

宇宙線反粒子識別を対象とした機械学習の応用と根拠の可視化

今福 拓海^{*1}, 石川 博^{*1}, 荒木 徹也^{*2}, 山本 幸生^{*3}, 福家 英之^{*3}
清水 雄輝^{*4}, 和田 拓也^{*5}, 中上 裕輔^{*5}

Application of Machine Learning and Visualization of Evidence for Cosmic-Ray Antiparticle Identification

IMAFUKU Takumi^{*1}, ISHIKAWA Hiroshi^{*1}, ARAKI Tetsuya^{*2}, YAMAMOTO Yukio^{*3}
FUKE Hideyuki^{*3}, SHIMIZU Yuki^{*4}, WADA Takuya^{*5}, NAKAGAMI Yusuke^{*5}

Abstract

The General AntiParticle Spectrometer (GAPS) aims to search for antideuterons originating from dark matter through highly sensitive observation of cosmic ray antiparticles. In GAPS, the use of machine learning is being considered for cosmic ray antiparticle identification, which requires high rejection power and identification accuracy. In a previous study, a three-dimensional convolutional neural network (CNN) model was used to achieve high rejection power for simulated data under limited conditions where the incident angle and position of the antiparticle are fixed. In this study, in order to seek the explanatory power of discrimination by analyzing the factors that contribute to the high accuracy of machine learning, the parts that the CNN model focuses on in discrimination are visualized, and the cases of incorrect discrimination are analyzed.

Key Words : 3DCNN, Grad-CAM, GAPS

概要

GAPS (General AntiParticle Spectrometer) は宇宙線反粒子の高感度観測によって暗黒物質を起源とする反重陽子の探索を目指す気球実験計画である。GAPS では、高い識別率や正確性が求められる宇宙線反粒子識別に対して機械学習の活用が検討されている。先行研究では、三次元の CNN (畳み込みニューラルネットワーク) モデルに用いることで、反粒子の入射角と入射位置を固定した限定的な条件下でのシミュレーションデータに対して高い識別精度が確認されている。本研究では機械学習の高い精度の識別における要因の分析によって識別の説明性を求めるため、CNN モデルが識別において注視している部分を可視化し、誤って識別したケースを分析した。

* 2021 年 11 月 30 日受付 (Received November 30, 2021)

^{*1} 東京都立大学大学院システムデザイン研究科 (Graduate School of System Design, Tokyo Metropolitan University)

^{*2} 群馬大学理工学部 (School of Science and Technology, Gunma University)

^{*3} 宇宙科学研究所学際科学研究系 (Institute of Space and Astronautical Science, Department of Interdisciplinary Space Science)

^{*4} 神奈川大学 (Kanagawa University)

^{*5} 青山学院大学大学院 (Aoyama gakuin University)

1 はじめに

GAPS (General AntiParticle Spectrometer) と呼ばれる宇宙線反粒子の高感度観測による反重陽子の検出によって暗黒物質の解明の手がかりの獲得を目指す気球実験計画がある [5, 10–12]. 反重陽子は暗黒物質の候補粒子の対消滅・崩壊によって生成される粒子の一つとして考えられており [3], 予想存在量が極微小で現在まで未発見であるが, 低エネルギー域で 1 イベントでも検出できれば暗黒物質起源の可能性が高い. GAPS では, 反粒子の検出に, シリコン半導体 (Si(Li)) 検出器群とそれを二重に取り囲むプラスチックシンチレーションカウンタ群 (TOF) [8, 14] から構成される GAPS 測定器を用いる. シリコン半導体検出器群は, 8 つのセグメントに分かれたシリコン半導体検出器が $12 \times 12 \times 10$ で並べられている. 反粒子が測定器に入射すると 2 層の TOF カウンタを通過後, 積層されたシリコン検出器群を通過中にエネルギー損失により減速・捕獲され, Si と励起エキゾチック原子を構成する. 励起エキゾチック原子はすぐ崩壊し, その過程で特性 X 線や反粒子と Si 原子核との対消滅によるパイ中間子や陽子のハドロン群が放出される [9]. 励起エキゾチック原子の崩壊過程やパイ中間子やロー中間子の生成数は捕獲された反粒子の種類に依存するため, これらの飛跡を測定することで入射反粒子を同定することができる [2, 11]. 反重陽子の検出において, 同様にエキゾチック原子を形成しうる反陽子が比較対象であり, 反重陽子に対する反陽子の存在比は 10^4 以上とされている. 反重陽子と反陽子の存在量の差の大幅な抑制と求められる高い識別能力に対し, GAPS ではシミュレーションによる反重陽子, 反陽子の検出器内の飛跡データと NN (ニューラルネットワーク) を用いた粒子識別法に着手している [13]. 和田らの研究 [13] では, 反重陽子と反陽子の検出器に対する入射角を固定したシミュレーションデータに対して 3 次元の CNN (畳み込みニューラルネットワーク) によって高い精度での識別を実現している. しかしながら, 高い精度の識別が実現できている要因については研究がなされていない. 機械学習ではモデルの説明性についての研究も行われている [6]. そこで本研究では, [13] の機械学習の高い精度の識別における要因の分析やデータから物理的事象の根本的な特徴を取り出すことによって識別の説明性を求めることを目指し, Grad-CAM (Gradient-weighted Class Activation Mapping) による CNN モデルの識別根拠部分の可視化を行い, 識別器が誤って識別した際の原因について調べた. また, 可視化により識別器が入射経路や対消滅後の粒子の広がりや形状を考慮していないことが分かった, そこで, シミュレーションデータ内の粒子の通過した位置情報のみで識別が可能かどうか計算機内での実験を行った.

本論文は以下の構成に従う. 2 章では, 先行研究 [13] について述べる. 3 章では, Grad-CAM による CNN モデルの識別根拠部分の可視化について述べる. 4 章では, シミュレーションデータの非 0 値の値を統一したデータによる識別器について述べる. 5 章では, 本論文のまとめと今後の課題について述べる.

2 先行研究

GAPS では宇宙線反粒子の識別器として NN を用いた識別手法を開発しており, 特に 3 次元の CNN モデルを用いることで高精度な識別が可能であることがわかっている [13]. 以下に GAPS での先行研究におけるデータセット, CNN モデルのネットワーク構成について述べる.

2.1 データセット

入力データとして反陽子と反重陽子それぞれの入射や放出されるハドロン群などのシミュレーションによってシリコン半導体検出器群に記録されるエネルギー損失値の三次元データを用いる. シリコン半導体検出器は 1 層に 12×12 で並べられたものが 10 層重なっており, 三次元配列の計 1440 チャンネルにそれぞれエネルギー

ギー損失値が記録される。ただし、シリコン半導体検出器の8つのセグメントは層によって向きが異なり配列に整形出来ないため、1つにまとめたものとして扱っている。反粒子の入射角は測定器に対して鉛直方向下向きに固定したデータと、測定条件としては現実的である測定器から見て上半球から一様に入射するデータが用意されている。以降本稿では、測定器に対して鉛直方向下向きに固定したデータを入射角固定、測定器から見て上半球から一様に入射するデータを入射角ランダムと呼ぶ。

2.2 ネットワーク構成

表1に先行研究における3次元CNNモデルのネットワーク構成を示す。モデルは3次元の畳み込み層(conv), プーリング層(max-pooling), 全結合層(dense)で構成されており、プーリング層と全結合層の間にバッチ正規化, ドロップアウト, 平滑化を行っている。各畳み込み層には出力にReLU関数, 最終出力層にはsigmoid関数, 損失関数には二値交差エントロピー, 最適化アルゴリズムにはAdamを用いている。

表1 先行研究における3次元CNNモデルのネットワーク構成

layer	kernel	stride	output size
input			12×12×10
conv	3×3×3	1×1×1	12×12×10×64
conv	3×3×3	1×1×1	12×12×10×128
conv	3×3×3	1×1×1	12×12×10×256
max-pooling	2×2×2	2×2×2	6×6×5×256
dense			512
dense			256
dense			128
dense			64
dense			32
output			1

3 Grad-CAMによるCNNモデルの識別根拠部分の可視化

識別精度向上のため、先行研究で用いられているCNNモデルが誤って識別したシミュレーションデータのどこを注視しているのかをGrad-CAMによって可視化した。識別器の構築, 学習, およびGrad-CAMはKeras [4]とTensorflow [1]を用いた。先行研究では反重陽子と反陽子を各200万件用いているが、本研究ではその内の各2万件を使用している。また、可視化に使用するデータは全て入射角固定とし、識別器の学習データは入射角固定の36000件(反重陽子:18000件, 反陽子:18000件)とした。データには1件あたり反重陽子もしくは反陽子の1イベントが記録されている。

3.1 Grad-CAM

Grad-CAM(Gradient-weighted Class Activation Mapping) [7]は、畳み込みニューラルネットワーク(CNN)の入力画像に対するクラス分類の根拠部分をヒートマップにして可視化する手法である。三次元データを入力としたGrad-CAMによるヒートマップ $H_{Grad-CAM}^c$ を求める式を以下に示す。

$$\alpha_l^c = \frac{1}{Z} \sum_i \sum_j \sum_k \frac{\partial y^c}{\partial A_{ijk}^l} \quad (1)$$

$$H_{Grad-CAM}^c = ReLU \left(\sum_l \alpha_l^c A^l \right) \quad (2)$$

ここで、式 (1) の y は出力層のあるクラス c における出力、 l は畳み込み層の特徴マップの番号、 A は l 番目の特徴マップ、 α は特徴マップに対する重みである。一般的に特徴マップは最も出力に近い畳み込み層のものが用いられる。式 (2) で特徴マップ A と特徴マップに対する重み α を掛け合わせたものに ReLU 関数を適用したものをヒートマップ $H_{Grad-CAM}^c$ としている。ReLU 関数を用いるのはクラス分類に影響するのは正の勾配であると考えられるためである。

3.2 正しく識別したシミュレーションデータとその Grad-CAM

図 1, 2 に先行研究における CNN モデルでそれぞれ正しく識別された反重陽子と反陽子のデータの粒子の入射経路に合わせた断面図 (左) と Grad-CAM の断面図 (中央), 及びそれら 2 枚をヒートマップとして重ね合わせたもの (右) を示す。ここで、入射経路に合わせた断面図とは、入射角固定のデータの大きさが $12 \times 12 \times 10$ である三次元配列において (6,6,0) から (6,6,10) に向かって入射しているため、(6,0,0) ~ (6,12,10) の範囲を切り出した断面 (縦 × 横 = 12×10) のことを指す。また、左図の白い部分が粒子のエネルギー損失値 [MeV] の大きさを表しており、粒子は左から右に向かって入射している。Grad-CAM はヒートマップの赤い部分が識別器が注視している部分を表しており、図 1, 2 から、識別器は反重陽子と反陽子に共通して、入射経路の末端である最もエネルギー損失値が大きい部分を注視していることがわかる。また、反重陽子と反陽子のデータから、対消滅などが起こるまでの反粒子の入射経路がおおよそ反重陽子の方が反陽子より長い。これは人の目でも判断が容易であり、入射角を固定したデータによる識別器の精度が高い要因であると考えられる。

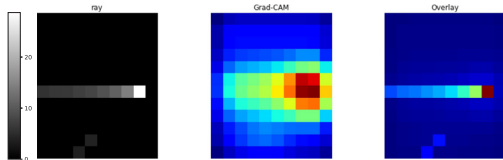


図 1 正しく識別された反重陽子のデータの断面図 (左) と GradCAM の断面図 (中央), 及びそれらの重ね合わせ (右)

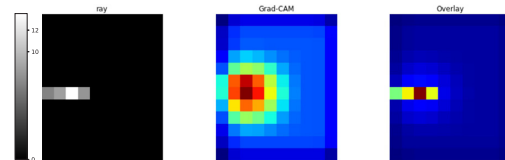


図 2 正しく識別された反陽子のデータの断面図 (左) と GradCAM の断面図 (中央), 及びそれらの重ね合わせ (右)

3.3 誤って識別したシミュレーションデータとその Grad-CAM

今回の実験において、入射角固定の場合では、誤って反重陽子と判定された反陽子と、誤って反陽子と判定された反重陽子は今回の実験のテストデータ4000件のうち各1件ずつであった。図 3 に誤って反重陽子と識別された反陽子のデータの粒子の入射経路に合わせた断面図 (左) と Grad-CAM の断面図 (中央), 及びそれら 2 枚をヒートマップとして重ね合わせたもの (右) を示す。3.2 節で述べたように反重陽子と反陽子には入射経路の長さの違いがあり、識別器は対消滅などが起きている入射経路の末端の、1 件のデータ内における最もエネルギー損失値が大きい部分を注視している。図 3 は、反陽子のデータであり、入射経路の長さが比較的短いはずであると考え、最もエネルギー損失値が大きい部分が本来入射経路の末端より先の入射経路上に現れてしまったため、識別器は反重陽子の入射経路の長さで誤って識別してしまったと考えられる。図 3 では、最もエネルギー損失値が大きい部分が本来入射経路の末端である対消滅などが起きた部分より先の入射経路上にあるため、識別器は反重陽子の入射経路の長さで誤って識別してしまっている。

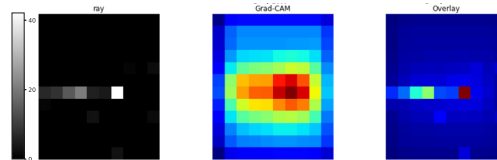


図3 誤って反重陽子と識別された反陽子のデータの断面図(左)と Grad-CAM の断面図(中央), 及びそれらの重ね合わせ(右)

図4に誤って反陽子と識別された反重陽子のデータの粒子の入射経路に合わせた断面図(左)と Grad-CAM の断面図(中央), 及びそれら2枚をヒートマップとして重ね合わせたもの(右)を示す. Grad-CAMを確認すると図4の入射経路の断面図((6,0,0) ~ (6,12,10))では何も確認できない部分に識別器が注視している. そこでデータの合計12枚の断面図(0,0,0) ~ (0,12,10), ..., (12,0,0) ~ (12,12,10)を全て足し合わせて識別器が注視している部分を可視化した(図5). 図5に誤って反陽子と識別された反重陽子のデータの粒子の入射方向と垂直に断面図を取り, 足し合わせたもの(左)と, 左図と同様にして Grad-CAM の断面図を足し合わせたもの(中央), 及びそれら2枚をヒートマップとして重ね合わせたもの(右)を示す. 図5によると入射経路の末端以外の部分にシミュレーションデータ内でのエネルギー損失値の最大値があり, Grad-CAMで識別器が最も注視している. 本来CNNモデルを導入した意図としてシミュレーションデータの入射経路や対消滅後の粒子の広がりなどの空間的な情報を識別器に認識させることが挙げられるが, 図5より識別器はエネルギー損失の最大値に注視しており入射経路などの空間的な情報を認識していないことが示された. そこで, 4章では, 識別器にエネルギー損失の最大値を注視せずに空間的な情報を認識させるために, シミュレーションデータ内のエネルギー損失値が0でない値を同じ値に統一して識別器を生成し, 精度を確認した.

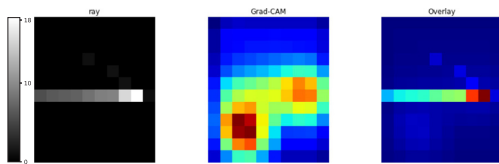


図4 誤って反陽子と識別された反重陽子のデータの断面図(左)と Grad-CAM の断面図(中央), 及びそれらの重ね合わせ(右)

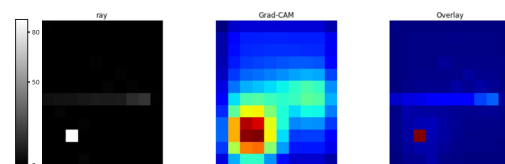


図5 誤って反陽子と識別された反重陽子のデータの断面図を足し合わせたもの(左)と Grad-CAM の断面図を足し合わせたもの(中央), 及びそれらの重ね合わせ(右)

4 シミュレーションデータの非0値の値を統一したデータによる識別器

3.3節で示したように, 識別器は入射経路や対消滅後の粒子の広がり形状を考慮しておらず, 各シミュレーションデータ内のエネルギー損失値が最大の部分に注視していた. そこで, シミュレーションデータのエネルギー損失値が非0である値を同じ値に統一し, 粒子の通過した位置情報のみで識別できないか実験した.

4.1 実験条件

データセットは2.1節に示す入射角固定, 入射角ランダムをそれぞれ使用し, エネルギー損失値の非0値を1.0に統一した. データは学習用, 検証用, 評価用として各36000件, 4000件, 4000件とした. データの反重陽子と反陽子の比率は全て約1:1である. 識別器のネットワーク構成は2.2節に示す構成と同じである.

4.2 実験結果

表 2 に非 0 値を元のままにしたデータを用いた識別器と、非 0 値を 1.0 に統一したデータでの識別器の正解率と、反重陽子を正 (positive) とした場合の適合率と再現率を示す。表 2 より、正解率、適合率、再現率はどれも元のデータを使用している場合より非 0 値を統一したデータの方が下がっている。これはエネルギー損失値の大きさの情報を失っているため自明である。しかし、位置情報のみであっても一定の精度で識別できており、入射角固定において、元のデータの場合で誤って識別した 2 件のデータの内、2 件とも正しく識別されていることを確認した。非 0 値統一の効果が明瞭に表れた例として図 6 に図 4 に示した誤って反陽子と識別した反重陽子のデータの非 0 値を 1.0 に統一した後の断面図 (左) と Grad-CAM の断面図 (中央)、及びそれらの重ね合わせ (右) を示す。図 6 によると本実験の識別器は粒子の通過した位置情報の全体を根拠とした上で識別している。また、入射角ランダムにおいても、元のデータの場合で誤って識別した 250 件のデータの内、98 件が正しく識別されていた。

表 2 非 0 値を統一した場合の識別精度の比較

	正解率	適合率	再現率
元のデータ (入射角固定)	0.9995	0.9995	0.9995
非 0 値を統一したデータ (入射角固定)	0.9943	0.9956	0.9931
元のデータ (入射角ランダム)	0.9390	0.9497	0.9269
非 0 値を統一したデータ (入射角ランダム)	0.8167	0.8203	0.8113

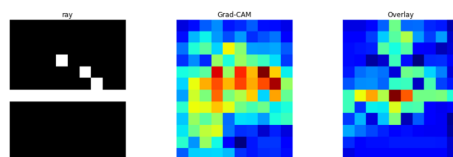


図 6 図 4 に示した誤って反陽子と識別した反重陽子のデータの非 0 値を 1.0 に統一した断面図 (左) と Grad-CAM の断面図 (中央)、及びそれらの重ね合わせ (右)

5 まとめ

本研究では、識別精度向上のために、Grad-CAM による CNN モデルの識別根拠部分の可視化とエネルギー損失値の非 0 値の統一によって粒子の通過した位置情報のみで識別できないか実験した。実験によって粒子の通過した位置情報のみでも一定の精度で識別可能であったが、従来の手法に比べ精度が落ちたことから、粒子の位置情報とエネルギーの大きさの双方が識別に利用できることが分かった。今後は、エネルギー損失値の大きさと位置情報に適切な重みをかけて識別器を学習させる手法や、入射角がランダムなシミュレーションデータから直線的な入射成分と対消滅後の粒子の拡散成分に分けクラスタリングできないかを検討する予定である。

謝辞

本研究は、東京都立大学傾斜的研究費 (全学分) 学長裁量枠国際研究環支援, JSPS 科研費 20K12081, 20K04002, 17H01136, 住友財団基礎科学研究助成費, 三菱財団科学研究助成費, 及び NASA/APRA 経費による。

参考文献

- [1] ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z., CITRO, C., CORRADO, G. S., DAVIS, A., DEAN, J., DEVIN, M., ET AL. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467* (2016).
- [2] ARAMAKI, T., CHAN, S., CRAIG, W., FABRIS, L., GAHBAUER, F., HAILEY, C., KOGLIN, J., MADDEN, N., MORI, K., AND YU, H. A measurement of atomic x-ray yields in exotic atoms and implications for an antideuteron-based dark matter search. *Astroparticle Physics* 49 (2013), 52–62.
- [3] BRÄUNINGER, C. B., AND CIRELLI, M. Anti-deuterons from heavy dark matter. *Physics Letters B* 678, 1 (2009), 20–31.
- [4] CHOLLET, F., ET AL. Keras documentation. *keras. io* 33 (2015).
- [5] FUKE, H., KOGLIN, J., YOSHIDA, T., ARAMAKI, T., CRAIG, W., FABRIS, L., GAHBAUER, F., HAILEY, C., JOU, F., MADDEN, N., MORI, K., YU, H., AND ZIOCK, K. Current status and future plans for the general antiparticle spectrometer (GAPS). *Advances in Space Research* 41, 12 (2008), 2056–2060.
- [6] GUIDOTTI, R., MONREALE, A., RUGGIERI, S., TURINI, F., PEDRESCHI, D., AND GIANNOTTI, F. A survey of methods for explaining black box models, 2018.
- [7] SELVARAJU, R. R., COGSWELL, M., DAS, A., VEDANTAM, R., PARIKH, D., AND BATRA, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. pp. 618–626.
- [8] 橋本岳, 大塚壮平, 小財正義, 清水雄輝, 竹内崇人, 福家英之, 蓑島温志, 吉田篤正, 吉田哲也, 渡邊翼, AND 和田拓也. 宇宙線反粒子探索 GAPS 実験用 TOF シンチレーションカウンタの開発. In 大気球シンポジウム: 平成 29 年度 (nov 2017), 宇宙航空研究開発機構宇宙科学研究所 (JAXA)(ISAS).
- [9] 小財正義, 浅尾義士, 井上剛良, 井上拓哉, 大塚壮平, 岡崎峻, 小川博之, 加藤千尋, 河内明子, 小池貴久, 崎本一博, 清水雄輝, 高橋克征, 高橋俊, 大丸拓郎, 永井大樹, 橋本岳, 福家英之, 蓑島温志, 宗像一起, 山田昇, 吉住雄大, 吉田篤正, 吉田哲也, 渡邊翼, 和田拓也, ARAMAKI, T., BOGGS, S., CRAIG, W. W., DOETINCH, P., FABRIS, R., GAHBAUER, F., HAILEY, C., MADDEN, N., MOGNET, S., MORI, K., ONG, R., PEREZ, K., ZIOCK, K., AND ZWEERINK, J. 南極周回気球による宇宙線反粒子探索実験 GAPS の現状報告. In 大気球シンポジウム: 平成 28 年度 (nov 2016), 宇宙航空研究開発機構宇宙科学研究所 (JAXA)(ISAS).
- [10] 福家英之, 坂東信尚, 小川博之, 岡崎峻, 莊司泰弘, 吉田哲也, 荒牧嗣夫, 森嘉野, 河合宏紀, AND 宮崎芳郎. GAPS プロトタイプ気球実験計画の全体報告. 大気球シンポジウム: 平成 23 年度 (2011).
- [11] 福家英之, 坂東信尚, 小川博之, 岡崎峻, 莊司泰弘, 高田淳史, 吉田哲也, AND 荒牧嗣夫. 宇宙線反粒子探索計画 GAPS のプロトタイプ気球実験計画. 大気球シンポジウム: 平成 22 年度 (2010), 132–135.
- [12] 福家英之, 小財正義, 小川博之, 岡崎峻, 西城大, 徳永翔, 吉田哲也, 中上裕輔, 竹内崇人, AND 和田拓也. 宇宙線反粒子探索 GAPS 実験計画の (特に日本チームの) 現状報告. 大気球シンポジウム: 2019 年度 (2019).
- [13] 和田拓也, 小財正義, 清水雄輝, 中上裕輔, 福家英之, AND 吉田哲也. ニューラルネットワークを用いた GAPS 用粒子識別法の開発. 第 20 回宇宙科学シンポジウム講演集.
- [14] 和田拓也, 大塚壮平, 小財正義, 清水雄輝, 橋本岳, 福家英之, 蓑島温志, 吉田篤正, 吉田哲也, AND 渡邊翼. GAPS 用 TOF シンチレーションカウンタの開発, jan 2017.