

27T

CNNを用いたTomo-e Gozenデータからの銀河抽出と分布解析

銀河学校B班: 佐藤 功基(高専4)【長野工業高等専門学校】、大野 智輝(2025年卒)【麻布高等学校】、堀 航士朗(高3)【武蔵高等学校】、齋藤 元(高3)【東大寺学園高等学校】、秋田 心菜(高3)【愛知県立旭丘高等学校】、森永 悠星(高2)【大智学園高等学校】、伊藤 弘也(高2)【長野県松本深志高等学校】、寺原 直希(高2)【川口市立高等学校】、浅田 一樹(高2)【東京都立日比谷高等学校】

研究の背景及び目的

宇宙大規模構造は、分光観測による赤方偏移測定や、超新星の光度-距離関係の解明など、観測技術の発展とともに明らかにされてきた。
本研究ではAI(深層学習)を用いて、単一バンドによる銀河画像のみから宇宙大規模構造をどこまで再現可能かという仮想的な課題を設定した。そこで、機械学習による画像処理技術として一般的なCNN(畳み込みニューラルネットワーク)を用いて観測領域内における銀河の位置検出を行い、さらに銀河までの距離推定の可能性を検討することを目的とした。
図1にSDSSによる銀河の三次元分布を天の赤道を含む平面で切った断面図を示す。宇宙の大規模構造を調べることで、宇宙や銀河の進化の解明につながる。

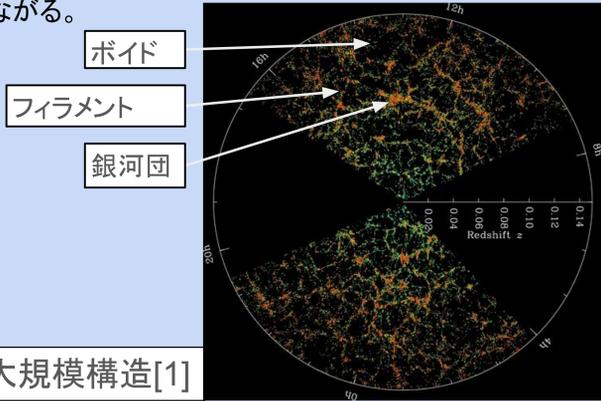


図1 宇宙の大規模構造[1]

観測

日時:2025年3月26日 20:43~21:23(JST)
使用装置:東京大学木曾観測所105cmシュミット望遠鏡(図2)+Tomo-e Gozen
観測領域:おとめ座・かみのけ座領域(図3)。(銀河が多く密集することが知られている)
観測設定:1ポジション当たり露光0.5s×20枚×6回(CMOSセンサー間のギャップを補完するためのデザイン観測)

望遠鏡及び観測装置の情報 [2]

有効口径:1,050 mm、焦点距離:3,300mm、F値:3.1、視野:直径9度、センサーフォーマット:2,000x1,128画素/チップ(感光画素)、センサーチップ数:84 チップ、画素サイズ:19 μm、画素スケール:1.189 秒角、視野角:39.7分角×22.4分角×84 チップ=20.8平方度、Highゲイン設定での限界等級:(暗夜、PSF FWHM = 3 秒角、測光開口 = 2 x PSF FWHM、S/N = 10)、限界等級:15.9 等級(0.1 秒露光)、17.7等級(1秒露光)、19.1等級(10秒露光)



図2 東京大学木曾観測所105cmシュミット望遠鏡

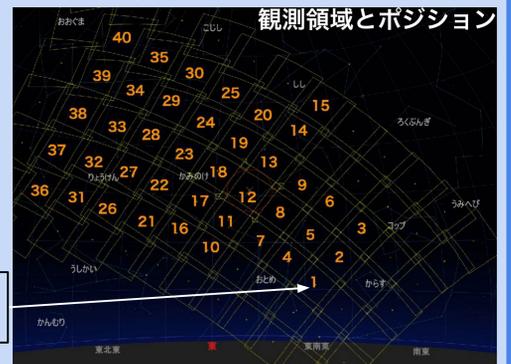


図3 観測領域とポジション

解析

- ①:デザイン観測が実施済みである1視野分のデータを256×256pixelの単位領域に分割
 - ②:ポジション33、34、39を教師画像に用い、目視により銀河を含む画像と含まない画像を選び、学習データを作成し、CNN(図4)に学習させる。データ数は元のデータが69、拡張後のデータが1035である。
 - ③:別の領域で256×256pixelの単位領域に分割し、CNNを用いて画像上の銀河を抽出。
 - ④:単位領域同士をつなげ合わせてヒートマップに表す。
 - ⑤:③を基に銀河までの距離を推定。
 - ⑥:三次元分布として表現。
- ※④、⑤は本研究では未実施。

CNNとは? [3]

CNN(Convolutional Neural Network):畳み込みニューラルネットワーク。画像の特徴を認識し、画像を分類する。

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 256, 256, 3)	0
conv2d (Conv2D)	(None, 124, 124, 64)	5,248
batch_normalization (BatchNormalization)	(None, 124, 124, 64)	256
dropout (Dropout)	(None, 124, 124, 64)	0
conv2d_1 (Conv2D)	(None, 60, 60, 128)	284,928
batch_normalization_1 (BatchNormalization)	(None, 60, 60, 128)	512
dropout_1 (Dropout)	(None, 60, 60, 128)	0
conv2d_2 (Conv2D)	(None, 30, 30, 256)	295,168
dropout_2 (Dropout)	(None, 30, 30, 256)	0
global_average_pooling2d (GlobalAveragePooling2D)	(None, 256)	0
dense (Dense)	(None, 128)	32,896
dense_1 (Dense)	(None, 64)	8,256
dense_2 (Dense)	(None, 1)	65

Total params: 547,329 (2.09 MB)
Trainable params: 446,945 (2.09 MB)
Non-trainable params: 384 (1.38 MB)

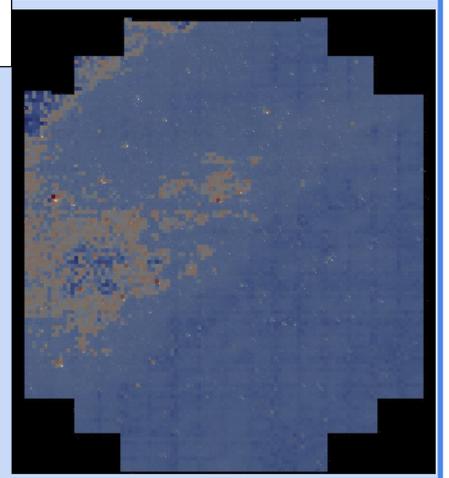
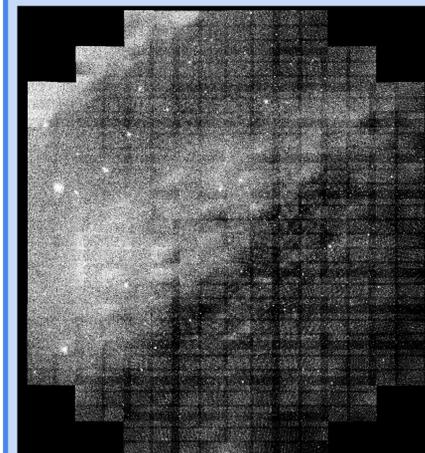


銀河が存在する確率
0 ←→ 1
図4 CNNのしくみ

結果・考察1

図5 銀河の抽出結果の一例

図5:ポジション2における銀河抽出の結果
背景光度にバイアスされた学習になっている可能性が示唆される(図6)。



銀河が存在する確率
0 → 1

図6 図5の領域における背景光度

再解析

- 背景光によるバイアスを取り除くために、以下の学習の改善を試みた。[4]
- (1) 選別銀河画像の改善(バルジが明瞭で、かつ直径が30ピクセル程度以上のものにする。)
 - (2) 学習回数:15→10
 - (3) 背景光度調整:教師画像の背景光度を(-5%,+10%)の範囲でランダムに調整し、CNNが背景光度をノイズと認識するようにする。(図7、8)
- ・図8及び図9に背景光度を操作する前後のMAE(Meaning Absolute Error) Curveを示す。

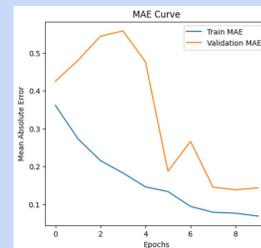


図7 改善前のMAE Curve

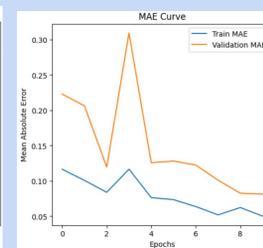


図8 背景光度に操作を加えた後のMAE Curve

結果・考察2

図9、10:再解析による結果
一定程度学習が改善されたと考えられる。

図9 目視で銀河を抽出した結果

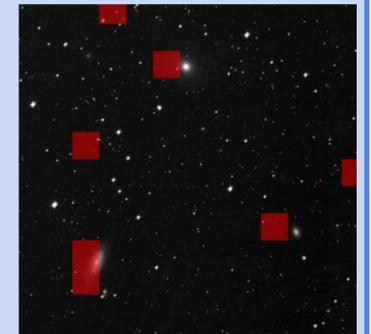
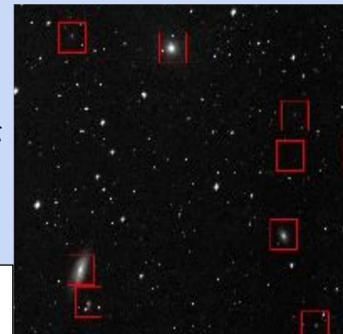


図10 改良したCNNにより銀河を抽出した結果(出力が0.8を超えたものに印を付している)

まとめと今後の展望

本研究では天球面上の銀河の分布を推定を試みる事ができた。しかし、CNNの改良はまだ途中であり、精度のさらなる向上が課題である。また、今後は銀河までの距離を推定し、3次元的な銀河分布を再現したい。

参考文献

- [1]宇宙の大規模構造 <https://astro-dic.jp/large-scale-structure-of-the-universe/>
- [2]The Tomo-e Gozen wide field CMOS camera for the Kiso Schmidt telescope <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10702/2310049/The-Tomo-e-Gozen-wide-field-CMOS-camera-for-the/10.1117/12.2310049.short>
- [3]Krizhevsky et al., 2012:ImageNet Classification with Deep Convolutional Neural Networks
- [4]CNNのデータ拡張による過学習防止 <https://zenn.dev/kusunoki576/articles/bbbd470dc3b6ec>

謝辞

本研究にご協力いただいた東京大学大学院理学系研究科天文学教育研究センター特任助教の今井正堯先生をはじめとした東京大学木曾観測所並びに銀河学校2025のスタッフの皆様がこの場を借りて深く御礼申し上げます。