

AI meets building inspections: defect diagnosis with vision language models

Defects in buildings can significantly undermine their safety and durability. Traditional inspection processes are often time-consuming, expensive, and subject to human error. To address these challenges, we developed a novel AI-based tool that combines visual imagery with textual information (vision-language) for more comprehensive defect analysis. This tool identifies visible defects, assesses their probable causes and proposes potential repair strategies. The model was trained on a custom dataset of crack images to ensure robust and context-sensitive performance.

Why the study:

- Building defects, particularly cracks, compromise structural integrity, lifespan, and occupant safety.
- Manual inspections are expensive, subjective, time-consuming, and hazardous, especially for mid- to high-rise structures.
- Existing computer vision models are limited as they mainly focus on image analysis without understanding defect context, causes, or remediation.

What we did:

- We developed a visual prototype of a vision-language model for building defects diagnosis using a custom dataset of building crack images.
- We designed the model to process both images and textual descriptions for context-aware defect analysis.

What we found:

- The model classifies different types of cracks and interpreted potential causes and remediation actions, when an image of a building crack is presented.
- It demonstrates improved diagnostic capabilities compared to traditional computer vision models that rely solely on images.
- The integration of vision and language enabled the system to provide more context-rich insights for defect assessment.

What this means:

- The model can significantly enhance building inspections by reducing inefficiencies and subjectivity.
- It offers a transformative, AI-driven approach to structural safety monitoring and maintenance planning.
- This research opens pathways for expanding to broader defect types, improving contextual reasoning, and supporting real-time assessments to assist human experts.

Lead researcher: Biyanka Ekanayake

Collaborators: Johnny Wong, Sara Wilkinson, Steve Ling (UTS), Vishal Thengane (Surrey University) and Srinath Perera (Western Sydney University)