



研究開発課題名: 災害対応のための視覚言語ベンチマーク基盤の構築

研究開発代表者: 横矢 直人 (東京大学・大学院新領域創成科学研究科・教授)

主たる共同研究者: 山野井 一輝(京都大学)・アドリアノ ブルーノ(東北大学)

グランドチャレンジへの挑戦・研究開発課題での達成目標:

物理シミュレーションと生成AIを統合した「災害×視覚と言語」ベンチマークを構築し、
実災害データの偏りと不足を補い、人と協働する災害対応AIの実現に資する評価基盤を確立する

研究概要:

- ・ マルチプラットフォーム・マルチセンサ・複数災害種を統合した視覚言語ベンチマークを設計・構築する
- ・ 土砂・洪水・津波の物理シミュレーションと生成AIを統合し、マルチプラットフォーム・マルチセンサ合成災害画像を作成する
- ・ 意味と物理の両面で整合性の取れた災害視覚データ向け自動アノテーション手法を開発する
- ・ 視覚質問応答や画像・地図上での指し示し応答によりAIの柔軟性を評価し、現場の実務者との協働で有効性と信頼性を検証する
- ・ 公開可能な合成・実災害ベンチマークを構築し、災害対応AIの研究を加速させる

想定する社会的インパクト:

- ・ 災害対応における状況認識・判断支援AIの開発と実装を加速する評価基盤となり、現場の人々に信頼されるAIの条件を明確化する
- ・ 災害時の視覚・言語情報の統合活用により、人とAIの協働と通じた高精度かつ柔軟な意思決定支援を実現する

広域視覚データ

衛星 地形 地図 空撮 地上



災害用視覚言語ベンチマーク

シミュレーション×生成AI
→ 多様な災害シナリオを
網羅する視覚言語データ



視覚言語による災害対応AI

避難所Aへの安全ルートと危険箇所は？



国道B号線が通行可能
県道Cは土石流で通行不可
D橋は冠水で使用不可



R&D Project Title: Building a Vision–Language Benchmark Framework for Disaster Response

Principal Investigator: Naoto Yokoya (Professor, GSFS, The University of Tokyo)

Co-PI: Kazuki Yamanoi (Kyoto University) • Bruno Adriano (Tohoku University)

Grand Challenge and Goal: Construct a “Disaster × Vision–Language” benchmark integrating physical simulation and generative AI, address the bias and scarcity of real–world disaster data, and establish an evaluation framework that enables human–AI collaboration in disaster response

Summary:

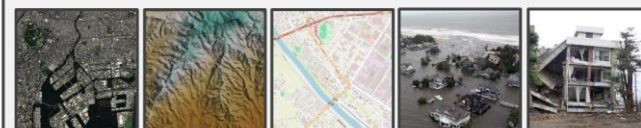
- Design and build a vision–language benchmark from integrated multi–platform, multi–sensor, multi–hazard data
- Create multi–platform, multi–sensor synthetic disaster imagery by coupling physics simulations with generative AI
- Develop auto–annotation methods for disaster imagery that ensure semantic and physical consistency
- Evaluate AI flexibility via VQA and visually grounded answers, and validate usefulness and trust with disaster–response practitioners
- Release an open synthetic + real disaster benchmark to accelerate disaster–response AI research

Social Impact:

- Accelerate disaster–response situational–awareness/decision–support AI and define what makes AI trusted by practitioners.
- Enable high–accuracy, flexible disaster decision support via vision–language integration and human–AI collaboration.

Large-scale geospatial image data

Satellite DEM Map UAV Ground



Disaster-response vision–language benchmark

Simulation × GenAI
→ Vision–language data covering diverse disaster scenarios



Vision–language disaster-response AI

Safe route to A and hazards?



Route B: open
Route C: closed (landslide)
D Bridge: closed (flooded)