**A machine learning approach to identifying and classifying volcanic deposits**

[M.N. Trouva](https://repository.tudelft.nl/person/Person_d0f04867-2f7e-45d9-ab70-ff47c064c47e)

**Published Date**

16-08-2024

Summary

The thesis develops a framework using exclusively Synthetic Aperture Radar (SAR) data to identify and classify volcanic deposits into Pyroclastic Density Currents (PDCs) and lava flows, which is crucial for planetary exploration where direct sampling is infeasible. Confirming the existence of PDCs on a planetary body provides insights into its geologic and atmospheric evolution, volatile content, and even potential for habitability. The study uses high-resolution, dual-polarised Level-1 Ground Range Detected (GRD) SAR C-band imagery from the Sentinel-1 mission. Case studies from Sinabung (Indonesia), Fuego (Guatemala), Kilauea (Hawaii), and Nyiragongo (Congo) offer a diverse dataset. The framework involves preprocessing steps suitable for highly variable topographic terrain, including calibration, despeckling, and radiometric and geometric corrections to produce reflectivity estimates in VV and VH polarisations. Texture features are derived using the Gray Level Co-occurrence Matrix (GLCM) method, including contrast, homogeneity, dissimilarity, and angular second moment (ASM). They are used as input to the first classifier alongside more straightforward metrics like polarisation ratio, Normalised Difference Index (NDI), and local variance. The first model distinguishes deposit pixels from the scene’s background. The deposit pixels are further classified as PDCs or lava flows by an extreme gradient boosting classifier using the same features. The models demonstrated high F1 scores, indicating a good balance between precision and recall, though both exhibited high RMSE. The first model’s high RMSE (0.35) is not concerning for this application, considering an accurate outline is unnecessary. The second model’s RMSE (0.42) coupled with a high False Negative rate (25%) reduces the method’s reliability. Nonetheless, when tested on unseen data from Soufriere Hills (Montserrat) and Pahoa (Hawaii), the models produced accurate results with at least double the pixels correctly identified as their respective deposit type. Including GLCM-derived texture features significantly enhanced accuracy, validating their use in volcanic deposit classification. This study’s contributions to remote sensing in volcanology provide a foundation for understanding remote worlds. Future research could explore the impact of resolution on classification and test different texture measures as complementary or substitute features. Integrating part of the workflow into satellite firmware could enable near real-time monitoring and enhance hazard assessment and mitigation strategies.