A Field-Deployable IoT-Edge-AI Framework for Landslide Prediction and Early Warning

Amrita Joshi *1 **Leonard Ng. Wei Tat** *1

*Equal contribution [1]*School of Materials Science and Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798*. Correspondence to: Leonard Ng. Wei Tat leonard.ngwt@ntu.edu.sg.

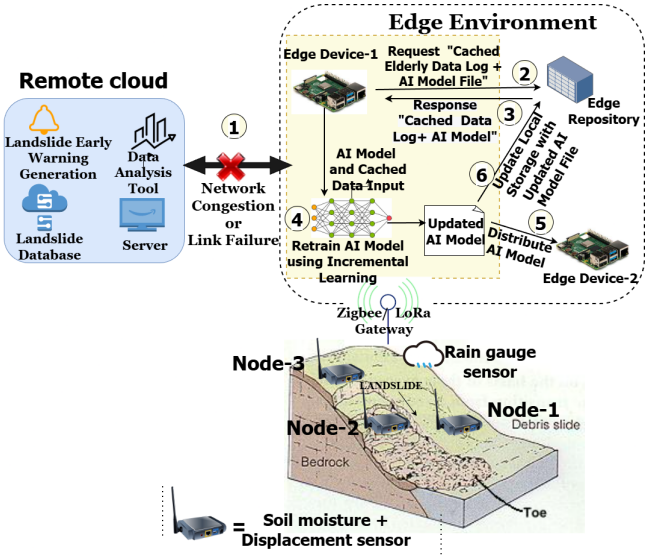

Fig. 1: IoT-EDGE-AI Architecture for landslide monitoring, Prediction, and Early Warning.

1. Introduction

Landslides represent one of the most destructive natural hazards, particularly in mountainous and high-rainfall regions, where sudden slope failures pose severe risks to human life, infrastructure, and the environment. Conventional landslide monitoring systems typically rely on centralized data acquisition and cloud-based analysis, resulting in high latency, dependence on continuous network connectivity, and delayed warning dissemination [1] [2] [3]. These limitations significantly reduce the effectiveness of early warning systems, especially in remote and infrastructure-scarce regions where landslides are most prevalent [4, 5].

Recent advances in Artificial Intelligence (AI) and embedded computing have enabled a paradigm shift toward edge intelligence, where data processing and decision-making occur close to the sensing source [6]. Edge AI offers the potential to reduce response time, improve system robustness, and minimize communication overhead by enabling local inference rather than continuous raw data transmission. However, integrating AI into resource-constrained IoT nodes while maintaining system reliability and scalability remains a significant challenge [7, 8].

In this work, we present a field-deployable IoT–Edge-AI architecture for landslide monitoring and prediction as shown in Fig. 1, where AI-driven inference is executed directly at the edge. The proposed system combines heterogeneous geotechnical and environmental sensing with lightweight 1 D CNN model deployed on embedded hardware. Unlike cloud-centric approaches, the system generates landslide risk warnings locally and transmits only event-driven alerts to the cloud. This design ensures low-latency response, reduced power consumption, and continued operation under intermittent network connectivity. The proposed framework is validated using real-world deployment scenarios, suitabile for practical landslide early warning applications.

1.1 Methodology: System Architecture and Edge Intelligence Framework

The proposed landslide monitoring system is designed as a distributed intelligence framework comprising three tightly coupled layers: the sensing layer, the edge intelligence layer, and the cloud layer. Each layer is assigned a distinct functional role, collectively enabling reliable, low-latency hazard detection.

At the sensing layer, multiple IoT sensor nodes are deployed in landslide-prone regions to continuously monitor critical parameters influencing slope stability as shown in Fig 1. These parameters include soil moisture content, rainfall intensity, pore-water pressure, and ground displacement. The sensors are selected to capture both short-term triggering factors (e.g., rainfall bursts) and long-term degradation indicators (e.g., gradual soil saturation). Sensor data are sampled at predefined intervals and forwarded to the edge processing unit for local analysis.

The edge intelligence layer forms the core contribution of this work. Multiple embedded edge devices hosts a lightweight AI model (1-D CNN) trained to recognize patterns indicative of landslide initiation. Instead of transmitting raw sensor data to the cloud, the edge node performs real-time inference and evaluates the predicted landslide probability against predefined risk thresholds. When normal conditions are detected, the system remains in a low-communication state, conserving bandwidth and energy. When elevated risk is identified using AI prediction, the edge node immediately generates a warning signal and initiates alert transmission. The cloud layer serves as a supervisory and archival component rather than the primary decision-maker. Alerts generated at the edge are forwarded to the cloud for visualization, logging, and dissemination to relevant authorities. Additionally, the cloud aggregates long-term data for model retraining, system diagnostics, and historical trend analysis. This division of responsibilities ensures that time-critical decisions are handled locally, while the cloud provides scalability and centralized oversight. This architecture deliberately

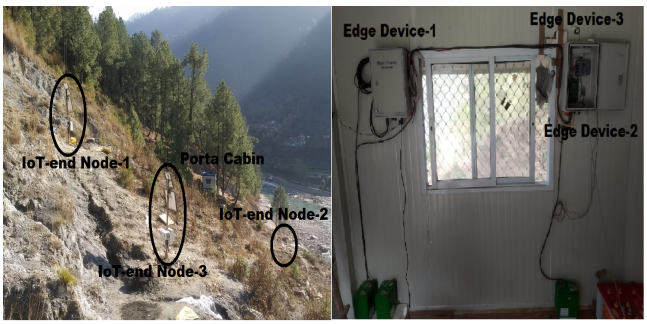

Fig. 2: Deployment of IoT-end nodes and edge environment (inside portacabin) at landslide location.

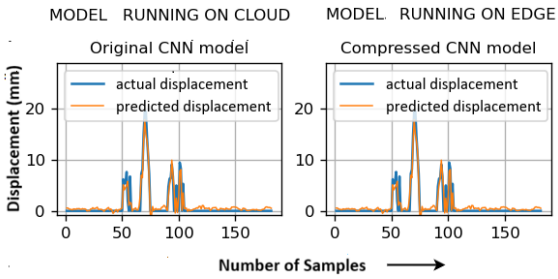

Fig. 3: Prediction of displacement in actual landslide location using original 1-D CNN running on Cloud and compressed 1-D CNN model running on Edge.

decouples decision latency from network reliability, enabling the system to function autonomously during communication outages—a critical requirement for disaster-prone and remote environments.

2. Experimental Validation and Results

The proposed IoT–Edge–AI architecture was experimentally validated using real-world deployments in landslide-prone region, as shown in Fig. 2. The evaluation focuses on highlighting the benefits of executing AI inference at the edge in comparison to conventional cloud-centric AI systems. Edge-AI on resource constraint edge device performance was evaluated in terms of response latency, communication efficiency, power consumption, and operational robustness in comparison to original cloud complex AI models. Edge-based inference enabled landslide risk detection and warning generation within approximately 150–200 ms, representing a substantial improvement over traditional cloud-dependent AI based systems that typically incur multi-second delays due to data transmission and centralized processing. This low-latency response is critical for early warning applications where even small delays can have severe consequences. Compared to continuous raw data streaming to cloud, the proposed system reduced data transmission volume by approximately 65–70% by enabling only high-risk events transmission, enabling scalable deployment across large geographic areas without overwhelming network infrastructure. Correspondingly, power consumption measurements indicated a 35–40% reduction in node-level energy usage,

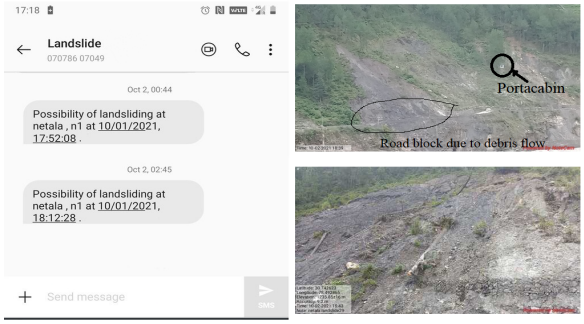

Fig. 4: Early warning generation by the edge device installed at the vicinity of the landslide location.

Table 1: Average energy consumption of original and compressed 1D CNN model.

Model	MSE	Average Power (W)	Execution Time (s)	Average Energy (J)
Original 1D CNN	1.4103	3.968	24.13	95.747
Pruned and Quantized 1 D CNN Model	1.4103	3.893	18	69.336

due to reduced wireless communication activity.

Figure 3 illustrates the 24-hour ahead landslide displacement predictions obtained from the compressed(prunning and Quantization) AI models deployed at the edge and the corresponding original AI models executed in the cloud. The predicted displacement values (expressed in millimeters) are evaluated against the observed ground-truth displacement measurements from real landslide location data. The results indicate that both the cloud-based original models and the edge-based compressed models closely track the actual displacement trends, demonstrating strong agreement with the measured values as shown in Table 1. as well. Also the table shows the efficacy of using compressed AI model over complex AI model.When executing five concurrent requests for inference generation at the edge device, the original 1D CNN model requires an average execution time of 24.13 s with a mean power consumption of 3.968 W, resulting in an energy usage of 95.232 J. In comparison, the compressed 1D CNN model completes the same workload in 18 s with an average power consumption of 3.853 W, consuming 69.336 J. Consequently, deploying the compressed model at the edge achieves a 27.58% reduction in energy consumption.

Also, the critical role of edge-AI in proposed framework became evident during an actual landslide event at the test site on 2nd October 2021, when cloud services were temporarily unavailable to issue early warning alert. In this scenario, edge autonomously generated early warning notifications based on AI prediction, shown in Fig. 4. Two warning alerts were triggered by edge 12 hours before the landslide occurrence, based on cumulative displacement thresholds. This observation shows that incorporating edge-AI in landslide monitoring systems enhances overall system reliability, a capability that cannot be achieved using a conventional IoT–cloud architecture.

References

[1] C. Liu *et al.* A new service mechanism for profit optimizations of a cloud provider and its users. *IEEE Transactions on Cloud Computing*, 9(1):14–26, 2017.

[2] T Hemalatha, Maneesha Vinodini Ramesh, and Venkat P Rangan. Effective and accelerated forewarning of landslides using wireless sensor networks and machine learning. *IEEE Sens. J.*, 19(21):9964–9975, 2019.

[3] D. Zhang *et al.* Capture and prediction of rainfall-induced landslide warning signals using an attention-based temporal convolutional neural network and entropy weight methods. *Sensors*, 22(16):6240, 2022.

[4] Amrita Joshi, Debi Prasanna Kanungo, and Rajib K. Panigrahi. Fault tolerant edge for landslide early warning system. In *2022 IEEE 19th India Council International Conference (INDICON)*, pages 1–7, 2022.

[5] S. Bagwari *et al.* Low-cost sensor based and lorawan opportunities for landslide monitoring systems on IoT platform: A review. *IEEE Access*, 2021.

[6] Partha Pratim Ray, Dinesh Dash, and Debashis De. Edge computing for internet of things: A survey, e-healthcare case study and future direction. *J. of Netw. and Computer Applications*, 140:1–22, 2019.

[7] He Li, Kaoru Ota, and Mianxiong Dong. Learning IoT in edge: Deep learning for the internet of things with edge computing. *IEEE network*, 32(1):96–101, 2018.

[8] S. Gill *et al.* AI for next generation computing: Emerging trends and future directions. *Internet of Things*, 19:100514, 2022.