

TSUBAME 共同利用 令和 6 年度 学術利用 成果報告書

深層学習に基づく小惑星上のボルダー検知技術の研究開発
Development of a Deep Learning-Based Technology for Boulder Detection on Asteroids

本田親寿¹, 許 竣豪², 富 宣超², 神山徹³, 吉川一朗²
Chikatoshi Honda¹, Hur Junho², Fu Xuanchao², Toru Kouyama³, and Ichiro Yoshikawa²

¹ 公立大学法人会津大学, ² 東京大学, ³ 産業技術総合研究所

¹The University of Aizu, ²The University of Tokyo, ³National Institute of Advanced Industrial Science and Technology

邦文抄録(300 字程度)

本研究では、小惑星探査画像における小さなボルダーの検出性能を向上させることを目的として、Mask R-CNN フレームワークにおける Backbone 部分を、従来の CNN 構造(ResNet-50)から、Transformer 構造をベースとする SOD-Former (Small Object Detection-Former) に置き換えるモデルを構築し、その性能を比較検証した。SOD-Former は、Transformer ブロックを 8 層組み込み、CNN 層である程度の特徴抽出した後であるが画像全体の文脈情報を活用する構造を持つ。性能評価には Recall、IoU、Overall Accuracy を用いて行い、一貫して Transformer ベースのモデルの方が安定して高い検出性能を持つことが確認された。本研究の結果は、Transformer ベースのモデルが小惑星表面の画像解析、特に小さなボルダーの検出において CNN ベースの従来手法よりも高精度かつ頑強であることを示唆するものである。

英文抄録(100 words 程度)

To accelerate the process of boulder detection that is crucial for both science and exploration mission purposes, machine learning models such as Mask R-CNN have been applied in recent studies. In this work, we aim to improve the detection of small objects by replacing the backbone of Mask R-CNN with a Transformer-based module from conventional CNN module. We confirmed that our approach shows better capability of detecting smaller features. This indicates high possibility of Transformer module for improving data analysis in asteroid exploration missions.

Keywords: Transformer, Mask-RCNN, Asteroid, Boulder detection, Remote sensing

背景と目的

近年、「はやぶさ2」や「OSIRIS-REx」、DART ミッションなどの小惑星探査ミッションにより、小惑星表面の高解像度画像が多数取得され、科学的議論を可能とさせる豊富なデータセットが整備されつつある。これにより、小惑星表面に存在するボルダーのサイズ、形状、方位などの統計的特徴を精密に解析し、天体の形成過程や進化史を比較研究することが可能になる (cf. Michikami et al., 2021)。たとえば、Ryugu と Bennu は一見するとアルベド、スペクトル、密度、さらにはリターンサンプルの CI 的な特徴を共有してもつことが見られるが、ボルダーの形状分布や方位の特性には違いが認められている。これらの違いは、それぞれの天体が異なる母天体由来であることや、形成時の物質分布の不均質性、あるいは進化経路の違いを示唆している

可能性がある。他方、小惑星リュウグウの表層画像はそのわずか 4%しか解析されず、人間が膨大なデータを処理するには時間的・経済的制約が大きい (Hirata et al., 2019)。さらに、従来の深層学習モデルでは特定の天体に対してのみ有効であり、別の対象にはそのまま適用できないといった汎用性の課題も存在する。

このようなサイエンス議論に耐えうる統計値の獲得には小惑星それぞれにおいて多数の接近画像から数万～数十万のボルダーの数え上げと輪郭の計測が求められる (Michikami et al., 2021) が、人間による手作業での解析は難しく、小惑星ごとに異なる表面特徴に対応できる頑強な自動解析技術の開発が課題となっている。

以上の観点から、本研究では一般写真において高い汎用性能が確認されている Meta 社が開発した基盤

モデル「Segment Anything(SAM)」(Kirillov et al., 2023)の惑星探査データへの応用可能性を検討することを最終目標とし、その第一段階として、Transformer 構造に基づく高性能な画像認識モデルの実装と応用に取り組んだ。特に、小惑星表面に存在するボルダー(岩塊)の自動検出を対象とし、従来の畳み込みニューラルネットワーク(CNN)と Transformer ベースのアプローチによる高精度化で Transformer の優位性の検証を目指した。

概要

本研究では、小惑星探査画像における小さなボルダーの検出性能を向上させることを目的として、Mask R-CNN フレームワークにおける Backbone 部分を、従来の CNN 構造(ResNet-50)から、Transformer 構造をベースとする SOD-Former (Small Object Detection-Former)に置き換えるモデルを構築し、その性能を比較検証した。SOD-Former は、Residual Hybrid Attention Group (RHAG)と呼ばれる Transformer ブロックを 8 層組み込んだ構造を持ち、CNN 層である程度の特徴抽出した後であるが画像全体の文脈情報を活用しながら、小さな物体の検出に最適化されている。

性能評価には Recall、IoU (Intersection over Union)、Overall Accuracy (OA)を用いて行い、小さなボルダーを見やすくするための画像のリスケージングの倍率(2 倍、4 倍)によって精度が異なることはあったものの、2 倍サイズの画像において、SOD-Former は Recall で約 2 倍、IoU で約 20 ポイント上回る性能を示した。4 倍サイズでは両モデル間の差は縮小されたが、それでも一貫して Transformer ベースのモデルの方が安定して高い検出性能を持つことが確認された。

本研究の結果は、Transformer ベースのモデルが小惑星表面の画像解析、特に小さなボルダーの検出において CNN ベースの従来手法よりも高精度かつ頑強であることを示唆するものである。SOD-Former は画像中のピクセル間の関係性を広範囲に学習できるため、複雑な地形や陰影が存在する小惑星表面においても有効であると考えられる。

結果および考察

本研究では、小惑星探査画像における小さなボルダーの検出性能を向上させることを目的として、Mask R-CNN フレームワークにおける Backbone 部分を、従来の CNN 構造(ResNet-50)から、Transformer 構造をベースとする SOD-Former (Small Object Detection-Former)に置き換えるモデルを構築し、その性能を比較検証した。

○使用データと前処理

データとして使用したのは、小惑星リュウグウに搭載された ONC-T カメラによって撮影された可視画像 228 枚である。撮影高度は 5km 以下に限定し、岩塊の形状が明瞭に視認できる画像を抽出した。解像度は 1024×1024 ピクセルであり、先行研究(Seki et al., 2024)に従い Sigma コントラストによる強調処理および Wiener フィルタによるぼやけ除去を前処理として行った。画像は、オリジナルサイズ、Half(512×512)、Quarter (256×256)に分割し、それぞれについて手動アノテーションデータ(Kouyama, 2025)と対応づけを行った。最終的に 117,676 個のアノテーションを含む学習データセットを構築し、学習とテストに 8:2 で分割した。

○モデル構造

本研究では、Mask R-CNN のフレームワークにおいて、ResNet-50 ベースのモデルと SOD-Former ベースのモデルを比較した。SOD-Former は、Residual Hybrid Attention Group (RHAG)と呼ばれる Transformer ブロックを 8 層組み込んだ構造を持ち(図 1)、CNN 層である程度の特徴抽出した後であるが画像全体の文脈情報を活用しながら、小さな物体の検出に最適化されている。

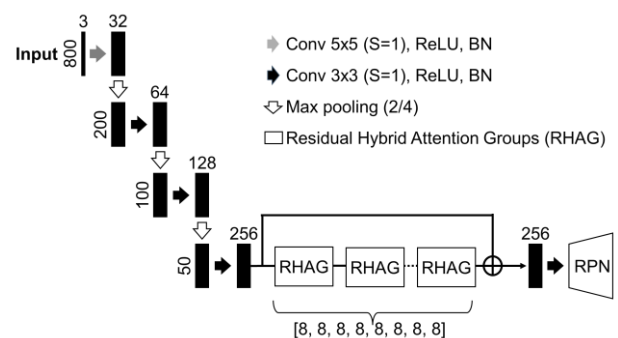


図 1 提案する RHAG ブロックを画像特徴抽出モジュール

また、RPN (Region Proposal Network)においては、小さな岩塊に対応する Anchor Box サイズの最適化を行い、実際のアノテーションのサイズ分布に合致するよう調整を加えた。

○評価指標と実験結果

図 2 に Quarter 条件でのボルダー検知結果の例を示す。

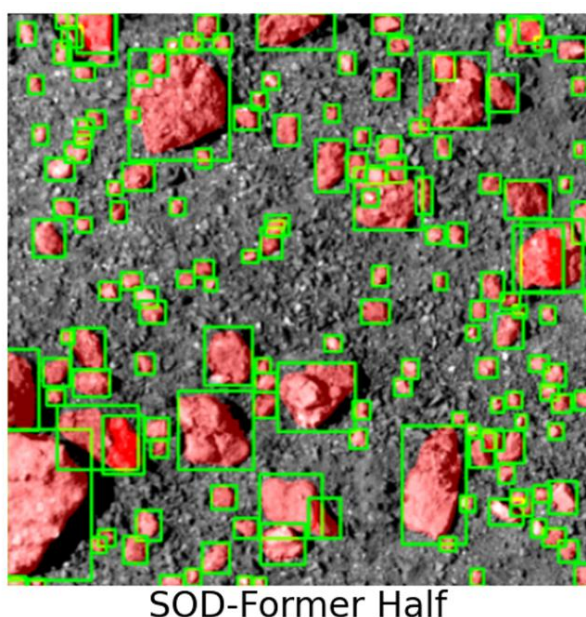


図 2 SOD-Former ベースのモデルを用いたボルダー検出結果例

性能評価には Recall、IoU (Intersection over Union)、Overall Accuracy (OA)を用いて行っている。HalfとQuarterのデータセットに対してそれぞれの評価指標の結果をまとめると以下のテーブルのようになる。

表 1 Resnet ベースと SOD-Former ベースのボルダー検知精度比較(SOD-Former を SOD-F と記す)。

	Resnet	SOD-F	Resnet	SOD-F
Dataset	Half	Half	Quarter	Quarter
Recall	28.98	59.55	68.86	72.09
IoU	46.91	68.48	85.44	83.23
OA	76.02	84.32	79.22	79.35

特に Half サイズの画像において、SOD-Former は Recall で約 2 倍、IoU で約 20 ポイント上回る性能を示した。Quarter サイズでは両モデル間の差は縮小されたが、それでも一貫して Transformer ベースのモデルの方が安定して高い検出性能を持つことが確認された。

○考察

本研究の結果は、Transformer ベースのモデルが小惑星表

面の画像解析、特に小さなボルダーの検出において CNN ベースの従来手法よりも高精度かつ頑強であることを示唆するものである。SOD-Former は画像中のピクセル間の関係性を広範囲に学習できるため、複雑な地形や陰影が存在する小惑星表面においても有効であった。

また、Quarter サイズの画像では CNN モデルとの差が縮まった背景には、極端に小さい物体では CNN でも十分な特徴抽出が可能となるケースがあることが考えられる。一方で、Half 画像のような中解像度環境では、Transformer の文脈理解能力がより効果的に機能し、顕著な性能向上につながっている。

まとめ、今後の課題

本研究では、小惑星探査画像における小さなボルダーの検出性能を向上させることを目的として、Mask R-CNN フレームワークにおける Backbone 部分を、従来の CNN 構造 (ResNet-50) から、Transformer 構造をベースとする SOD-Former に置き換えるモデルを構築し、その性能を比較検証し、Transformer 構造を取り込むことで精度向上が見られることを明らかにした。今後は、さらに他の小惑星画像への適用を進めるとともに、Segment Anything Model (SAM) などの基盤モデル技術を取り入れ、地球上の類似データと組み合わせた転移学習による精度向上を試みる予定である。