

## 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】

### 要旨

我々は東京大学木曾観測所105cmシュミット望遠鏡および広視野観測装置Tomo-e Gozenにより取得された観測データを用い、CNNを活用して銀河を自動抽出することで、既知の宇宙大規模構造の再現を試みた。

#### 1.目的

宇宙大規模構造は、観測技術の発展により明らかにされてきた。本研究では、観測技術の高度化に先立って AI(深層学習)が確立されていた場合、銀河画像のみから宇宙大規模構造の再現は可能かという課題を設定した。そこで、CNNによる銀河の位置検出と、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。

#### 2.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シュミット望遠鏡及びTomo-e Gozen

観測領域:銀河が密集しているおとめ座・かみのけ座周辺の銀河密集領域を網羅的に観測した(図1)。露光条件は、過去のTomo-e Gozen観測を参考に0.5s×20枚とし、CMOSセンサー間のギャップを補完するためデザイン観測を行った。図1に示す1ポジションあたりの観測時間は約3分である。

#### 3.解析

取得した全画像データを256×256 pixelの単位領域に分割した。ポジション39において、目視により抽出した、銀河を含む130枚の画像と、ほぼ同数の銀河を含まない画像から学習データを作成し、Google Colaboratory上で構築したCNNにこれらを学習させた。第一段階として、観測領域全体について単位領域ごとに銀河の有無を確率スコアで評価し、その結果を可視化する。第二段階として、距離既知の銀河カタログを教師データとして、抽出した銀河の距離をCNNを用いて推定し、銀河の3次元的な分布を調べる。

#### 4.結果・考察

結果1:CNNを用いて画像上から銀河を抽出し第一段階の解析を行った結果、右の画像のようになった(図2)。なお、右の画像はポジション3の例であり、今回の解析では銀河がフラットに分布しているなどといった結果が得られた。

考察1:結果から、CNNの精度が悪く、正しく銀河を抽出できていないことがわかった。背景光度と結果で得られた分布に関係性が見られたため、過学習によりCNNが銀河を構造ではなく、背景光度に基づき抽出した可能性が示唆される。そこで、過学習を解消するために、次に示す3つの方法で学習の改善を試み、再度解析を行った。(1)教師データの背景光度を少し揺らし、CNNが背景光度をノイズと認識するようにする。(2)学習回数を減らす。(3)教師データに用いるデータの選別を厳しくする。

結果2:CNNの精度を検証するため、バルジが明確で、直径が30ピクセル程度を超えているという条件から目視で銀河を抽出した結果(図3)と、改良したCNNにより銀河を抽出した結果(図4)を示す。なお、図4ではCNNの出力が0.8を超えるものに印を付けている。

考察2:図3と図4の比較から、学習の改善により、CNNによる銀河の抽出の精度が上がったと考えられる。

#### 5.まとめと今後の展望

CNNの精度のさらなる向上が課題である。また、本研究では解析で示した、第一段階の天球面上の銀河の分布の推定にとどまったため、今後は第二段階を行い3次元的な分布を推定したい。

#### 6.参考文献

[1]CNNのデータ拡張による過学習防止

<https://zenn.dev/kusunoki576/articles/bbbd470dc3b6ec>(2025年12月28日閲覧)

#### 7.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亮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のスタッフの皆様に深く御礼申し上げます。



図1 番号をふった範囲を観測した。

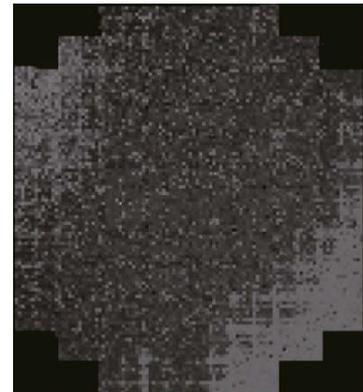


図2 CNNによって銀河を抽出した結果  
銀河がある確率が、黒で低く、白で高い



図3 改良した手動抽出データ

図4 改良したCNNの出力