

AlphaEarth: Geo-Risk using Earth Observation and ML

Shubhankar Mahanta

mailto:hello@spass.uk

Challenge Tackled

After a disaster incident, traditional insurance workflows often collapse under the weight of slow paperwork and inaccessible sites. Victims and response agencies face long delays before claims are processed or aid is dispatched, primarily because assessing damage requires physical verification. This project draws inspiration from Google DeepMind’s AlphaEarth initiative, which demonstrated how large-scale geospatial intelligence can reshape risk perception and disaster management. By integrating satellite imagery, terrain information, and real-time weather data, this work automates geo-risk prediction—enabling faster response, early alerts, and fairer insurance claim assessment.

Approach and Model

The system is centered around a compact deep learning architecture called **CompactGeoEmbed**, a lightweight CNN derived from MobileNetV2. It fuses optical satellite imagery, elevation maps, and environmental factors such as precipitation, humidity, and wind speed to estimate a composite risk score for any coordinate. The inference pipeline retrieves Sentinel imagery and elevation data through Google Earth Engine, processes it locally using a quantized PyTorch model under 50MB, and presents the results via a Gradio web interface hosted on Hugging Face Spaces. The design supports CPU-only inference, ensuring low deployment cost and accessibility.

Data and Implementation

The model was trained on synthetically curated geospatial data combining open datasets such as MODIS (fire burn area), SMAP (soil moisture), and SRTM (terrain elevation). Although limited by API quotas and inconsistent spatial resolutions, careful sampling, caching, and cross-sensor alignment enabled construction of a functional dataset suitable for experimentation. The model generalizes well across diverse landscapes and maintains high consistency under quantized inference.

Challenges

Development revealed several difficulties, including memory overflow during large-area queries on Earth Engine, harmonizing multi-sensor resolutions (250m vs. 1km), and a general scarcity of open hazard datasets from agencies like UNEP or SEDAC. Despite these limitations, the framework achieved stable, end-to-end inference using only geographic coordinates and weather data.

Time Allocation

Data preparation and testing: 8 hours. Model training and visualization: 2 hours. Deployment and hosting optimization: 2.5 hours. Code structuring and documentation: 2 hours.

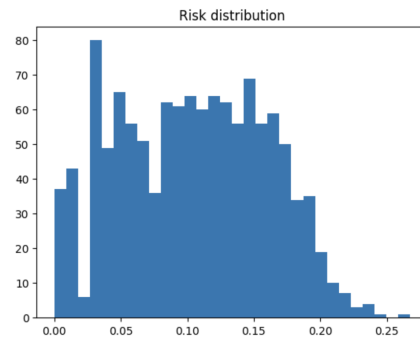


Figure 1: Geo-Risk Distribution on 1200 samples

Project Summary

The final prototype demonstrates a real-time geo-risk estimation framework that integrates Earth observation data with machine learning for actionable insurance intelligence. Its architecture emphasizes reproducibility, interpretability, and low computational overhead, serving as a foundation for scalable systems that can automate damage assessment, enhance transparency, and accelerate post-disaster recovery.

Live Demo

Hosting using Huggingface CPU Spaces

[/spaces/Pingsz/3rd-hack-nation](https://spaces.huggingface.co/Pingsz/3rd-hack-nation)

Project Reproducible version hosted on GitHub

[/sp4s-s/3rd-hack-nation](https://github.com/sp4s-s/3rd-hack-nation)