

Application of Deep Learning Methods Combined with Physical Background in Wide-Field-of-View Imaging Atmospheric Cherenkov Telescopes

Authors: Ao-Yan Cheng, Hao Cai, Tian-Lu Chen, Yi-Qing Guo, Hao Cai, Tian-Lu Chen, Yi-Qing Guo

Date: 2024-03-10T00:00:00+00:00

Abstract

The HADAR experiment, which will be constructed in Tibet, China, combines the wide-angle advantages of traditional EAS array detectors with the high sensitivity advantages of focused Cherenkov detectors. Its physics objective is to observe transient sources such as gamma-ray bursts and counterparts of gravitational waves. The aim of this study is to utilize the latest AI technology to enhance the sensitivity of the HADAR experiment. We have built training datasets and models with distinctive creativity by incorporating relevant physical theories for various applications. They are able to determine the kind, energy, and direction of incident particles after careful design. We have obtained a background identification accuracy of 98.6 %, a relative energy reconstruction error of 10.0 %, and an angular resolution of 0.22-degrees in a test dataset at 10 TeV. These findings demonstrate the enormous potential for enhancing the precision and dependability of detector data analysis in astrophysical research. Thanks to deep learning techniques, the HADAR experiment's observational sensitivity to the Crab Nebula has surpassed that of MAGIC and H.E.S.S. at energies below 0.5 TeV and remains competitive with conventional narrow-field Cherenkov telescopes at higher energies. Additionally, our experiment offers a fresh approach to dealing with strongly connected scattered data.

Full Text

Preamble

Application of Deep Learning Methods Combined with Physical Background in Wide Field of View Imaging Atmospheric Cherenkov Telescopes

Ao-Yan Cheng,¹ Hao Cai,^{1,*} Shi Chen,² Tian-Lu Chen,^{3,†} Xiang Dong,¹ You-Liang Feng,³ Qi Gao,³ Quan-Bu Gou,⁴ Yi-Qing Guo,^{4,5,‡} Hong-Bo Hu,^{4,5} Ming-Ming Kang,⁶ Hai-Jin Li,³ Chen Liu,⁴ Mao-Yuan Liu,³ Wei Liu,^{4,5} Fang-Sheng Min,⁴ Chu-Cheng Pan,¹ Bing-Qiang Qiao,⁴ Xiang-Li Qian,⁷ Hui-Ying Sun,⁷ Yu-Chang Sun,¹ Ao-Bo Wang,¹ Xu Wang,⁷ Zhen Wang,⁸ Guang-Guang Xin,⁹ Yu-Hua Yao,^{4,10} Qiang Yuan,¹¹ and Yi Zhang¹¹

¹School of Physics and Technology, Wuhan University, Wuhan 430072, People's Republic of China

²School of Physics and Astronomy, Yunnan University, Yunnan 650091, People's Republic of China

³The Key Laboratory of Cosmic Rays (Tibet University), Ministry of Education, Lhasa 850000, Tibet, People's Republic of China

⁴Key Laboratory of Particle Astrophysics, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, People's Republic of China

⁵University of Chinese Academy of Sciences, 19 A Yuquan Road, Shijingshan District, Beijing, 100049, People's Republic of China

⁶College of Physics, Sichuan University, Chengdu, 610064, People's Republic of China

⁷School of Intelligent Engineering, Shandong Management University, Jinan, 250357, People's Republic of China

⁸Tsung-Dao Lee Institute, Shanghai Jiao Tong University, 200240 Shanghai, People's Republic of China

⁹Suzhou Aerospace Information Research Institute, Suzhou 215123, People's Republic of China

¹⁰College of Physics, Chongqing University, No.55 Daxuecheng South Road, High-tech District, Chongqing, 401331, People's Republic of China

¹¹Key Laboratory of Dark Matter and Space Astronomy, Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, People's Republic of China

The High Altitude Detection of Astronomical Radiation (HADAR) experiment, constructed in Tibet, China, combines the wide-field advantages of traditional extensive air shower array detectors with the high-sensitivity capabilities of focused Cherenkov detectors. Its objective is to observe transient sources such as gamma-ray bursts and gravitational wave counterparts. This study aims to leverage the latest AI technology to enhance the sensitivity of HADAR experiments. By incorporating relevant physical theories for various applications, we constructed training datasets and models with distinctive creativity. These models can determine the type, energy, and direction of incident particles through careful design. We achieved a background identification accuracy of 98.6%, a relative energy reconstruction error of 10.0%, and an angular resolution of 0.22 degrees in a test dataset at 10 TeV. These findings demonstrate the significant potential for enhancing the precision and reliability of detector data analysis in astrophysical research. Using deep learning techniques, the HADAR experiment's observational sensitivity to the Crab Nebula has surpassed that of MAGIC and H.E.S.S. at energies below 0.5 TeV and remains competitive with conventional

narrow-field Cherenkov telescopes at higher energies. In addition, our experiment offers a new approach for dealing with strongly connected, scattered data.

Keywords: VHE Gamma-Ray Astronomy; HADAR; Deep Learning; Convolutional Neural Networks

Introduction

The investigation of ultrahigh-energy gamma-ray astronomy [1] is pivotal for understanding a range of extreme astrophysical phenomena. These ultrahigh-energy gamma rays predominantly originate from highly active galactic nuclei, cataclysmic supernova explosions, neutron stars, and black holes. Importantly, these observations form an empirical cornerstone for addressing some of the most compelling scientific questions, including the detection of dark matter and the identification of cosmic-ray sources.

Space observatories such as the Fermi Gamma-ray Space Telescope [2][3] require long observation times to gather statistically relevant data because high-energy gamma-ray flux decays quickly and follows a power-law distribution [4]. Consequently, large-area ground detectors are required for detecting high-energy gamma rays.

The Earth's atmosphere has become a substantial interaction medium for high-energy photons in energy ranges exceeding gigaelectronvolts [5], effectively preventing their direct detection. In this regime, high-energy γ -rays interact with the Earth's upper atmospheric layers [6], leading to the generation of electron-positron e^-e^+ pairs, which in turn initiate electromagnetic cascades. Within these cascades, relativistic electrons and positrons produce concentrated beams of Cherenkov radiation. These beams serve as the principal observational targets for specialized ground-based detection systems, enabling nuanced studies of the high-energy universe.

Over the last few decades, gamma-ray astronomy has become a frontier discipline for studying the universe's highest-energy astrophysical phenomena [7][8]. The Cherenkov Telescope Array (CTA) is an international initiative that seeks to better understand high-energy gamma rays in the universe by detecting atmospheric Cherenkov radiation [9]. The 20–300 TeV energy range covered by CTA aims to close the observational gap currently present in this energy range. A group of five Cherenkov telescopes, called the High Energy Stereoscopic System (H.E.S.S.), is now operational in Namibia and used to observe cosmic TeV energy rays. It offers crucial information for comprehending gamma-ray sources [10] such as pulsars and supernova remnants [11]. The Major Atmospheric Gamma Imaging Cherenkov (MAGIC) telescopes, two telescopes situated at La Palma in the Canary Islands, are another notable project. By detecting Cherenkov radiation from cosmic gamma-ray sources [12], the physical characteristics and origins of cosmic rays can be investigated [13]. These telescopes, which have a small field of view, are primarily employed to make accurate observations of known high-energy astrophysical objects, providing important knowledge and

observational data for a comprehensive understanding of these extreme cosmological occurrences.

However, measurable signals from many extreme celestial occurrences only last for brief periods because of their intrinsically transient characteristics. Thus, there is an urgent demand for wide-field detectors that can instantly capture high-energy particle signals across a broad spatial range because large detector arrays require time to recalibrate their directed reception. Wide-field Cherenkov imaging telescopes are particularly useful in such situations.

The High Altitude Detection of Astronomical Radiation (HADAR) experiment is one such technology [14][15], comprising an array of imaging atmospheric Cherenkov refractive telescopes based on atmospheric Cherenkov principles. Positioned at an impressive altitude of 4300 m in Yangbajing, Tibet (30.0848°N, 90.5522°E), the experiment incorporates four water lenses, each spanning a diameter of 5 m, strategically placed at the vertices of a square with a side length of 100 m [16]. These lenses, ensconced within a hemispherical glass shell 5 m in diameter, are housed in a steel-structured tank with a diameter of 8 m. This tank, rising to a height of 7 m, is filled with high-purity water, and cameras outfitted with photomultiplier tubes are attached to its base. Upon the interception of Cherenkov radiation, a quartet of detectors meticulously logs the charge deposition on the photomultiplier tubes within a predefined temporal frame. All the collected data are subsequently archived in an external storage apparatus.

To effectively detect very high-energy (VHE) γ -ray sources, analytical methodologies must execute several critical tasks:

Background Suppression: The identification of distinct shape features within detector images is pivotal. This allows the isolation of target γ -rays from the overwhelmingly abundant background of cosmic rays, which are predominantly composed of protons.

Energy Reconstruction: The accurate estimation of the original energy of incident particles is achieved by correlating variables such as deposited charge and the relative spatial coordinates within the detectors.

Direction Reconstruction: Leveraging the stereoscopic images acquired from the detectors, the axis direction of the resultant particle shower must be reconstructed. This facilitates the estimation of the original direction from which the incident particles originated.

Deep learning techniques are becoming increasingly integral to the data analysis frameworks employed in Cherenkov telescope experiments. A variety of computational algorithms, particularly those based on convolutional neural networks (CNNs), have shown unparalleled success in addressing the multifaceted analytical challenges endemic to this scientific domain. Prominent initiatives, such as H.E.S.S. [17] and CTA [18], have harnessed the robust capabilities of CNNs to analyze their observational datasets, yielding significant advancements. These

technological strides corroborate the findings of the present study, reinforcing the compelling case for the widespread utility and robustness of deep learning methodologies within this specialized field of research.

This study is structured as follows: The use of deep learning is briefly introduced in Section 2, along with our data production process. A more thorough overview is provided in Sections 3-5, and the pertinent data are used as the basis for training the background suppression, energy reconstruction, and incident direction models. The modeling calculations of the flux and observational sensitivity to the Crab Nebula at various energies are presented in Section 6. Finally, Section 7 presents a thorough summary of the materials used.

II. Deep Learning Approaches for HADAR Data

Deep learning is an important branch of artificial intelligence (AI) that has emerged as a growing field in recent years. With the significant increase in computational power, especially with GPU chips, the real-time computation of large parallel data (e.g., high-dimensional matrices) has become possible. This has led to the convergence of deep learning in various data analysis industries, simplifying tasks that were previously difficult for humans to perform. As a type of AI, deep learning methods simulate the multi-layer neural network structure of the human brain to solve problems. Similar to human learning, deep learning methods continuously update network parameters through a step-by-step understanding of data. This process ultimately allows the network to extract as many useful features as possible from the inputs and obtain model predictions through neural network computations.

In this study, we adopted CNNs. As excellent models for image recognition, CNNs can spontaneously extract local relationships at different positions within the same dimension and across different dimensions based on the input signals, ultimately yielding optimal results. Fundamentally, a CNN is composed of multiple layers consisting of a series of convolutional pooling layers connected to fully connected layers. Owing to this layered structure, errors from one layer propagate to the next, potentially leading to an exponential growth in errors. To mitigate this problem, we used a specialized CNN known as a residual CNN. The advantage of this network is that each layer receives not only processed signals from the previous layer but also original signals from the previous layer. This significantly enhances the model's ability to fit the data.

By leveraging the power of CNNs, specifically residual CNNs, we aim to address the three main challenges in detecting VHE gamma-ray sources: background suppression, energy reconstruction, and direction reconstruction. Our experimental results, which are detailed in the subsequent sections, demonstrate the effectiveness of deep learning methods in these high-stakes, complex analytical tasks.

Throughout the research process, all the datasets used for training the neural networks were generated using the widely adopted CORSIKA Monte Carlo pro-

gram [19]. In the simulations, the set altitude was 4300 m, which corresponded to an atmospheric depth of 606 g/cm². Geomagnetic coordinates were set in Yangbajing, Tibet. The simulated primary cosmic-ray particles included both gamma rays (serving as a signal) and protons (serving as a background), with energies ranging from 20 to 10 TeV. The incident zenith angle was set to 20° and the azimuthal angle ranged from 0 to 360°. All events were uniformly scattered within a circle centered on the HADAR array with a radius of 400 m. Following shower simulation, appropriate software packages were used to simulate the response of each HADAR detector, which was placed at each of the four vertices of a square with a side length of 100 m. The size of the field of view determines the detector camera size. When storing a cluster shot example, we chose a 400 × 400 storage range to maintain integrity. Each photomultiplier tube had a diameter of 5.1 cm [20]. After storing the example, its use is analyzed to reconstruct the scene.

To implement the model incorporating the associated algorithms, we utilized a deep-learning framework based on PyTorch. Training and testing were performed on a machine equipped with two NVIDIA GeForce GTX 3090 GPUs and a high-performance computing cluster furnished with four NVIDIA V100 tensor-core GPUs, achieving nearly identical results on both platforms.

For diverse research tasks, specialized datasets were curated, a range of neural network architectures were adopted, and various loss functions were employed. The images from the HADAR detector can be used in these various combinations as input (the model accepts an image matrix as input, which varies slightly under different preprocessing efforts), and the neural network computations produce the relevant parameters of interest. Detailed descriptions of these components are thoroughly discussed in the following three sections, where we articulate the underlying rationale for these choices. Our codebase, called GPLearn, is publicly available on GitHub. This package features modular functionalities, enabling users to tailor the module parameters to specific needs. With GPLearn, readers have the capability not only to reproduce the results presented in this paper using our supplied datasets but also to adapt the code for actual detector data analysis. We anticipate that these contributions will be broadly applicable to the field of astrophysics.

III. Background Suppression

The flux of cosmic rays typically exceeds that of high-energy gamma rays by several orders of magnitude, thereby posing a significant challenge for detection. Consequently, the Cherenkov radiation generated by these high-energy cosmic rays acts as the predominant source of background noise. Given that approximately 90% of cosmic rays are protons, this study primarily focuses on evaluating the influence of protons on photon detection. Proton signals with energies three times greater than those of photon signals are particularly problematic, as they are especially difficult to distinguish. To address this challenge, we conducted simulations involving particles with various energies to effectively

segregate the target photon signals from the interfering proton background.

The prevailing methodology for discriminating between high-energy gamma-rays and background cosmic rays leverages the intrinsic differences between electromagnetic and hadronic showers. When photons impinge upon the Earth's atmosphere, they predominantly induce electromagnetic showers, which, in contrast to the hadronic showers triggered by protons, yield more uniform detector images. Specifically, in the data recorded by our detectors, the images were represented as pixel arrays (Fig. 1 [Figure 1: see original paper]). For a fitted ellipse, conventional methods compute the image moments of the pixels. If a shower is composed of photons or protons, it can be determined by comparing the major and minor axes of the fitted ellipses. Because photon-generated electromagnetic showers have a more concentrated energy distribution, their pixel layouts follow a linear trajectory. In contrast, hadronic showers caused by protons exhibit more core hits, resulting in significantly spread and uneven pixel patterns. The main goal of existing background suppression work is to accurately distinguish between the two categories of image structures. In particular, CNNs have produced exceptional results in image identification using deep learning approaches. Consequently, compared to conventional curve-fitting techniques, they offer more accurate discrimination capabilities.

For the background suppression task, we adopted the ResNet-18 residual CNN model, as depicted in Fig. 2 [Figure 2: see original paper], to meet our specific objectives. Unlike the canonical ResNet-18 model, our customized version employs smaller convolutional kernels with designated strides for initial data convolution. This choice was motivated by the fact that the region of interest in our detector images was generally confined to a narrow central zone. The use of smaller kernels was advantageous for capturing the critical shape and edge features within this focal region, as substantiated by the improved performance in subsequent evaluations.

The output of our model is formulated as a two-dimensional vector that quantifies the probability of a particle being categorized as either a signal (a photon) or background noise (a proton). This vector undergoes automatic normalization to ensure that its components sum to one, thereby enhancing interpretability. The normalized probabilities s_k are computed using the following equation:

$$s_k = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}}$$

Here, z_k denotes the raw output value corresponding to the k th class, and K denotes the total number of classes. In our application, $K = 2$ represents the signal and noise categories. This normalization ensures that the output probabilities are mutually exclusive and exhaustive, summing up to one, which aids in yielding more interpretable outcomes.

To evaluate the divergence between the model's predicted and actual probability

distributions, we employed cross-entropy as our loss function, which is expressed as follows:

$$CE(y, p) = - \sum [y_g \ln(p_{gi}) + y_p \ln(p_{pi})]$$

where $y_g = 0$ and $y_p = 1$ indicate the labeling of protons and photons, respectively, and p_{gi} and p_{pi} reflect the probabilities of being predicted as a photon and proton, respectively, for the i th instance model output. In deep learning, the cross-entropy loss function is the most widely used function for binary classification tasks. This helps the model train more quickly and is highly sensitive to changes in its predicted probabilities.

The physical context was considered during the data preparation. Considering that electromagnetic and hadron clusters primarily create differences in the degree of image standardization, we combined the actual background of the current work. Consequently, the absolute positions between different detectors become less important. Therefore, we retained only a small portion of the images (Fig. 3(a) [Figure 3: see original paper]) centered on the actual signal area (Fig. 3(b) [Figure 3: see original paper]) and stitched the remaining image together (Fig. 3(c) [Figure 3: see original paper]), so that the resulting 128×128 matrix was taken as an input to the model. Finally, the model processed the data accurately and precisely, providing a satisfactory answer to this physical problem based only on the intensity and relative positions.

To demonstrate the effectiveness and performance of the model, we introduced a threshold factor, denoted by ζ_0 , and set its value to 0.5 as a criterion for interpreting the classification of incoming particles. In this context, ζ refers to the second component of the two-dimensional output vector of the model. In an ideal scenario, the photon signal output should be aligned at $\zeta_\gamma = 0$, whereas the proton signal output should converge at $\zeta_{\text{proton}} = 1$. However, in practice, any intersection between the distribution curves of ζ_γ and ζ_{proton} signifies the error rate of the model in the classification, as illustrated in Fig. 4 [Figure 4: see original paper].

We retained the events identified as photons and excluded those identified as protons. Simultaneously, we recorded the correct identification rates for both photon and proton events. The quality factor Q is defined as follows:

$$Q = \frac{\varepsilon_{\gamma s}}{\sqrt{\varepsilon_{ps}}}$$

Here, $\varepsilon_{\gamma s}$ represents the survival rate of photons after discrimination, and ε_{ps} represents the survival rate of protons after discrimination. The higher the correct identification rate of the model, the higher the quality factor, indicating better background suppression.

Figure 5 [Figure 5: see original paper] presents the accuracy of the model for particle identification across various energy thresholds. The yellow curve, which corresponds to protons, consistently demonstrates high identification rates across the entire energy spectrum. Conversely, the blue curve, representing photons, exhibits a more variable identification rate with a marked improvement at elevated energy levels. Previous research indicated that the threshold energy for the production of Cherenkov radiation is approximately 50 GeV. Near this energy threshold, conventional electromagnetic showers are prone to generating anomalous results. As the energy level increased, these showers exhibited increasing regularity, leading to a consequent enhancement in the model's identification accuracy.

The results from the test set indicated that the identification rates for both photons and protons exhibited a marked increase as the particle energy increased. At lower energy levels, particles generate smaller atmospheric showers, resulting in a limited number of pixels being captured by ground-based detectors. This scarcity of data hampers the ability of the model to render accurate identification. Conversely, as the particle energy increases, the quality of the images captured by the detectors improves significantly. This augmentation in data quality, coupled with an adequate number of parameters, led to a significant increase in the identification accuracy of the model.

The yellow curve in Figure 6 [Figure 6: see original paper] illustrates the performance gains achieved through the utilization of deep learning techniques, in stark contrast to the blue curve, which represents the efficacy of traditional methods. In the domain of low-energy particles, the quality of identification markedly increased with increasing energy levels, leading to a swift incline in the quality factor curve. Conversely, at higher energy thresholds, the model has ample raw parameters available for making accurate judgments, resulting in a plateauing of the result curve. Notably, the model's identification rate has already approached a robust 98% at this stage.

IV. Energy Reconstruction

The outcome may not be satisfactory if the model's output is solely labeled as particle energy without any changes to the model's inputs or structure. The reconstructed particle energy is related to the incidence zenith angle, distance from the cluster core, and quantity of charge deposited in the detector after accounting for the pertinent physical background. Traditional approaches employ simulated data to determine the energy distribution function at each location and actual observational data to fit the following function [21]:

$$E_{\text{erc}} = f(q, R, \theta)$$

Because the absolute positional information of the relative detector array is required to calculate the distance and direction information of the signal to the

actual incident center, and because this information was removed during the dataset construction process for background suppression, data retaining only the relative positional information is insufficient for reconstructing the final energy.

As a result, we added the relative position of the pixel centroid and the absolute position of the detector array as new inputs to the model based on the data preprocessing employed in the background suppression. That is, we added a 4×4 matrix to the 128×128 matrix used for background suppression to represent the absolute and relative coordinates of the image centroid for each detector. The model was created as depicted in Figure 7 [Figure 7: see original paper], considering that the energy resolution does not require particle classification work. Four detector photos were processed independently and combined after a series of fully connected layers, and the final result was represented by a one-dimensional vector. This model, which does not include convolutional layers, not only retains a high level of reconstruction accuracy but also significantly accelerates the calculations.

Expanding upon the data preprocessing techniques deployed in our background suppression task, we incorporated additional variables: the relative positions of the pixel centroids and the absolute coordinates of the detector array. This is in recognition of the fact that energy is fundamentally correlated with the amount of charge deposited. Accordingly, our model architecture, illustrated in Figure 7 [Figure 7: see original paper], processes images from four individual detectors in parallel. These processed outputs are subsequently fused through a sequence of fully connected layers, culminating in a one-dimensional output vector.

In the test dataset, we employed both the model-predicted energy and the actual particle energy to compute the associated energy resolution, denoted as $\Delta E/E$. As illustrated in Figure 8 [Figure 8: see original paper], our findings reveal a trend of improving energy resolution with increasing particle energy.

To provide an intuitive assessment of the efficacy of our model, we analyzed 2,500 samples from the test dataset, each representing a different energy level. We generated a distribution curve to illustrate the energy prediction performance of the model. The predicted energy values are plotted on the x-axis, and the y-axis represents the probability associated with each specific energy prediction. To facilitate easier interpretation, we normalized the data for each discrete energy level, ensuring that the integral of each curve across all energy points was equal to unity.

Figure 9 [Figure 9: see original paper] illustrates the distribution curves for the energy prediction at 500 GeV intervals. The areas of overlap between these curves signify regions where the model's estimations may deviate from the actual conditions, constituting a principal constraint on the accuracy of the energy resolution. More importantly, our model achieved a relative error of only 13.0% at an energy level of 1 TeV. Remarkably, at higher energy levels, such as 10 TeV, the relative error narrowed further to 10.0%. This value approaches the

theoretical limits of the energy resolution achievable with Cherenkov telescopes, underscoring the significance of our research.

V. Direction Reconstruction

After the atmospheric Cherenkov radiation reaches the ground-level detectors, it is refracted through specialized water lenses before being captured on the image plane. Calculating the path of refracted light is complex. Our initial efforts were focused on estimating the projected position of the shower core on this image plane. By calculating the fitting ellipse of the four detector images and identifying the location where clustering occurs based on the weighted intersection of the long axes of each ellipse, conventional methods (such as the Hillas reconstruction [22]) can determine where clustering occurs and, ultimately, the direction of the incidence of γ -rays [23][24].

Unlike the energy reconstruction task, both the intensity and absolute position of each pixel are intricately linked, serving as indirect indicators of the incident direction of the gamma rays. However, generating two-dimensional matrices with large input sizes is improbable, and it is implausible to create pixel matrices based solely on actual positions, according to the description of the detector array in Section 2. We continued to use the cropping method in response to this.

Unlike background suppression, each cropped pixel point is considered a three-dimensional dataset, which is actually a $4 \times 64 \times 64 \times 3$ matrix. The first dimension of this matrix indicates the detector's serial number; the second and third dimensions indicate the relative position of the corresponding signal on the detector (with respect to a single detector); and the fourth dimension indicates the signal intensity detected by the photomultiplier tube at this position, along with the point's absolute position (with respect to the ground). Our tests demonstrate that the performance of the model is significantly enhanced when the absolute positioning information of the pixels is retained.

To accommodate the specialized input dataset, we employed a CNN model featuring three-dimensional convolutional layers, as shown in Figure 10 [Figure 10: see original paper]. The raw data underwent a sequence of six upsampling convolutional layers, each followed by a pooling layer. The processed data were then fed into a linear regression layer, culminating in a two-dimensional vector output. After normalization, this vector served as an azimuthal direction vector in Earth's plane. Subsequently, the model computed the corresponding incident angle based on this vector.

The error angle distance Ω was calculated from the azimuthal angle using a spherical coordinate transformation:

$$\cos \Omega = \sin \theta \sin \phi \cdot \sin \theta \sin \phi' + \sin \theta \cos \phi \cdot \sin \theta \cos \phi' + \cos^2 \theta$$

Here, ϕ represents the true incident azimuthal angle, and ϕ' represents the azimuthal angle reconstructed by the model. All our simulations were based on an incident zenith angle of $\theta = 20^\circ$. The results plotted in Figure 11 [Figure 11: see original paper] reveal that the deep learning method we employed yields an approximately 60% improvement in accuracy over traditional methods, as cited in [25]. Notably, the reconstruction error decreases substantially with increasing energy levels. This enhanced accuracy is attributable to the higher number of photons collected by the telescope at higher energy levels, which allows for a more precise reconstruction of the event direction. In the energy range exceeding 1 TeV, our model reduced the angular error to less than 0.2° , providing a robust foundation for subsequent research.

VI. Simulated Measurements of the Crab Nebula

A. Sensitivity measurements

The Crab Nebula serves as a prototypical supernova remnant, emitting across a wide range of wavelengths, including radio, X-ray, and gamma-ray bands. The study of this celestial object is pivotal for gaining insight into the physical mechanisms underlying supernova explosions, the generation and acceleration of cosmic rays, and for addressing important scientific queries, such as the nature of dark matter. Among the numerous physical attributes associated with the Crab Nebula that require precise measurement, sensitivity is a crucial parameter. The level of sensitivity directly impacts the quality—both in terms of accuracy and precision—of our estimates concerning various physical phenomena related to the Crab Nebula.

In the preceding sections, we delved into the deployment of deep learning methodologies for sophisticated analysis of detector data, achieving a marked improvement in accuracy. In this section, our focus shifts to the specific application of deep learning to enhance the sensitivity measurements related to the Crab Nebula. We illustrate how these advanced computational techniques outperform traditional methods and offer a more precise and reliable assessment of this crucial parameter. Additionally, we juxtapose our findings with those obtained from other Cherenkov imaging telescope experiments to contextualize the efficacy and innovation of our approach.

To calculate the sensitivity, we conducted a simulation that accounted for both gamma rays and scattered protons across an energy spectrum ranging from 100 GeV to 10 TeV. The simulated sampling area was set to $S_{\text{sample}} = 800 \times 800 \text{ m}^2$, and a fixed zenith angle of 20° was used. The azimuthal angle in the simulation was varied between 0° and 180° . Additionally, we integrated the geographical factors specific to the Yangbajing site and the celestial trajectory of the Crab Nebula. Based on these considerations, we estimated the annual observation time dedicated to the Crab Nebula to be approximately 320 h, which translates to $T_{\text{crab,obs}} = 1.152 \times 10^6 \text{ s}$.

For gamma rays, we can calculate the effective number of events coming from

the Crab Nebula received by the HADAR experiment in one year at energy i by using the following equation:

$$N_{1\text{yr}}[i] = T_{\text{crab,obs}} \cdot S_{\text{sample}} \cdot F_{\text{crab}} \cdot \frac{N_{\gamma,\text{sim}}[i]}{N_{\gamma,\text{sim,all}}[i]} \cdot \varepsilon_{\gamma}$$

Here, $F_{\text{crab}} = \int_i^{\infty} 2.83 \times 10^{-11} \text{ ph cm}^{-2} \text{ s}^{-1} \text{ TeV}^{-1} \cdot (E/\text{TeV})^{-2.62} dE$ represents the integrated flux of the Crab Nebula for energies greater than i . $N_{\gamma,\text{sim}}[i]$ is the number of effective simulated photon events in energy range i , $N_{\gamma,\text{sim,all}}[i]$ is the total number of simulated photon events, and ε_{γ} represents the ratio of events located within an angular resolution set at 68%.

For scattered protons, the effective number of observed events for one year can be calculated as:

$$N_{1\text{yr}}^{\text{CR}}[i] = T_{\text{crab,obs}} \cdot S_{\text{sample}} \cdot F_{\text{CR}} \cdot \frac{N_{\text{CR,sim}}[i]}{N_{\text{CR,sim,all}}[i]} \cdot \Omega_i$$

Here, F_{CR} represents the integrated flux of protons in the cosmic rays. $N_{\text{CR,sim}}[i]$ is the number of effective simulated proton events in the energy range i , $N_{\text{CR,sim,all}}[i]$ is the total number of simulated proton events, and Ω_i represents the solid angle within the photon angular resolution range.

After the implementation of background suppression techniques and the exclusion of proton-like events, we can derive the integrated significance across the designated energy spectrum using Equation [26]:

$$S[i] = \frac{N_{1\text{yr}}[i]}{\sqrt{N_{1\text{yr}}^{\text{CR}}[i]}}$$

Utilizing the framework of Gaussian statistics, sensitivity is defined as the minimal flux from the Crab Nebula required for the detector to register a signal at a 5-sigma significance level. Consequently, the specific sensitivity of the Crab Nebula for the HADAR experiment was calculated using the following formula:

$$F_{\text{sensitivity}}[i] = F_{\text{crab}} \cdot \frac{5}{S[i]}$$

To offer a comprehensive perspective, we graphically delineated the calculated sensitivity metrics both prior to and following the incorporation of the deep learning algorithms. These were juxtaposed with the corresponding data from other pertinent experiments [27][28][29], as depicted in Figure 12 [Figure 12: see original paper].

Following the integration of advanced deep learning algorithms, the HADAR experiment exhibited a substantial enhancement in sensitivity within the low-energy domain. Although it has yet to attain the level of sensitivity exhibited by CTA, the revamped HADAR experiment now stands as a formidable contender to established initiatives such as MAGIC and H.E.S.S. Importantly, HADAR boasts a unique attribute: an expansive field of view not typically afforded by conventional reflective Cherenkov telescopes. This distinct advantage facilitates real-time capture of transient celestial sources within its observational purview, rendering the outcomes particularly compelling.

B. Observation of the Pulsar in the Crab Nebula

The primary objective of this section is to perform a rigorous analysis of the observational data pertaining to the Crab Nebula pulsar in the VHE regime, as acquired through the MAGIC telescope [30]. The significance of the observations from the HADAR experiment was calculated to lay the groundwork for future inquiries into pulsar emissions at high-energy wavelengths.

We determined the spectral properties of the primary pulsar (P1) and interpulsar (P2) of the Crab Nebula using data obtained from the MAGIC telescope [31], as shown graphically in Figure 13 [Figure 13: see original paper]. The formula for the integral flux of each energy band is as follows:

$$F_{P1} = 1.1 \times 10^{-11} \cdot \left(\frac{E}{\text{TeV}} \right)^{-3.2} \text{ cm}^{-2}\text{s}^{-1}$$

$$F_{P2} = 2 \times 10^{-11} \cdot \left(\frac{E}{\text{TeV}} \right)^{-2.9} \text{ cm}^{-2}\text{s}^{-1}$$

Following the implementation of the background filtering procedures, only events characterized by gamma-like profiles were retained for subsequent analyses of the pulse signals of the primary pulsar of the Crab Nebula (P1) and interpulsar (P2). To quantify the observational significance S of the HADAR experiment, we employed the following equation:

$$S = \frac{N_{\text{on}} - \alpha N_{\text{off}}}{\sqrt{\alpha(N_{\text{on}} + N_{\text{off}})}}$$

Here, N_{on} signifies the photon count emanating from the pulsar, and N_{off} represents the background photon count. Remarkably, at the 100 GeV energy threshold, the dominant background source in the observations of the Crab Nebula pulsar transitions from protons to gamma rays emanating from the Crab Nebula itself. Consequently, N_{off} encapsulates both proton and non-pulsar gamma-ray backgrounds. The observation time ratio α is defined as $\alpha = t_{\text{on}}/t_{\text{off}}$, where t_{on} and t_{off} denote the durations of the HADAR detector observations for the source and background, respectively.

Tables 1 and 2 list the observational data sets acquired from the MAGIC and HADAR experiments, respectively. These tables encapsulate the derived significances corresponding to distinct energy bands for both the main pulsar (P1) and interpulsar (P2) of the Crab Nebula, as corroborated by references [32] and [33].

Capitalizing on the robust data processing capabilities afforded by deep learning methodologies, the HADAR experiment demonstrated a notably superior level of significance within the low-energy domain compared with the MAGIC experiment. Nevertheless, as the energy spectrum increases, certain constraints, including equipment specifications and geographical considerations, result in a discernible performance disparity compared with MAGIC. Moving forward, we aim to augment the sensitivity of our apparatus through a multi-faceted approach, encompassing the expansion of the detector's effective observational area as well as the refinement of computational models to enhance the angular resolution.

VII. Outlook

With the emergence of large-scale AI language models, AI has been progressively incorporated into many industrial applications. In physics, data analytics are a burgeoning arena for the expansion of AI technologies. By leveraging its computational advantages, AI has the potential to uncover latent relationships within complex datasets, accelerate data-fitting procedures, and simulate realistic models of experimental measurements.

In high-energy physics, the synergy between AI and big data allows for the precise reconstruction of energy, momentum, and mass metrics for particles, which is a critical step towards elucidating the underlying properties and behaviors of subatomic particles. Additionally, exact spatial reconstructions are indispensable for tracking the decay products of heavy particles and capturing rare events that could signal novel particles or interaction mechanisms. In the field of astronomy, AI technologies facilitate the real-time processing of astronomical signals, enabling more efficient studies of distant cosmic events and contributing to our understanding of the origin and composition of the universe.

In this study, we effectively addressed complicated physical issues by integrating real-world physical contexts using appropriate data preparation and models. Additionally, the excellent results from deep learning also serve to confirm the accuracy of the corresponding physical theories. The training process, which was built in accordance with real-world physical theories, both theoretically supported and verified the deep learning results.

During the actual research process, we demonstrated the following:

1. The relative position and shape information of Cherenkov radiation can be used to discriminate particle types; at the same time, information contained in data that is not sensitive to absolute position will not be lost

after slicing and cropping.

2. The energy information from the Cherenkov radiation reaction is not sensitive to the relative position information of pixels; moreover, energy can be approximately represented by a linear function of charge amount, incidence angle, and relative center distance.
3. The directional information of Cherenkov radiation is sensitive to the absolute position of pixels; therefore, retaining the absolute position information of pixels is crucial for correctly inferring the initial direction of particles when processing image data.

Deep learning techniques have simplified and improved labor-intensive physical reconstruction procedures. However, because physical data features are typically sensitive to a limited scale range, adopting small-scale convolutional kernels aids neural networks in acutely recognizing features. A network can gather global feature information more efficiently by expanding its dimensions. The cropping of data was made possible for geographically distinct trials by adding the necessary positional dimensions, which also suggests a new approach for data analysis in large detector arrays.

VIII. Bibliography

- [1] G. G. Fazio, High-energy gamma-ray astronomy. *Nature* 225, 0049/187/2/460 905-911 (1970). doi: 10.1038/225905a0
- [2] A. A. Abdo, M. Ackermann, M. Ajello, et al., The first fermi large area telescope catalog of gamma-ray pulsars. *Astrophys. J. Suppl. Ser.* 187, 460-494 (2010). doi: 10.1088/0067-
- [3] J. A. R. Cembranos, A. de la Cruz-Dombriz, A. L. Maroto, et al., Detection of branon dark matter with gamma ray telescopes. *Phys. Rev. D* 85, 043505 (2012). doi: 10.1103/Phys-RevD.85.043505
- [4] E. P. Mazets, S. V. Golenetskii, V. N. Il' Inskii, et al., Catalog of cosmic gamma-ray bursts from the KONUS experiment data. *Astrophys. Space Sci.* 80, 3-83 (1981). doi: 10.1007/BF00649140
- [5] Luis Anchordoqui, María Teresa Dova, Analisa Mariazzi, et al., High energy physics in the atmosphere: phenomenology of cosmic ray air showers. *Ann. Phys.* 314, 145-207 (2004). doi: 10.1016/j.aop.2004.07.003
- [6] Bernard Degrange and Gérard Fontaine, Introduction to high-energy gamma-ray astronomy. *C. R. Phys.* 16, 587-599 (2015). doi: 10.1016/j.crhy.2015.07.003
- [7] C. Fichtel, Future prospects for gamma-ray astronomy. *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.* 301, 693-701 (1981). doi: 10.1098/rsta.1981.0153
- [8] FAN Zihao, LIANG Juncheng, LIU Haoran, et al., Influence of wall effect on detection efficiency of beta-emitters in TDCR-Cerenkov method. *NUCLEAR TECHNIQUES* 46, 060504 (2023). doi: 10.11889/j.0253-3219.2023.hjs.46.060504
- [9] M. Actis, G. Agnetta, F. Aharonian, et al., Design concepts for the cherenkov

- telescope array cta: an advanced facility for ground-based high-energy γ -ray astronomy. *Exp. Astron.* 32, 193–316 (2011). doi: 10.1007/s10686-011-9247-0
- [10] F. Aharonian, A. G. Akhperjanian, A. R. Bazer-Bachi, et al., The h.e.s.s. survey of the inner galaxy in very high energy gamma rays. *Astrophys. J.* 636, 777 (2006). doi: 10.1086/498013
- [11] F. Aharonian, A. G. Akhperjanian, A. R. Bazer-Bachi, et al., Observations of the crab nebula with hess. *Astron. Astrophys.* 457, 899–915 (2006). doi: 10.1051/0004-6361:20065351
- [12] J. Albert, E. Aliu, H. Anderhub, et al., Vhe gamma-ray observation of the crab nebula and its pulsar with the magic telescope. *Astrophys. J.* 674, 1037–1055 (2008). doi: 10.1086/525270
- [13] J. Albert, E. Aliu, H. Anderhub, et al., Discovery of very high energy gamma radiation from ic 443 with the magic telescope. *Astrophys. J.* 664, L87 (2007). doi: 10.1086/520957
- [14] Guang-Guang Xin, Hao Cai, Yi-Qing Guo, et al., A novel trigger algorithm for wide-field-of-view imaging atmospheric cherenkov technique experiments. *Nucl. Sci. Tech.* 33, 25 (2022). doi: 10.1007/s41365-022-01003-3
- [15] Hui-Ying Sun, Tian-Lu Chen, Danzengluobu, et al., Prospective study on observations of gamma-ray emission from active galactic nuclei using the hadar experiment. *Acta Phys. Sin.* 72, 049501 (2023). doi: 10.7498/aps.72.20221976
- [16] H. Cai, Y. Zhang, C. Liu, et al., Wide field-of-view atmospheric Cherenkov telescope based on refractive lens. *J. Instrum.* 12, P09023 (2017). doi: 10.1088/1748-0221/12/09/P09023
- [17] I. Shilon, M. Kraus, M. Büchele, et al., Application of deep learning methods to analysis of imaging atmospheric cherenkov telescopes data. *Astropart Phys.* 105, 44–53 (2019). doi: 10.1016/j.astropartphys.2018.10.003
- [18] Salvatore Mangano, Carlos Delgado, María Isabel Bernardos, et al., Extracting gamma-ray information from images with convolutional neural network methods on simulated cherenkov telescope array data. *Artificial Neural Networks in Pattern Recognition* 243–254. doi: 10.1007/978-3-319-99978-4_{19}
- [19] Tadeusz Wibig, Testing the superposition model in small corsika shower simulations. *J. Phys. G-Nucl. Part. Phys.* 49, 035201 (2022). doi: 10.1088/1361-6471/ac4da7
- [20] H. Cai, T.-L. Chen, Q. Gao, et al., Performance of a wide field-of-view atmospheric Cherenkov telescope prototype based on a refractive lens. *Nucl. Instrum. Methods Phys. Res. Sect. A-Accel. Spectrom. Dect. Assoc. Equip.* 927, 46–53 (2019). doi: 10.1016/j.nima.2019.02.020
- [21] V. A. Acciari, M. Beilicke, G. Blaylock, et al., Veritas observations of the gamma-ray binary ls i +61 303. *Astrophys. J.* 679, 1427–1432 (2008). doi: 10.1086/587736
- [22] A. M. Hillas, The origin of ultra-high-energy cosmic rays. *Annu. Rev. Astron. Astrophys.* 22, 425–444 (1984). doi: 10.1146/annurev.aa.22.090184.002233
- [23] H. Abdalla, R. Adam, F. Aharonian, et al., Resolving acceleration to very high energies along the jet of centaurus a. *Nature* 582, 356–359 (2020). doi: 10.1038/s41586-020-2354-1

- [24] F. Aharonian, Q. An, Axikegu, et al., Observation of the crab nebula with lhaaso-km2a - a performance study *. Chin. Phys. C 45, 025002 (2021). doi: 10.1088/1674-1137/abd01b
- [25] Guang-Guang Xin, Yu-Hua Yao, Xiang-Li Qian, et al., Prospects for the detection of the prompt very-high-energy emission from gamma-ray bursts with the high altitude detection of astronomical radiation experiment. Astrophys. J. 923, 112 (2021). doi: 10.3847/1538-4357/ac2df7
- [26] Ti-Pei Li and YuQian Ma, Analysis methods for results in gamma-ray astronomy. Astrophys. J. 272, 317-324 (1983). doi: 10.1086/161295
- [27] Collaboration Lhaaso, Zhen Cao, F. Aharonian, et al., Peta-electron volt γ -ray emission from the crab nebula. Science 373, 425-430 (2021). doi: 10.1126/science.abg5137
- [28] Felix Aharonian, Sergey Bogovalov, and D. Khangulyan, Abrupt acceleration of a 'cold' ultrarelativistic wind from the crab pulsar. Nature 482, 507-509 (2012). doi: 10.1038/nature10793
- [29] Alice K. Harding and Constantinos Kalapotharakos, Synchrotron self-compton emission from the crab and other pulsars. Astrophys. J. 811, 63 (2015). doi: 10.1088/0004-637X/811/1/63
- [30] J. Aleksić, E. A. Alvarez, L. A. Antonelli, et al., Observations of the crab pulsar between 25 and 100 gev with the magic i telescope. Astrophys. J. 742, 43 (2011). doi: 10.1088/0004-637X/742/1/43
- [31] Magic Collaboration S. Ansoldi, Lucio Angelo Antonelli, P. Antoranz, et al., Teraelectronvolt pulsed emission from the crab pulsar detected by magic. Astron. Astrophys. 585, 1-6 (2015). doi: 10.1051/0004-6361/201526853
- [32] J. Aleksić, E. A. Alvarez, L. A. Antonelli, et al., Phase-resolved energy spectra of the crab pulsar in the range of 50-400 gev measured with the magic telescopes. Astron. Astrophys. 540, A69 (2012). doi: 10.1051/0004-6361/201118166
- [33] J. Aleksić, E. A. Alvarez, L. A. Antonelli, et al., Performance of the magic stereo system obtained with crab nebula data. Astropart Phys. 35, 435-448 (2012). doi: 10.1016/j.astropartphys.2011.11.007

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.