Using an On-Device VLM-Based Edge Computing System for Real-Time Disaster Detection

Song, Seungmin (School: Yeouido High School)
Park, Juwon (School: Seoul International School)

This study aims to build an on-device edge computing system that leverages Vision-Language Models (VLMs) in place of the conventionally used Convolutional Neural Networks (CNNs) for real-time disaster detection, enabling not only the detection of disaster situations but also subsequent natural language-based response generation. To this end, the AIDER disaster image dataset was employed, with an ImageNet-pretrained CNN model serving as the baseline. A CNN model fine-tuned on AIDER and a zero-shot VLM, Gemma3-4B, were configured as experimental groups for comparative evaluation. Experiments were conducted under seed-controlled conditions across five different seeds to ensure reproducibility and fairness, and all evaluations were performed in an on-device environment using a Raspberry Pi 5. The results demonstrate that the VLM model outperforms the CNN model by more than approximately 10 percentage points in terms of accuracy, confirming that VLMs are sufficiently viable for real-time disaster detection. Furthermore, since the system operates on-device, it ensures efficiency in terms of both energy consumption and network dependency. Additionally, implementing the system as an edge computing platform enables the handling of natural language-based downstream tasks, which aligns well with recent trends in intelligent edge systems.

Awards Won: