

深層学習を用いた相互作用銀河の形態分類と特徴量表現の解析

山谷 剛瑠 (高専3) 【福島工業高等専門学校】

要 旨

本研究では、相互作用銀河(interaction)を検出するためHyper Suprime-Cam(HSC) [1]アーカイブ画像から畳み込みニューラルネットワーク(CNN) [2]による二値分類器を構築し、相互作用銀河の研究に役立つ検出方法の開発を目指した。

1. はじめに

相互作用銀河は銀河進化を理解する上でとても重要である。一方で大規模サーベイでは対象数が多く、相互作用銀河の探索が課題となる。そこで本研究ではHSC画像から相互作用銀河を自動的に検出するプログラムの開発を目指す。また、相互作用銀河の理解への新たなアプローチとして深層学習を用いた特徴量の抽出を試みた。

2. 研究手法

HSCのiバンドの撮像画像のアーカイブデータをGALAXY CRUISE[3]のカタログに基づいて取得し、非相互作用銀河(non_interaction)・相互作用銀河(interaction)の2クラスに分類した。分類方法は相互作用の投票率 $P(\text{interaction})$ の値に応じて分類し、 $P(\text{interaction}) > 0.79$ の銀河をinteraction、 $P(\text{interaction}) < 0.15$ のものをnon_interactionとして定義した。3層のCNNを用いて二値分類を行いテストデータに対して混同行列と正解率(Accuracy,)各クラス再現率で性能を評価した。また、SHAP値の算出を行った。

3. 結果

CNNの学習曲線と混同行列、SHAP値を右図1, 2, 3に示した。Accuracy=0.894, interaction再現率=0.877, non_interaction再現率=0.908という結果となった。

4. 分析

図1から、学習曲線では訓練損失が低下した一方で検証損失が途中から改善しにくく、訓練と検証の間で学習に差が見られた。これはデータが少ないことによる過学習の可能性を示している。図2の混同行列からは再現率がnon_interactionに偏っていることがわかる。SHAPの出力では銀河中心部および、画像の領域内に写る対象ではない天体による寄与が目立った(図3)。そのため、目的としていた相互作用銀河の形態的特徴である、潮汐尾や銀河の形状の非対称性などがあまり明瞭ではなかった。表面輝度の低い連続的な構造は前処理や対象外の天体の影響によって埋もれやすく、CNNによって捉えにくい可能性があるため、このような構造の強調や分離が今後の課題となる。以上のことを踏まえて、现阶段での課題はサンプル不足に起因する過学習、低表面輝度の連続的な構造の検出の難しさ、対象天体周囲の他天体による影響、の3点であると考えた。

5. 今後の展望

本研究で明らかになった課題を改善するため3つの案を考案した。1つ目は画像の前処理としてセグメンテーションに基づくマスク処理[4]を行い、恒星や背景銀河など対象の銀河ではない天体が分類に影響を及ぼさないようにすることだ。これにより、モデルが対象天体由来の輝度分布を学習しやすくなるため低表面輝度の連続的な構造を検出しやすくなると思った。二つ目はサンプル数不足を解決するため、銀河衝突の数値シミュレーションデータも用いて学習を行うことだ。シミュレーションは時間発展を追うことができるため、様々なタイムフェーズでの相互作用銀河の形状を学習でき、汎化性能が向上が期待できる。3つ目はiバンド以外の撮像画像も学習に用いることである。本研究では計算時間の制約からiバンド画像のみを学習に用いたが、g, r, z, yバンドも併用し、マルチバンドの情報を取り入れることで分類性能の向上を図る。今後は、この3点を実装して効果を検証したい。

謝辞・参考文献

本研究はすばる望遠鏡HSC-SSP公開データ (PDR3) およびHSCアーカイブを用いた。形態分類にはGALAXY CRUISE Season 1 カタログ[3]を用い、Tanaka et al. (2023)[5]を参照した。

- [1] Hyper Suprime-Cam, <https://hsc-release.mtk.nao.ac.jp/doc/> (2026年1月6日閲覧)
- [2] はじめてのディープラーニング Pythonで学ぶニューラルネットワークとバックプロパゲーション, 我妻幸長 (2018年8月30日出版), SBクリエイティブ株式会社
- [3] GALAXY CRUISE Season 1 Data Release, https://galaxycruise.mtk.nao.ac.jp/en/dr_season1.html
- [4] Haigh et al. 2021 A&A 645 A107
- [5] Tanaka et al. 2023 PASJ 75 986

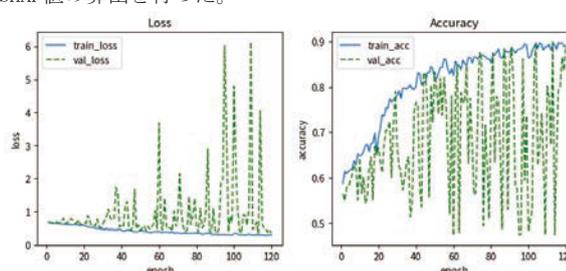


図1：学習における損失と正解率の推移 (train/val)

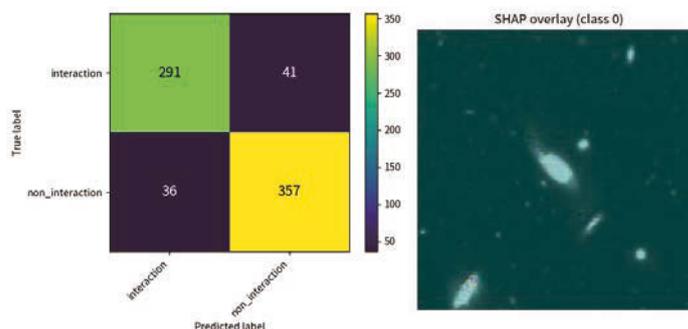


図2：衝突銀河の分類におけるCNNの混同行列

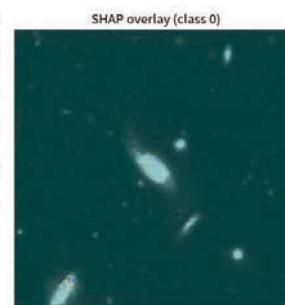


図3：CNN分類器のSHAP値 (クラス0)