Detection of Submarine Volcanic Eruptions Using Satellite Ocean-Colour and Machine Learning

Oliwia Pedrycz

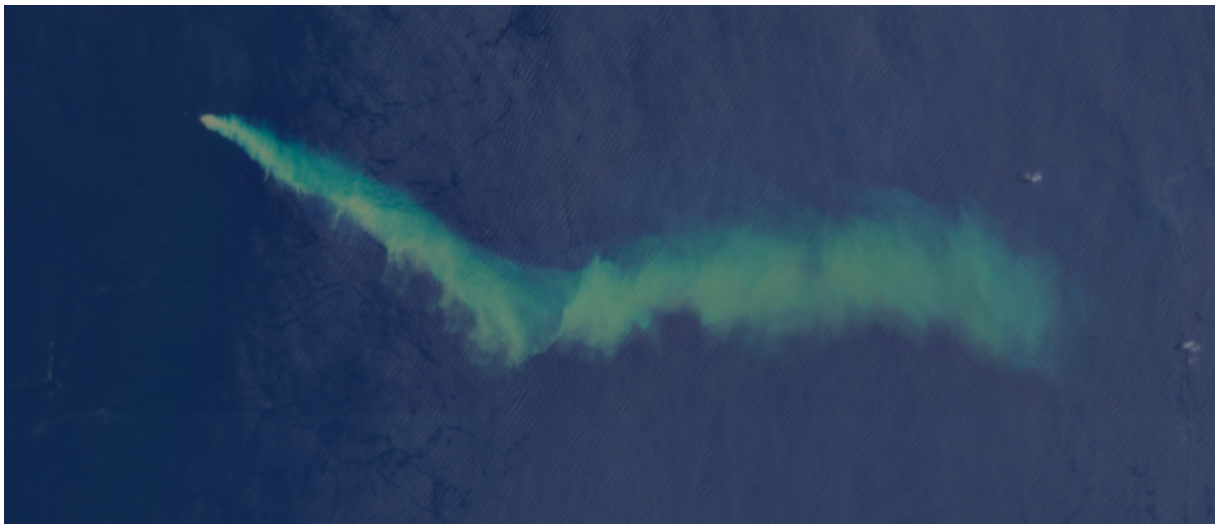
Supervisors: Rui Song, Don Grainger, Isabelle Taylor, Antonín Knížek

University of Oxford

September 2025

Abstract

Submarine volcanic eruptions are challenging to monitor due to their remote location. These eruptions impact the environment by changing nutrient levels and emitting greenhouse gases. Shallow submarine eruptions release heat, gases, and minerals, producing plumes that can alter the colour of the surrounding ocean water. We use a U-Net machine learning architecture, a convolutional neural network designed for image segmentation, and Sentinel-2 satellite data to detect submarine eruptions and reveal events that have previously gone unnoticed. Applying the model to the highly active Kavachi volcano in the Solomon Islands successfully identified eruptions that occurred between 2016 and 2025.



1 Introduction

Around 80% of present-day volcanic eruptive activity occurs in the ocean [1] . Submarine eruptions, despite being the dominant form of Earth's volcanism, are relatively understudied due to the challenges related to their monitoring, observing, and sampling of submarine geological features caused by their remote location [2] . These eruptions affect the environment, influencing nutrient levels, releasing greenhouse gases, and affecting seabed habitats.

The Fe-hydroxide-rich environment of the hydrothermal vents of the Kama'ehuakanaloa Seamount in Hawaii led to studies of novel microbial communities, resulting in the identification of a new Proteobacteria genus, *Mariprofundus ferrooxydans* [3]. Sharks were found to live around the highly active underwater volcano Kavachi in the Solomon Islands, a highly unique environment due to the heat, acidity, and turbidity of the water. The largest submarine eruption of the century, the Hunga Tonga-Hunga Ha'apai 2022 eruption, led to a dramatic bloom of phytoplankton. Satellite measurements estimated a 10-fold increase in phytoplankton in the proximity of the volcano [2]. Learning to spot such events is important for improving the understanding of ocean volcanism and associated marine ecosystems.

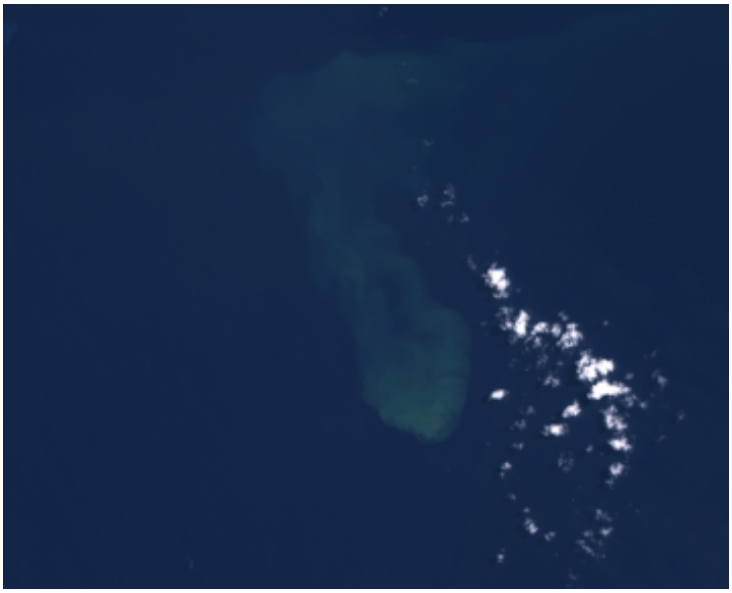


Figure 1: Eruption at Kavachi volcano in the Salomon Islands on the 21st of March 2023. A Sentinel-2 Level-2A image was obtained from the Copernicus Browser [4]. Eruption coordinates are included in the appendix.

Volcanic activity at shallow submarine volcanoes or those located near coastlines can be identified by the associated ocean-water discolouration. Such ocean discolourations often last longer than the main eruptive events and can be detected with remote sensing. Depending on the chemical components released during the eruption, the colour of the water can vary considerably, ranging from white and green to reddish-brown [5].

Machine learning, and particularly convolutional neural networks, is increasingly applied to the analysis of satellite imagery. Among these, the U-Net architecture has proven highly effective for image segmentation tasks [6]. In previous studies, this enabled efficient land cover classification and automatic identification of urban features with the use of Sentinel-2 data [7].

This project aimed to build a U-Net model using Sentinel-2 data to automatically identify ocean water-colour shifts caused by submarine eruptions and detect events that have gone unreported.

2 Methods

2.1 Data

Sentinel-2 satellite data was obtained from Google Earth Engine. Level-1 product was used as Google Earth Engine does not provide the full archive of Level-2A data. Band B10, the cirrus cloud detection band, has been excluded from the dataset, as it does not contribute to the objectives of the study.

2.2 Training Data

To train the model to recognise ocean discolouration due to submarine volcanic eruptions, a dataset of such eruptions has been produced. The database was created based on information on submarine volcanoes and eruptions available from the Smithsonian Institute's Global Volcanism Program and other online sources [8]. The Global Volcanism Program was helpful in identifying eruptions, as it provided data on Holocene volcanoes and their coordinates. It also provided information about most recent known eruptions, which were not always up-to-date for submarine eruptions that we investigated. To find relevant eruptions, near-surface active submarine volcanoes were considered.

The aim was to obtain a dataset of ocean discolourations with little cloud cover and no island visible. We trained the final model on 30 different discolourations at 6 different volcanoes: Kavachi (Solomon Islands), Volcano F (Tonga), Home Reef (Tonga), East Epi (Vanuatu), Metis Shoal (Tonga), and Kaitoku Seamount (Japan). We have also identified 22 additional eligible discolourations at some of these volcanoes, as well as Ruby Seamount (Northern Mariana Islands), Ahyi Seamount (Northern Mariana Islands) and Kick'em Jenny (Grenada). Although visible ocean discolourations have been found to occur at Hunga-Tonga (Tonga) and Tagoro (Canary Islands), they have not been included due to respectively the visibility of the island and extensive cloud cover. Detailed lists of identified eruptions are available in the appendix.

2.3 Training Data Preparation

For each considered eruption, a mask indicating the area of the discoloured water was created manually. The mask was loaded together with the 12-band satellite data file and binarised to create annotations indicating whether a pixel was eruptive.

Both files were split into 256×256-pixel square patches to increase the amount of data and decrease training time. Each patch had an overlap of 26 pixels to ensure that areas on edges, where the model could potentially be less accurate due to its tendency to best generalise central areas, would be accurately considered. When an image could not be split into an equal number of 256-pixel squares, it was cropped to an appropriate size, the edges being discarded. Patches at the end of the split image had large overlapping areas. Final output pairs were saved in a compressed Numpy archived file format.

2.4 Model Training and Testing

Model performance was investigated by loading the training dataset and excluding a set of random dates to evaluate as test data. Satellite files were then normalised by division by 1,000, and patches were classified to indicate whether an eruption was present or not. The dataset was undersampled to a 50:50 ratio of eruptive to non-eruptive patches by randomly selecting an equal number of non-eruptive patches to match the number of eruptive ones. The data was then split into training and validation, to obtain an overall approximate split of 8:1:1 between training, validation, and test data.

Training and validation data were considered in batches of 32, the first being shuffled and augmented. The last batch, if the training data were not divisible by 32, would be smaller. This ensured that the model did not use blank patches for training and consequently improved performance.

The U-Net architecture model was built with the use of the *Segmentation Models* package in Python [9]. After each training epoch, the validation data was evaluated by the measure of loss, indicating the error between the predictions and the actual value, optimised by the Adam optimiser. The loss function

applied was the combined Dice and Focal losses, which helped the model focus on both the size of the discoloured area and the hard-to-classify pixels [10]. The best model, one with minimal loss, was then saved and tested on the unseen eruptions.

Test data was loaded in a similar format to the training data, with the same number of bands and preparation steps, differing only in size. The input was tested by applying a sliding window to consider a square of 256 pixels, which was then moved by a step of 230 pixels to consider the next 256-pixel square, detecting eruptive pixels in the file. The newly generated mask was then saved and manually assessed against the true colour satellite image.

2.5 Model Parameters

Multiple parameters were considered for tuning. Performance of multiple pretrained backbones from the *Segmentations Models* repository, parameter responsible for feature extraction, was investigated (Table 1). The 20 backbones of varying complexity that were considered were chosen from all 8 available model families and showed considerable differences in performance.

densenet201	mobilenetv2	resnext101	seresnet50
efficientnetb1	resnet101	senet154	seresnext101
efficientnetb3	resnet152	seresnet152	seresnext50
efficientnetb7	resnet34	seresnet18	vgg16
inceptionresnetv2	resnet50	seresnet34	vgg19

Table 1: Model backbones considered.

Performance of backbones was first assessed by training the model on 11 eruptions at Kavachi volcano in the Solomon Islands with 500 epochs and a learning rate of 0.001. Training took approximately 40 minutes on the Orchid GPU (A100) within the UK supercomputer facility JASMIN. Best performing models have been tested on four different eruptions and manually assessed on whether they detected an eruption, how wide the area of detection was and the number of false positives caused by cloud cover. While the metrics of precision, recall, and the intersection of the union, widely used for evaluating model performance, were used, these metrics have been monitored for validation data to track whether training was stable rather than treated as a measure of success (Figure 2). That is because ocean discolouration extent varies depending on the volcano, proximity, time passed since eruption and eruption strength. Therefore, creating an accurate mask for an eruption was challenging as annotating extremely subtle discolourations could potentially cause more false positive detections. Additionally, it was often challenging to manually identify which small-scale areas were more discoloured than others due to the colour gradient. Finally, the model proved to detect discolourations sometimes extending the area that we have manually considered to indicate an eruption.

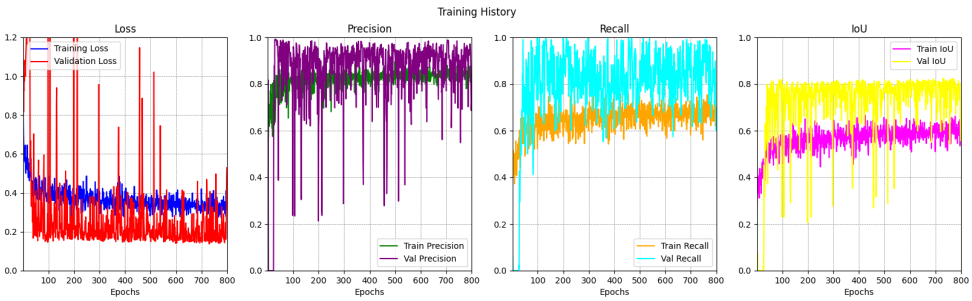


Figure 2: Metrics used to monitor model performance did not effectively reflect detection success.

Only two models, ones with backbones seresnet18 and efficientnetb1, were able to detect the small-sized eruption at Kavachi on the 21st of March 2023 (Figure 3). Other learning rates, smaller and larger than

0.001, were considered for both, showing the best performance of seresnet18 at 0.001, and comparable performance of efficientnetb1.

Finally, both models were trained on the final dataset of eruptions and assessed against 3 varying test eruptions. A few models were trained for 500 to 1000 epochs, with validation data evaluated after each epoch. The model achieving the best validation performance was saved as the final model. Seresnet18 with 800 epochs was chosen as the best performing model due to the high number of successfully detected eruptive pixels, but fewer false positive predictions were attributed to cloud cover.

Out of 10 random eruptions that the model was tested on, it detected all and showed no obvious false positives, especially those related to cloud cover.

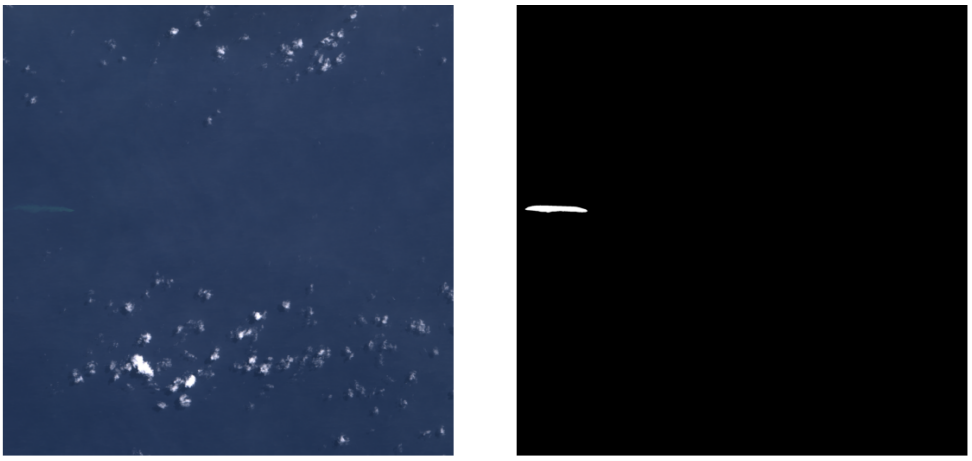


Figure 3: RGB image (left) and the corresponding mask (right) of the eruption at Kavachi volcano on the 26th of March 2021, challenging to detect for initial models.

2.6 Test Region

The final model was applied to the test area at Kavachi volcano in the Solomon Islands. Time series between the 1st of January 2016 and the 1st of January 2025 were considered and fed into the model. Masks for each available date were generated and manually assessed to consider further training steps.

Further improvements to the model were considered on the basis of model performance in the test area. That was approached by increasing the number of training data to include problematic cloud cover or features. A list of non-eruptive dates with extensive and problematic cloud cover, included in the training dataset for the final model at the current stage, is provided in the appendix. As problematic dates generally included heavy cloud cover, all potential dates have been carefully inspected with the use of the Copernicus browser before being added to training [4]. This is because numerous detections of subtle eruptions, initially thought to be false positives, were later confirmed to be actual eruptions. The training and application of the model to the test area took approximately 6 hours using the Orchid GPU within JASMIN.

A few eruptions at Kavachi volcano, which proved to be particularly difficult for models to detect, have been used to assess model performance. Each of these eruptions had varying detection success depending on the parameters and training data used. A list of these eruptions is included in the appendix.

3 Results

The model successfully detected eruptions at Kavachi volcano. Visual inspection showed that the model detected a significant majority of eruptions, as shown for illustrative eruptive periods between May and June and August and September of 2022 in Figure 4 and Figure 5.

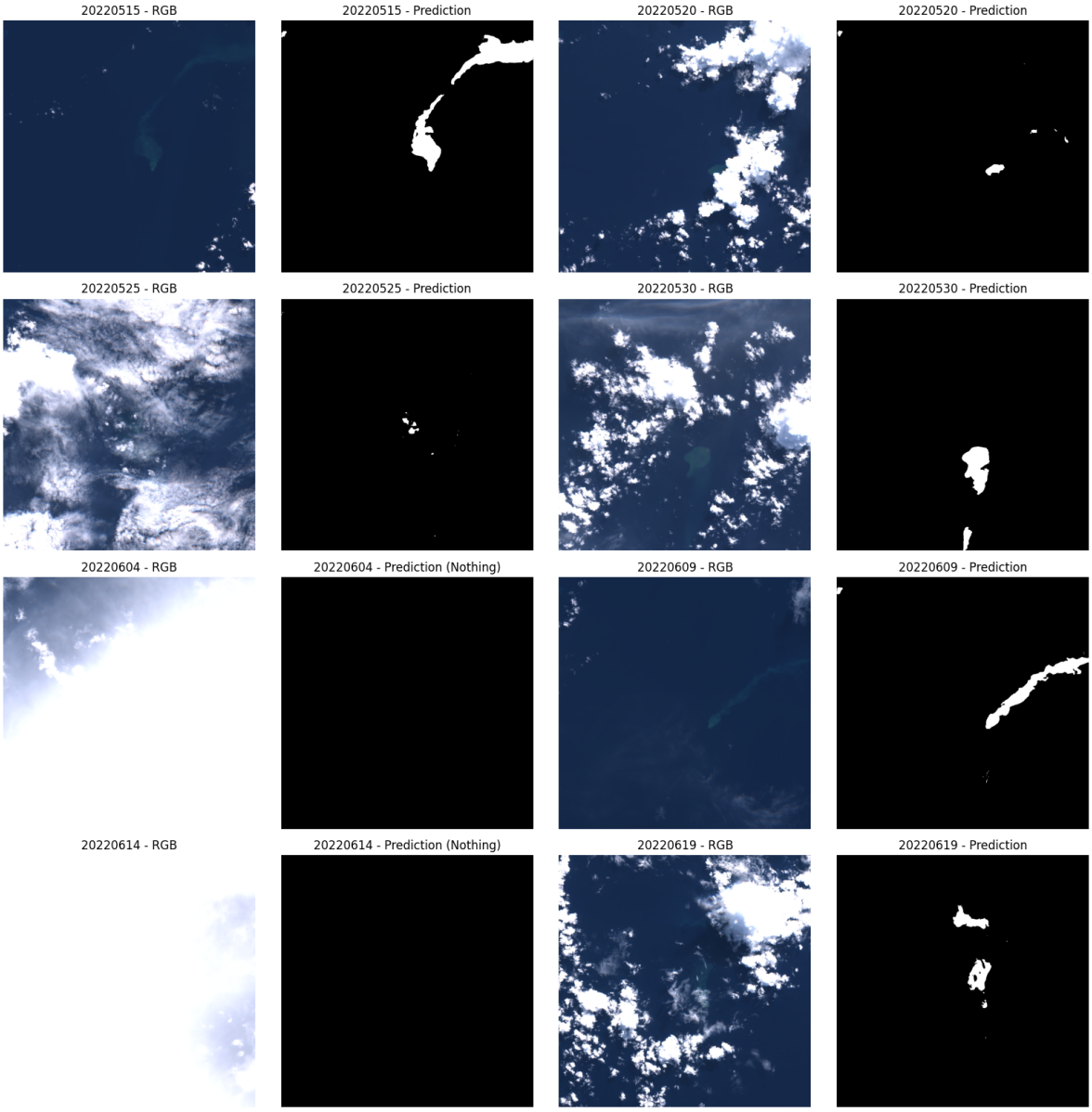


Figure 4: Detections at Kavachi volcano in May and June of 2022.

Moreover, the model was often able to detect ocean discolouration through light and moderate cloud mask (Figure 6). Occasionally, due to extensive cloud cover on some days, it was not possible to verify whether the few predicted pixels were accurate. These clusters would sometimes look similar to the ones on the mask that was generated for the marginally detected eruption on the 18th of October 2018 (Figure 7), but would be likely false positives at other times.

Although the model showed very good performance in detecting eruptions, it would sometimes struggle with generating false positive detections due to cloud cover. That has been improved significantly by adding 9 problematic dates with an empty mask to the training. However, as the data was undersampled to match the number of eruptive patches, only a small number of such problematic patches would randomly be present for training. The problem was not fully resolved. Masks also sometimes included an incorrect eruptive pixel that was not visible due to its size. That has been monitored by specifically indicating whether a mask was truly empty or not when assessing the results.

The model also generated a false positive detection for a nearby shallow area within the ocean, similar in colour and shape to an eruptive plume, located north-west of the volcano. To approach that, 6 patches including the shallow area were added to the training data. That was done after the undersampling step,

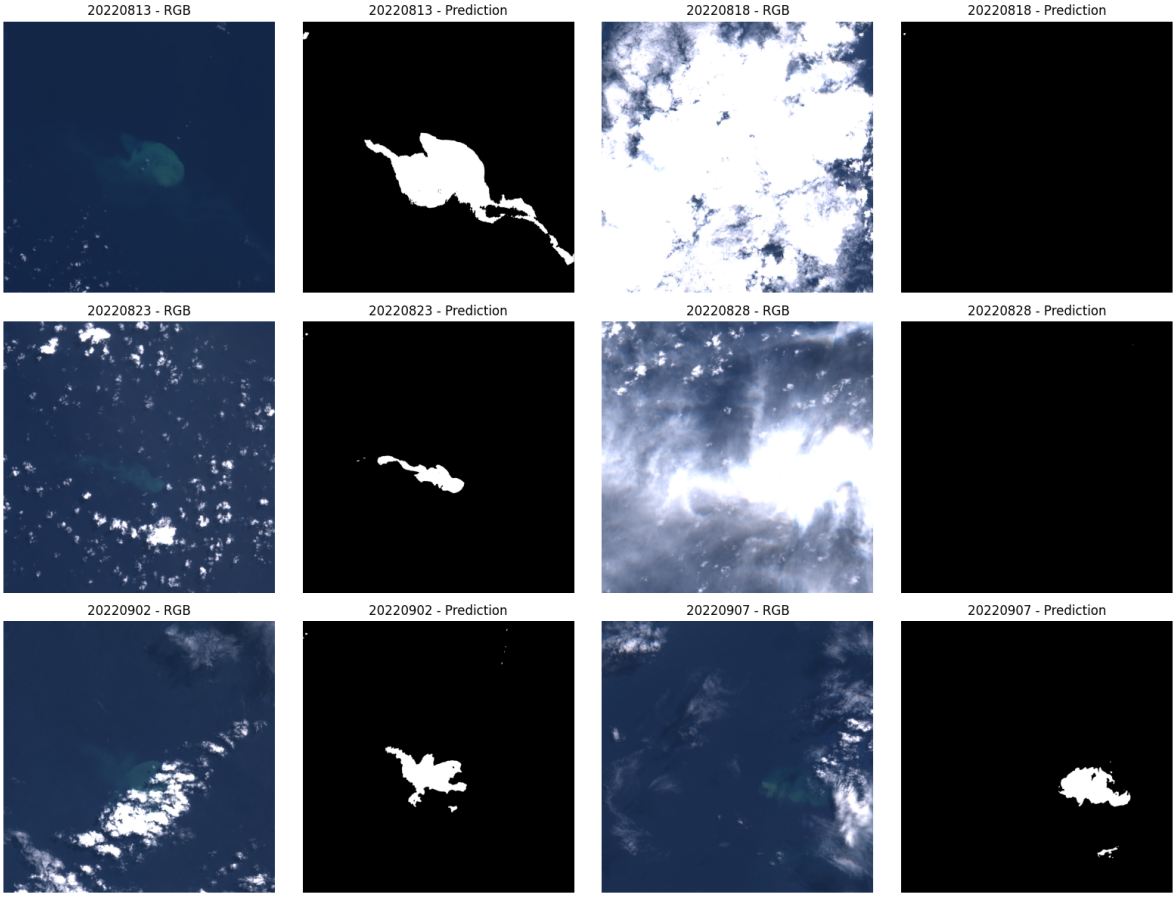


Figure 5: Detections at Kavachi in August and September of 2022.

to ensure that they were included. However, that has not improved the performance. Although the model has not yet been tested on other volcanoes, it could potentially cause issues, particularly for volcanoes in close proximity to land.

Different models showed varying sensitivity to plankton blooming, similar in colour to a discolouration caused by an eruption. The final model falsely classified one date of extensive plankton blooming as an eruption, the 21st of November 2021. That date has not yet been added to the dataset for further training to explore its impact.

4 Discussion

The model effectively detected submarine eruptions at Kavachi, generally handling well easily visible discolourations, as well as those that were more challenging to detect. A few eruptions have been missed by the model, these being under significant cloud cover. However, three alternative models were tested on the full test dataset to assess the outcome against the model deemed to be best-performing. None of the alternative models showed better overall performance, however, some of the eruptions missed were successfully detected by an alternative model. Unfortunately, an alternative model would miss other eruptions instead.

The model with the second-best-performing backbone, efficientnetb1, showed good sensitivity to eas ily visible eruptions and sometimes these under cloud cover, but would be less sensitive to small-sized discolourations. The second and third models, with a proportion of 38:62 and 44:56 of eruptive and non-eruptive patches, decreased the number of false positive detections but worsened detection success. Therefore, the final model was found to have the best performing parameters. Additionally, a visual inspection of satellite data of the full region with the use of the Copernicus browser showed multiple

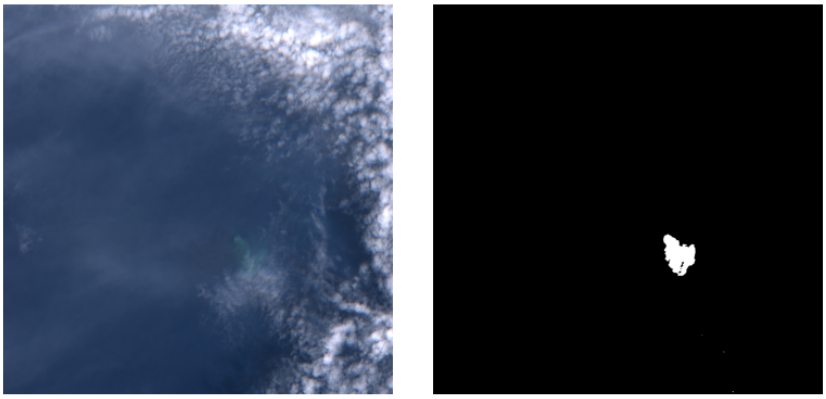


Figure 6: The model detected an ocean discolouration under the cloud cover during the eruption on the 2nd of December 2018.

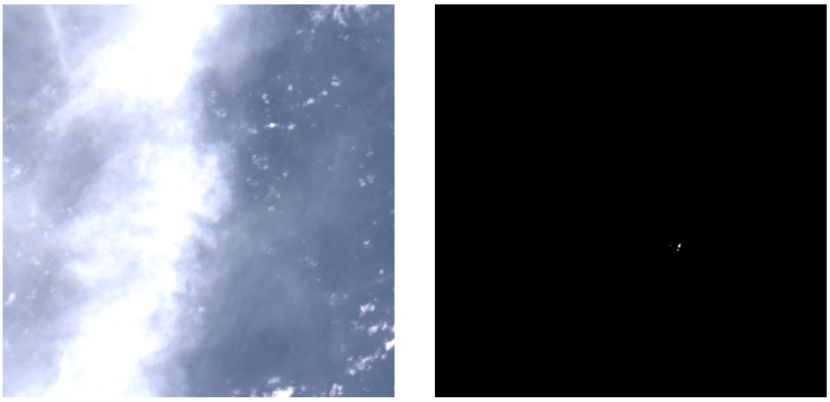


Figure 7: The model detected a subtle discolouration on the 18th of October 2018.

times that a cluster of a few pixels, looking like a false positive detection, was indeed a true eruption.

Although false positive pixels due to cloud cover have been detected, adding specific dates that caused problems to the training has decreased the issue and improved performance. Further increase in training data would strengthen the model. Particularly, manually extracting and including problematic patches rather than including the full satellite image could prove beneficial. That is to ensure that the patch is included in training rather than being randomly discarded at the undersampling stage. As each image is split into 100 patches, identifying the problematic patch can sometimes be straightforward, as considered for the shallow area northwest of Kavachi. Additionally, exploring the impact of adding the false positive detection due to plankton blooming on the 21st of November 2021 to training data would be appropriate to assess whether it impacts detection success. That is important as plankton blooming might cause challenges at other volcanoes.

Adding patches that included the shallow area northwestern of Kavachi which was falsely being detected as an eruption, did not improve performance. While increasing the number of such patches could potentially benefit the model, the discolouration is similar in colour and shape to a submarine volcano's plume, and the model might continue struggling. Depending on whether that is an issue on other volcanoes, an increase in the number of classes beyond eruption and non-eruption could help address the issue in the future.

The use of seismic data could be beneficial to help assess the model's performance, especially during days with extensive cloud cover, for which verification is particularly challenging. Running the model with data from a geostationary satellite could help detect eruptions in more frequent periods.

The model showed good performance at Kavachi, and further evaluation on other volcanoes would provide more certain conclusions. Most of the training data used included eruptions at Kavachi volcano, and it

is of interest whether the model applied to other volcanoes would be equally effective due to different colours and shapes. However, testing the model on eruptions at other volcanoes during the initial stages did not show a difference in performance. If necessary, a list of unused eruptions at volcanoes other than Kavachi is provided in the appendix.

Finally, when the model reaches a satisfactory level of performance, the data can be visualised by analysing the number of eruptions per year to understand when periods of activity occurred. Including the images, where the number of pixels is larger than a set number, would allow for obtaining truly eruptive dates, while not including dates where a few random false positive pixels were generated.

5 Conclusions

Applying the convolutional neural network U-Net architecture model to Sentinel-2 imagery proved successful in detecting ocean discolouration caused by submarine eruptions. The best-performing model was able to successfully detect both large and small eruptions even under light to moderate cloud cover.

Although the model detected a significant majority of eruptions, it missed those that were under extensive cloud cover. The model also struggled a little with false positive detections due to cloud cover, which has been significantly decreased by adding more cloud cover imagery to the training data. Finally, further work is needed to decrease the false positive detections attributed to a shallow-ocean area of shape and colour similar to a submarine volcanic plume.

At the Kavachi volcano, the model demonstrated very promising performance. Applying the model to other shallow submarine volcanoes, as listed in the appendix, would provide insight into model performance on other volcanoes.

Appendix

Ocean discolourations used for training

Volcano	Location	Latitude	Longitude	Width (km)	Date
East Epi	Vanuatu	-16.6806	168.38518	3	04/08/2017
East Epi	Vanuatu	-16.67	168.38295	5	09/07/2022
Home Reef	Tonga	-18.9898	-174.7579	5	06/06/2021
Kaitoku Seamount	Japan	26.155	141.14239	30	23/10/2022
Kavachi	Solomon Islands	-8.98	157.96	20	19/02/2024
Kavachi	Solomon Islands	-9.0129	158.0676	30	26/03/2021
Kavachi	Solomon Islands	-8.994	157.96179	25	02/12/2016
Kavachi	Solomon Islands	-8.9961	157.95	15	22/12/2017
Kavachi	Solomon Islands	-8.9806	157.9252	25	01/01/2018
Kavachi	Solomon Islands	-8.9808	157.9631	25	16/01/2018
Kavachi	Solomon Islands	-8.9948	157.9675	5	20/02/2018
Kavachi	Solomon Islands	-9.0391	157.959	20	25/02/2018
Kavachi	Solomon Islands	-8.998	157.9538	15	02/03/2018
Kavachi	Solomon Islands	-8.9976	157.9627	15	26/04/2018
Kavachi	Solomon Islands	-8.98798	157.959	15	25/06/2018
Kavachi	Solomon Islands	-8.98757	157.97823	8	08/10/2018
Kavachi	Solomon Islands	-8.97037	157.987	22	25/04/2020
Kavachi	Solomon Islands	-8.9797	158.04794	13	02/10/2020
Kavachi	Solomon Islands	-9.02717	158.03988	32	12/10/2021
Kavachi	Solomon Islands	-8.99	158	15	19/02/2022
Kavachi	Solomon Islands	-8.97105	158	22	21/03/2022
Kavachi	Solomon Islands	-9.003	157.95954	13	20/04/2022
Kavachi	Solomon Islands	-8.92476	157.99872	25	15/05/2022
Kavachi	Solomon Islands	-8.99207	157.9961	15	09/07/2022
Kavachi	Solomon Islands	-8.97529	157.96829	20	13/08/2022
Kavachi	Solomon Islands	-9.031	157.93	30	27/10/2022
Metis Shoal	Tonga	-19.17436	-174.855	4	01/08/2017
Metis Shoal	Tonga	-19.193	-174.86269	10	08/01/2020
Metis Shoal	Tonga	-19.16445	-174.82098	7	23/03/2020
Volcano F	Tonga	-18.28152	-174.4	12	08/08/2019

Ocean discolourations not used for training

Volcano	Location	Latitude	Longitude	Width (km)	Date
Ahyi Seamount	Northern Mariana	20.44591	145.03707	7	01/12/2022
Ahyi Seamount	Northern Mariana	20.43763	145.02803	6	05/01/2023
Ahyi Seamount	Northern Mariana	20.41881	145.00919	7	26/12/2022
Ahyi Seamount	Northern Mariana	20.41881	145.00919	14	25/01/2024
Home Reef	Tonga	-18.964	-174.742	7	21/06/2021
Home Reef	Tonga	-18.9898	-174.7579	10	21/07/2021
Kaitoku Seamount	Japan	26.1	141.1	12	27/12/2022
Kaitoku Seamount	Japan	26.1183	141.05656	70	01/01/2023
Kaitoku Seamount	Japan	26.122	141.1022	30	16/01/2023
Kavachi	Solomon Islands	-8.98	157.771	50	30/04/2020
Kavachi	Solomon Islands	-8.94578	157.9549	22	21/03/2023
Kavachi	Solomon Islands	-9	157.96452	8	20/04/2023
Kavachi	Solomon Islands	-8.99377	157.95816	12	14/06/2023
Kavachi	Solomon Islands	-9	157.95559	17	21/12/2023
Metis Shoal	Tonga	-19.18538	-174.841	6	21/05/2024
Metis Shoal	Tonga	-19.17792	-174.8493	6	15/06/2024
Metis Shoal	Tonga	-19.1714	-174.84307	6	16/06/2020
Metis Shoal	Tonga	-19.18251	-174.84973	8	06/07/2020
Metis Shoal	Tonga	-19.17792	-174.8493	10	21/07/2021
Ruby Seamount	Northern Mariana Islands	15.7536	145.816	55	19/09/2023
Volcano F	Tonga	-18.4	-174.49	40	11/08/2019
Kick 'em Jenny	Grenada	-61.627868	12.289688	3	29/04/2019

Ocean discolourations unsuitable for training

Volcano	Location	Latitude	Longitude	Width (km)	Date	Exclusion
Hunga Tonga	Tonga	-20.58153	-175.40874	7	20/09/2018	Island visible
Hunga Tonga	Tonga	-20.57815	-175.392	7	12/02/2019	Island visible
Hunga Tonga	Tonga	-20.4952	-175.3679	60	28/12/2021	Island visible
Metis Shoal	Tonga	-19.16052	-174.85179	8	17/04/2020	Cloudy
Tagoro	Canary Islands	27.55181	-17.95389	5	05/02/2018	Cloudy

Non-eruptive dates at Kavachi volcano used for training

30/05/2020	04/07/2020	14/07/2020	19/07/2020	19/05/2024
31/12/2020	04/07/2021	22/09/2020	16/05/2019	

Eruptive dates at Kavachi volcano used for model assessment

09/08/2018	19/08/2018	18/10/2018	23/10/2018	25/02/2019

References

[1] Vasco M. Mantas, A.J.S.C. Pereira, and Paula V. Morais. Plumes of discolored water of volcanic origin and possible implications for algal communities. the case of the Home Reef eruption of 2006 (Tonga, Southwest Pacific Ocean). *Remote Sensing of Environment*, 115(6):1341–1352, 2011.

[2] B. Barone, R. M. Letelier, K. H. Rubin, and D. M. Karl. Satellite detection of a massive phytoplankton bloom following the 2022 submarine eruption of the Hunga Tonga-Hunga Haapai volcano. *Geophysical Research Letters*, 49(17):e2022GL099293, 2022.

[3] D. Emerson, J.A. Rentz, T.G. Lilburn, R.E. Davis, H. Aldrich, C. Chan, and C.L. Moyer. A novel lineage of proteobacteria involved in formation of marine fe-oxidizing microbial mat communities. *PLOS ONE*, 2007.

[4] Copernicus Data Space Ecosystem. Copernicus browser, 2025. Accessed: October 1, 2025.

[5] Minoru Urai and Shoichi Machida. Discolored seawater detection using aster reflectance products: A case study of Satsuma-Iwojima, Japan. *Remote Sensing of Environment*, 99(1):95–104, 2005. Scientific Results from ASTER.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional networks for biomedical image segmentation. arXiv preprint arXiv:1505.04597, May 2015. Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).

[7] I. Kotaridis and M. Lazaridou. Semantic segmentation using a U-Net architecture on Sentinel-2 data. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XLIII-B3-2022, pages 6–11, Nice, France, June 2022. XXIV ISPRS Congress.

[8] Global Volcanism Program. Volcanoes of the world, version 5.3.1. Accessed: 6 Aug 2025, August 2025. Database. Distributed by Smithsonian Institution, compiled by Venzke, E.

[9] Pavel Iakubovskii. Segmentation models. `https://github.com/qubvel/segmentation_models`, 2019. GitHub repository.

[10] Wentao Zhu, Yufang Huang, Hui Tang, Zhen Qian, Nan Du, Wei Fan, and Xiaohui Xie. Anatomynet: Deep 3D squeeze-and-excitation U-Nets for fast and fully automated whole-volume anatomical segmentation. *CoRR*, abs/1808.05238, 2018.