“A Simulation Driven Deep Learning Approach for Separating Mergers and Star Forming Galaxies: The Formation Histories of Clumpy Galaxies in all the CANDELS Fields”

Leonardo Ferreira et al.　　arXiv:2204.05786v1

GAゼミ 2022/05/23 坂田　晴

【アブスト】

銀河の形成と進化に関する理解を深めるためには、最近大規模な合体をした銀河と、激しい星形成が起こっている銀河を区別することが重要である。本研究では、IllustrisTNG100の16万枚のシミュレーション画像を用いて、畳み込みニューラルネットワーク（CNN）に基づく機械学習の枠組みを開発した。そうした結果、Asymmetryを用いた従来の分類に比べて25％以上改善されている。また、Merger後の銀河はdisk優勢からbulge優勢の系に移行していることがわかった。

【ゴール】　Merger 銀河と星形成銀河の分類　　【オリジナリティ】 機械学習を用いている点

【ロジック】

・サンプル:IllustrisTNGのTNG100-1シミュレーションの模擬銀河を利用。

$0.5 < z < 3, M^* > 10^{9.5} M_\odot, M_{DM}/M_{Total} > 0.1$の条件で絞り込み。

過去500Myr以内に$\mu = M_2/M_1 \geq 0.25$のMergerが起きているものをpost-mergers と定義。

CANDELSのHST観測データから銀河を分類することを目的とする。

・学習:Ferreira et al. (2020) の方法を改善。

2つのネットワーク→1つのネットワークで二項分類。

Morfometryka を用い、すべての銀河のSérsic profileをフィット。

$A > 0.1$で対称銀河を除外。

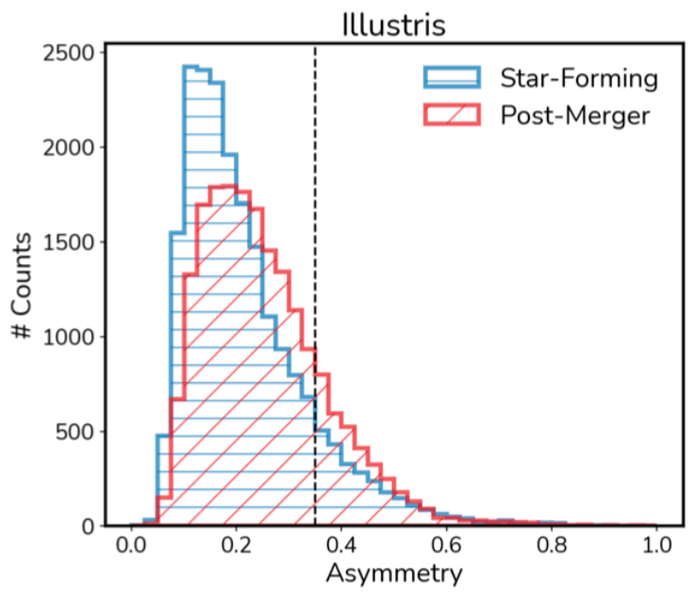


図7

【結果】

・図7:TNG100-1サンプル銀河のMorphometrykaで測定した非対称度Aの分布

・図8:学習した4つのモデルの性能指標と、シミュレーション画像に対する古典的なAsymmetry指数Aとの比較

左：学習したネットワークをデータセットに適用したROCカーブ

→Asymmetryを用いた分類に比べて少なくとも25％改善

右：CNNとAsymmetryの閾値と精度-完全性

→Asymmetryの分類で精度を出そうとすると、完全性は非常に低い。

・図9:CANDELSフィールドとcleanな場合の分類精度

・図10:背景と赤方偏移がモデルの性能に与える影響

・図12:CANDELSフィールドにおけるpost-merger銀河と星形成銀河の赤方偏移に対する分率

・図15:post-merger銀河と星形成銀河のSérsic指数分布

→Merger後は中心集中度高い

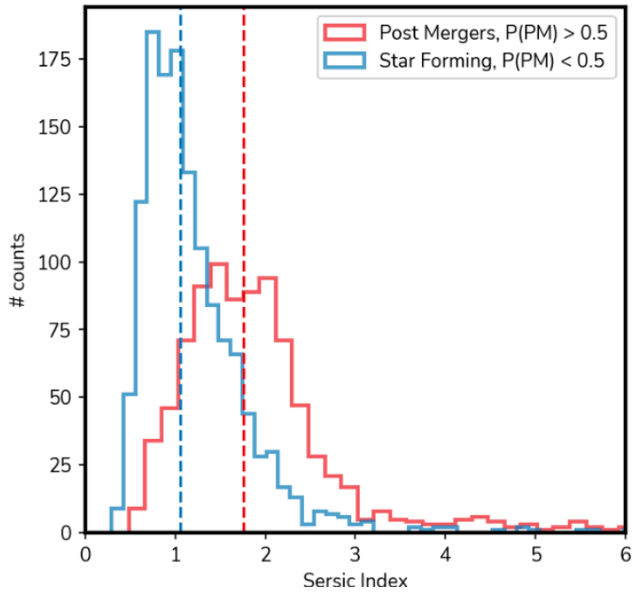


図15

【コメント】

・この方法でminor mergerについて議論するにはどの程度シミュレーションや観測の精度が必要になるのか？

・教師なし学習でクラスタリングさせた場合、どのような銀河のクラスが得られるか？

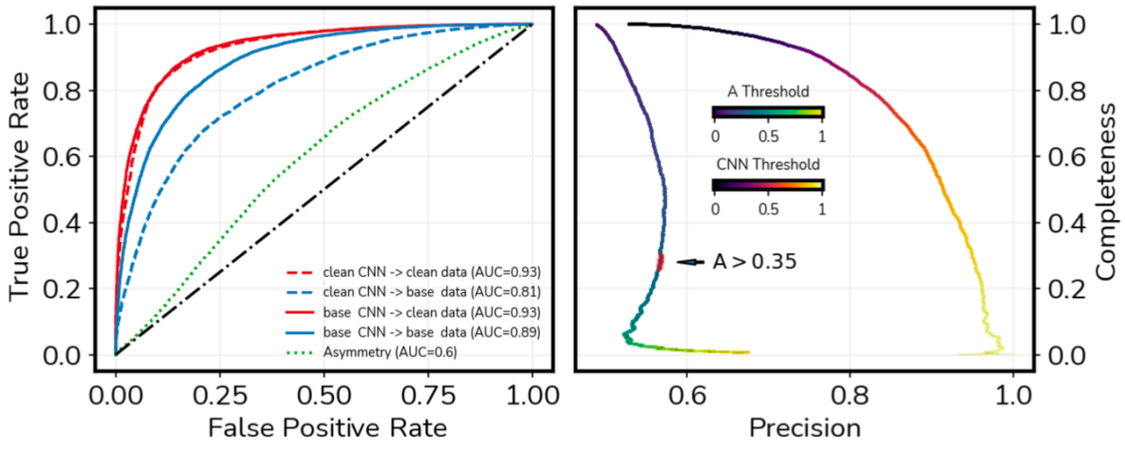

True Positive Rate
False Positive Rate
clean CNN -> clean data (AUC=0.93)
clean CNN -> base data (AUC=0.81)
base CNN -> clean data (AUC=0.93)
base CNN -> base data (AUC=0.89)
Asymmetry (AUC=0.6)
A Threshold
CNN Threshold
A > 0.35
Precision
Completeness

図8

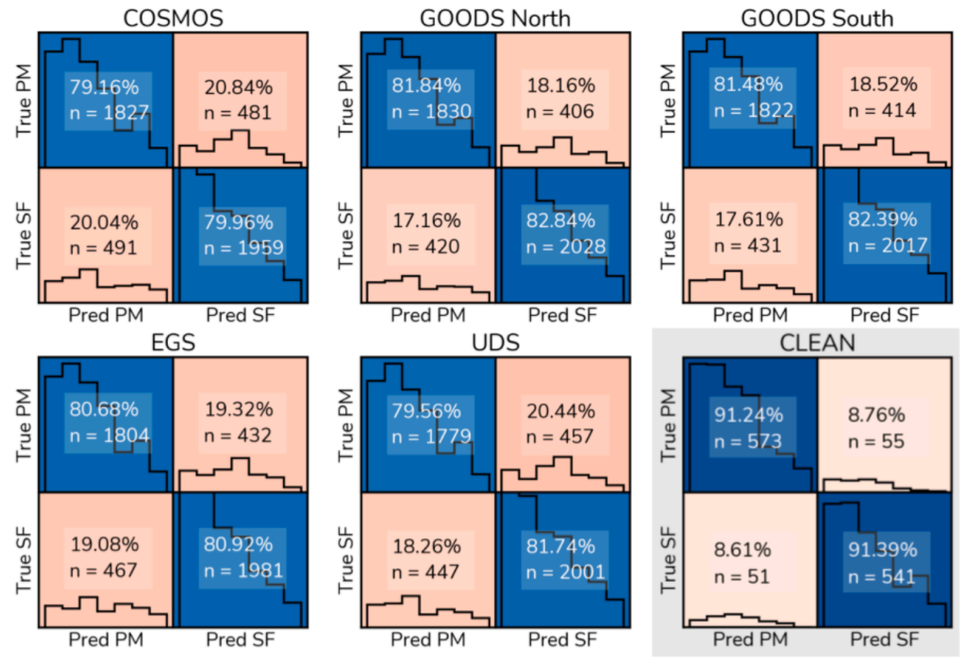

COSMOS
79.16%
n = 1827
20.84%
n = 481
20.04%
n = 491
79.96%
n = 1959
GOODS North
81.84%
n = 1830
18.16%
n = 406
17.16%
n = 420
82.84%
n = 2028
GOODS South
81.48%
n = 1822
18.52%
n = 414
17.61%
n = 431
82.39%
n = 2017
EGS
80.68%
n = 1804
19.32%
n = 432
19.08%
n = 467
80.92%
n = 1981
UDS
79.56%
n = 1779
20.44%
n = 457
18.26%
n = 447
81.74%
n = 2001
CLEAN
91.24%
n = 573
8.76%
n = 55
8.61%
n = 51
91.39%
n = 541
True PM
True SF
Pred PM
Pred SF

図9

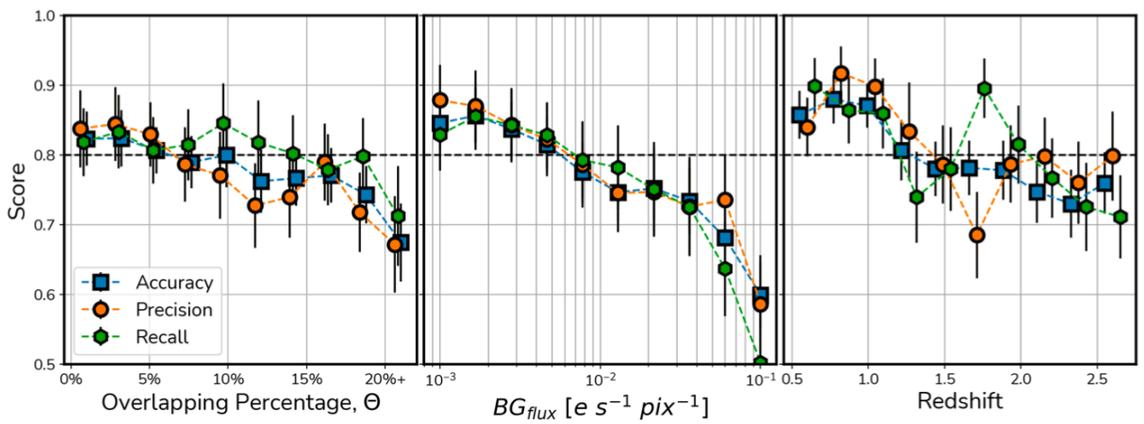

Score
Accuracy
Precision
Recall
Overlapping Percentage, Θ
BG_flux [e s^-1 pix^-1]
Redshift

図10

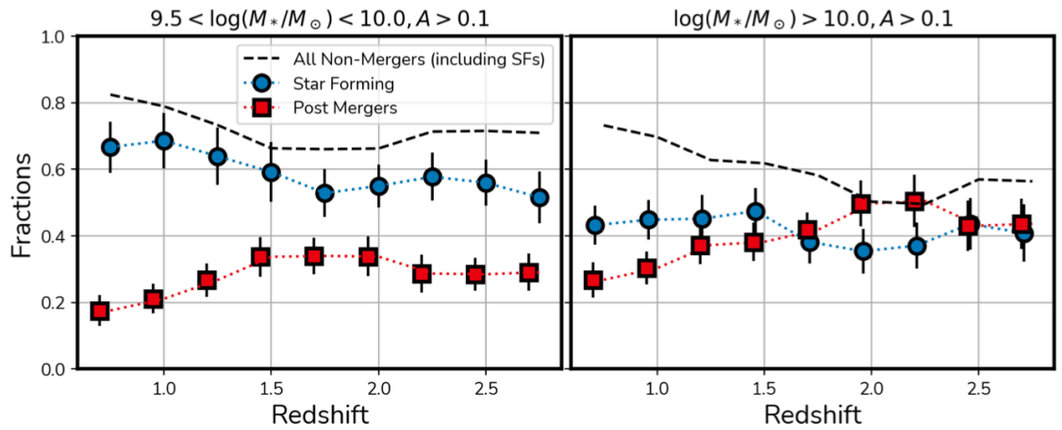

9.5 < log(M*/M⊙) < 10.0, A > 0.1
log(M*/M⊙) > 10.0, A > 0.1
All Non-Mergers (including SFs)
Star Forming
Post Mergers
Fractions
Redshift

図12