

課題名：
大規模地理空間データに基づくAI・基盤モデル開発とその応用

実施時期：2025年4月～2026年3月
所属機関名：産業技術総合研究所
代表者氏名：堤 千明

成果概要：
衛星画像を用いた基盤モデルの開発とその応用の観点で、だいち2号搭載の合成開口レーダ(SAR)、PALSAR-2データに基づく基盤モデルに継続事前学習技術を活用することで、だいち2号とは異なる小型SAR衛星の小規模なデータセットでも、転移学習の性能を高める基盤モデルの構築が行えることを確認した。

成果のポイント：
・小規模な小型SAR衛星データセットに対する継続事前学習効果の確認
緯度・経度・標高・時間といった地理空間と紐づいたデータは現在社会の基盤データの一つとなっているが、実利用時においてセンサ種別やタスク目的ごとに都度学習を行う場合、学習コスト(計算にかかる費用・時間、教師データ準備にかかるコストなど)が問題になる。この問題に対し、大量の衛星データを前提に基盤モデルの構築を行い、転移学習による効率的なAI利用が提案されている。本研究では少量のデータセットでも目的に即した精度を出せるAIモデルを効率的(低コスト・短時間)に作る方法の確立を目的に、産総研がJAXAとの連携により独自に整備しているPALSAR-2データアーカイブから構築したSAR基盤モデルを活用し、少量の小型SAR衛星データから継続事前学習により基盤モデルの更新を行った(図1)。転移学習(建物検知, 図2)後の性能比較から、継続事前学習の実施により、基盤モデルを利用しないケースや継続事前学習を実施しないケースに比べ高い精度が達成できることを確認し、また継続事前学習を早期に切り上げて(短い計算時間でも)高い精度を維持できることを確認した(図3)。

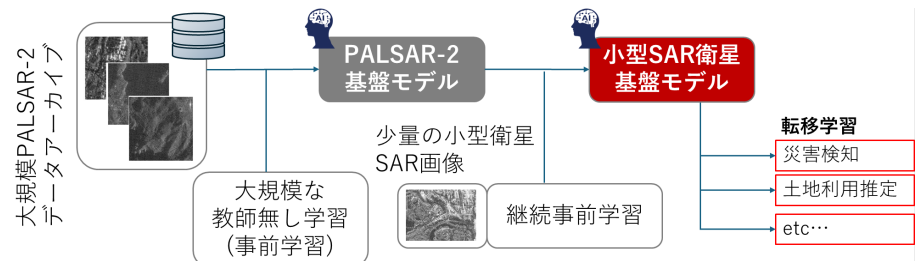


図1: PALSAR-2基盤モデルを起点に、小型SAR衛星データによる継続事前学習の概念図。PALSAR-2画像はJAXA提供(JAXA-産総研協定により)

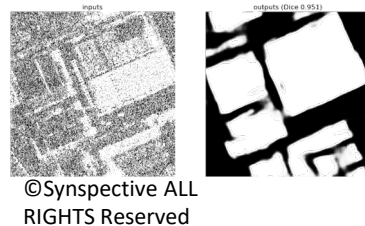


図2: 転移学習で性能評価に用いた(左)小型SAR衛星画像と(右)建物検知例。

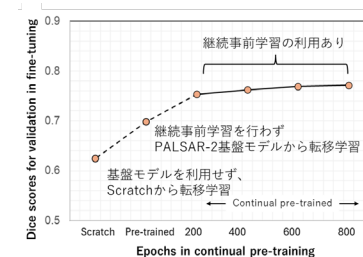


図3: 建物検知の精度で評価をした継続事前学習の効果。Scratchは基盤モデル利用無し、Pre-trainedは継続事前学習の利用無しでの精度を示す。

成果についてより詳細な情報を提供しているWebページ、発表論文などの情報：

- Sugimoto et al., Masked Siamese Autoencoder Using Multi-Frequency SAR Data, APSAR (2025)
- Kouyama et al., Evaluating the Effectiveness of Continual Pre-training of Foundation Model for Analysing Small SAR Satellite Data, ISRS (accepted)