＜年次報告書＞
第 1 年度（2025 年度）
日 ASEAN 科学技術・イノベーション協働連携事業（NEXUS）
日本－シンガポール「AI」研究領域

研究課題名（和文）	都市ヒートアイランド現象のモニタリングと緩和のための資源効率の高い基盤モデルの開発
研究課題名（英文）	Development of Resource-Efficient Foundation Models for Urban Heat Island Monitoring and Mitigation
研究期間	2025 年 4 月 1 日～2028 年 3 月 31 日

氏名	所属 部署　役職
横矢 直人	東京大学 大学院新領域創成科学研究科　教授
押尾 晴樹	東京電機大学 理工学部　准教授

1. 研究進捗状況と成果

① 研究構想にかかる成果

2025 年度は、高解像度都市ヒートアイランド（UHI）モニタリングのための視覚基盤モデルの開発に向けて、東京およびシンガポールを対象とした都市環境データの整備と、視覚基盤モデルの初期開発を進めた。特に、サブメートル解像度の土地被覆地図および高さデータの整備を中心に、都市の 3 次元構造、緑地・建物・道路等の配置を高精細に解析するための基盤を構築した。東京については、土地被覆地図および高さデータに加えて、QGIS UMEP/SOLWEIG を用いた Mean Radiant Temperature（MRT）データを複数時期について整備した。これにより、都市構造情報と熱環境指標を対応付けた解析・評価データを構築し、UHI 現象の要因分析や緩和策評価に向けた基盤を整えた。視覚基盤モデルについては、土地被覆分類および単眼高度推定を中心に初期モデルの構築を進めた。既存の高解像度リモートセンシングデータセットおよび関連モデルを活用し、RGB 画像から土地被覆や高さを推定し、それらを熱環境指標の推定へ接続する方向性を明確化した。これにより、次年度以降に予定している UHI 解析・緩和策評価のための視覚言語モデル（VLM）および AI エージェント研究への展開に向けた基礎を整えた。

さらに、都市環境理解や 3D 地理空間推論に関する関連研究も進展した。航空画像から 3D 地理空間推論能力を評価する Geo3DVQA を構築し、VLM が高さ、天空率、土地被覆等の 3 次元的手掛かりを統合的に扱えるかを評価する枠組みを提案した。これらの情報は、日射、放射環境、緑陰、風通し等を通じて MRT や気温分布に影響するため、UHI 要因分析・緩和策評価に重要である。また、リモートセンシング画像とストリートビュー画像を統合した都市解析モデル CityVLM を開発し、建物密度、緑地分布、道路構造、街路環境など、UHI の発生要因や緩和策に関わる都市環境情報を多視点から解析する枠組みを構築した。

② 国際研究交流の促進にかかる成果

国際研究交流については、南洋理工大学（NTU）の研究参加者と日本側若手研究者を中心に、オンラインミーティングを継続的に実施し、共同研究を推進した。ミーティングでは、都市空間理解の基盤となる 3D 大規模マルチモーダルモデル、点群データを用いた質問応答、方向依存の空間推論などについて議論し、都市環境を 3 次元的に理解するための基礎技術の検討を進めた。

特に、点群データを用いた空間推論において、観測者の位置・姿勢が欠落していると、左右・前後・奥行き等の方向依存の問いが不定になるという課題に着目し、RGB-D 映像からエゴポーズを復元し、点群を当該ポーズに整合させる手法を検討した。この共同研究は、3D 大規模マルチモーダルモデルの方向依存推論に関する成果につながった。

2. 代表的な成果事例

1. M. Tsujimoto, J. Wang, W. Xuan, and N. Yokoya, “Geo3DVQA: Evaluating vision-language models for 3D geospatial reasoning from aerial imagery,” Proc. WACV, 2026.
2. J. Wang, W. Xuan, H. Qi, Z. Chen, H. Chen, Z. Zheng, J. Xia, Y. Zhong, N. Yokoya, “CityVLM: Towards sustainable urban development via multi-view coordinated vision-language model,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 232, pp. 62-74, 2026.
3. Q. Liu, W. Xuan, J. Wang, N. Yokoya, L. Shao, and S. Lu, “Direction-aware 3D large multimodal models,” Proc. CVPR, 2026.