (1985). Testing and estimating change-points in time series. *Advances in Applied Probability*, *17*(4), 841–867. https://doi.org/10.1017/s0001867800015433

Powell, T. W., & Neuberg, J. (2003). Time dependent features in tremor spectra. *Journal of Volcanology and Geothermal Research*, *128*(1–3), 177–185. https://doi.org/10.1016/s0377-0273(03)00253-1

Prejean, S. G., & Brodsky, E. E. (2011). Volcanic plume height measured by seismic waves based on a mechanical model. *Journal of Geophysical Research*, *116*(B1), B01306. https://doi.org/10.1029/2010jb007620

Reiss, M. C., Massimetti, F., Laizer, A. S., Campus, A., Rümpker, G., & Kazimoto, E. O. (2023). Overview of seismo-acoustic tremor at Oldoinyo Lengai, Tanzania: Shallow storage and eruptions of carbonatite melt. *Journal of Volcanology and Geothermal Research*, *442*, 107898. https://doi.org/10.1016/j.jvolgeores.2023.107898

Ripepe, M., & Marchetti, E. (2002). Array tracking of infrasonic sources at Stromboli volcano. *Geophysical Research Letters*, *29*(22), 33–41. https://doi.org/10.1029/2002gl015452

Ripepe, M., Marchetti, E., Bonadonna, C., Harris, A. J. L., Pioli, L., & Ulivieri, G. (2010). Monochromatic infrasonic tremor driven by persistent degassing and convection at Villarrica Volcano, Chile. *Geophysical Research Letters*, *37*(15), L15303. https://doi.org/10.1029/2010gl043516

Ripepe, M., Marchetti, E., & Ulivieri, G. (2007). Infrasonic monitoring at Stromboli volcano during the 2003 effusive eruption: Insights on the explosive and degassing process of an open conduit system. *Journal of Geophysical Research: Solid Earth*, *112*(B9), B09207. https://doi.org/10.1029/2006jb004613

Ripepe, M., Poggi, P., Braun, T., & Gordeev, E. (1996). Infrasonic waves and volcanic tremor at Stromboli. *Geophysical Research Letters*, *23*(2), 181–184. https://doi.org/10.1029/95gl03662

Roman, D. C. (2017). Automated detection and characterization of harmonic tremor in continuous seismic data. *Geophysical Research Letters*, *44*(12), 6065–6073. https://doi.org/10.1002/2017GL073715

Rouet-Leduc, B., Hulbert, C., McBrearty, I. W., & Johnson, P. A. (2020). Probing slow earthquakes with deep learning. *Geophysical Research Letters*, *47*(4), e2019GL085870. https://doi.org/10.1029/2019gl085870

Saunders-Shultz, P., Lopez, T., Dietterich, H., & Girona, T. (2024). Automatic identification and quantification of volcanic hotspots in Alaska using HotLINK: The hotspot learning and identification network. *Frontiers in Earth Science*, *12*, 1345104. https://doi.org/10.3389/feart.2024.1345104

Scandone, R., & Malone, S. D. (1985). Magma supply, magma discharge and readjustment of the feeding system of Mount St. Helens during 1980. *Journal of Volcanology and Geothermal Research*, *23*(3–4), 239–262. https://doi.org/10.1016/0377-0273(85)90036-8

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision* (pp. 618–626).

Seydoux, L., Shapiro, N. M., de Rosny, J., Brenguier, F., & Landès, M. (2016). Detecting seismic activity with a covariance matrix analysis of data recorded on seismic arrays. *Geophysical Journal International*, *204*(3), 1430–1442. https://doi.org/10.1093/gji/ggv531

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 1–48. https://doi.org/10.1186/s40537-019-0197-0

Smith, C. M., McNutt, S. R., & Thompson, G. (2016). Ground-coupled airwaves at Pavlof Volcano, Alaska, and their potential for eruption monitoring. *Bulletin of Volcanology*, *78*(7), 1–12. https://doi.org/10.1007/s00445-016-1045-0

Soubestre, J., Shapiro, N. M., Seydoux, L., de Rosny, J., Droznin, D. V., Droznina, S. Y., et al. (2018). Network-based detection and classification of seismovolcanic tremors: Example from the Klyuchevskoy volcanic group in Kamchatka. *Journal of Geophysical Research: Solid Earth*, *123*(1), 564–582. https://doi.org/10.1002/2017JB014726

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*(1), 1929–1958.

Steinke, B., Jolly, A. D., Carniel, R., Dempsey, D. E., & Cronin, S. J. (2023). Identification of seismo-volcanic regimes at Whakaari/White Island (New Zealand) via systematic tuning of an unsupervised classifier. *Journal of Geophysical Research: Solid Earth*, *128*(3), e2022JB026221. https://doi.org/10.1029/2022jb026221

Tan, D. & Fee, D. (2024). darren-tpk/voiss-net: Detection and characterization of volcanic tremor using machine learning (Version v1.0.2) [Computer software]. *Zenodo*. https://doi.org/10.5281/ZENODO.11672651

Theys, N., Hedelt, P., De Smedt, I., Lerot, C., Yu, H., Vlietinck, J., et al. (2019). Global monitoring of volcanic SO_2 degassing with unprecedented resolution from TROPOMI onboard Sentinel-5 Precursor. *Scientific Reports*, *9*(1), 2643. https://doi.org/10.1038/s41598-019-39279-y

Titos, M., Bueno, A., García, L., Benítez, M. C., & Ibañez, J. (2018). Detection and classification of continuous volcano-seismic signals with recurrent neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(4), 1936–1948. https://doi.org/10.1109/tgrs.2018.2870202

Tkachenko, M., Malyuk, M., Shevchenko, N., Holmanyuk, A., & Liubimov, N. (2020). Label studio: Data labeling software. Open source software, Retrieved from https://github.com/heartexlabs/label-studio

Unglert, K., Radić, V., & Jellinek, A. M. (2016). Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra. *Journal of Volcanology and Geothermal Research*, *320*, 58–74. https://doi.org/10.1016/j.jvolgeores.2016.04.014

U.S. Geological Survey, E. H. P. (2017). Advanced national seismic system (ANSS) comprehensive catalog of earthquake events and products: Various. https://doi.org/10.5066/F7MS3QZH

Wassermann, J. (2012). Volcano seismology. In *New manual of seismological observatory practice 2 (NMSOP-2)* (pp. 1–77). Deutsches GeoForschungsZentrum GFZ.

Watson, L. M. (2020). Using unsupervised machine learning to identify changes in eruptive behavior at Mount Etna, Italy. *Journal of Volcanology and Geothermal Research*, *405*, 107042. https://doi.org/10.1016/j.jvolgeores.2020.107042

Watson, L. M., Johnson, J. B., Sciotto, M., & Cannata, A. (2020). Changes in crater geometry revealed by inversion of harmonic infrasound observations: 24 December 2018 eruption of Mount Etna, Italy. *Geophysical Research Letters*, *47*(19), e2020GL088077. https://doi.org/10.1029/2020gl088077

Waythomas, C. F., Haney, M. M., Fee, D., Schneider, D. J., & Wech, A. (2014). The 2013 eruption of Pavlof Volcano, Alaska: A spatter eruption at an ice-and snow-clad volcano. *Bulletin of Volcanology*, *76*(10), 1–12. https://doi.org/10.1007/s00445-014-0862-2

Waythomas, C. F., Haney, M. M., Wallace, K. L., Cameron, C. E., & Schneider, D. J. (2017). The 2014 eruptions of Pavlof Volcano, Alaska (Technical Report No. 2328-0328). *US Geological Survey*.

Wech, A. G., & Creager, K. C. (2008). Automated detection and location of Cascadia tremor. *Geophysical Research Letters*, *35*(20), L20302. https://doi.org/10.1029/2008gl035458

Witsil, A., Fee, D., Dickey, J., Peña, R., Waxler, R., & Blom, P. (2022). Detecting large explosions with machine learning models trained on synthetic infrasound data. *Geophysical Research Letters*, *49*(11), e2022GL097785. https://doi.org/10.1029/2022gl097785

Woulff, G., & McGetchin, T. R. (1976). Acoustic noise from volcanoes: Theory and experiment. *Geophysical Journal International*, *45*(3), 601–616. https://doi.org/10.1111/j.1365-246x.1976.tb06913.x

Yamada, T., Aoyama, H., Nishimura, T., Iguchi, M., & Hendrasto, M. (2017). Volcanic eruption volume flux estimations from very long period infrasound signals. *Geophysical Research Letters*, *44*(1), 143–151. https://doi.org/10.1002/2016gl071047

Yang, T., Zhou, S., Wang, Y., Lu, Y., & Zheng, N. (2022). Test-time batch normalization. arXiv preprint arXiv:2205.10210.

Ye, L., Bai, Y., Si, D., Lay, T., Cheung, K. F., & Kanamori, H. (2022). Rupture model for the 29 July 2021 MW 8.2 Chignik, Alaska earthquake constrained by seismic, geodetic, and tsunami observations. *Journal of Geophysical Research: Solid Earth*, *127*(7), e2021JB023676. https://doi.org/10.1029/2021jb023676

Zali, Z., Mousavi, S. M., Ohrnberger, M., Eibl, E. P., & Cotton, F. (2024). Tremor clustering reveals pre-eruptive signals and evolution of the 2021 Geldingadalir eruption of the Fagradalsfjall Fires, Iceland. *Communications Earth & Environment*, *5*(1), 1. https://doi.org/10.1038/s43247-023-01166-w

Zali, Z., Ohrnberger, M., Scherbaum, F., Cotton, F., & Eibl, E. P. S. (2021). Volcanic tremor extraction and earthquake detection using music information retrieval algorithms. *Seismological Research Letters*, *92*(6), 3668–3681. https://doi.org/10.1785/0220210016

Zhao, Y. X., Li, Y., & Wu, N. (2022). Data augmentation and its application in distributed acoustic sensing data denoising. *Geophysical Journal International*, *228*(1), 119–133. https://doi.org/10.1093/gji/ggab345

Zhu, W., & Beroza, G. C. (2019). PhaseNet: A deep-neural-network-based seismic arrival-time picking method. *Geophysical Journal International*, *216*(1), 261–273.