Boletim da Sociedade Astronômica Brasileira, **36**, no. 1, 58-60

Convolutional neural networks for gamma and hadron event classification

D. J. Silva[1] & A. J. T. S. Mello[1]

[1] Universidade Tecnológica Federal do Paraná e-mail: derleis@alunos.utfpr.edu.br, e-mail: ajmello@utfpr.edu.br

Abstract. Every second, countless particles from space reach our planet, some interacting with the Earth's atmosphere, which acts as a shield. This is the case for cosmic rays and gamma rays, which generate a shower of lighter and less energetic particles as they penetrate the atmosphere. Due to the opacity of the atmosphere to gamma rays, Imaging Atmospheric Cherenkov Telescopes such as HESS, VERITAS, MAGIC, and the upcoming CTA, are essential for observing these events. A particular characteristic of gamma rays is that they travel through the Universe without the deflection suffered by charged particles, thus carrying information about the location where they were generated, allowing for detailed studies of these sources. This work explores the use of deep learning techniques, specifically convolutional neural networks, for analyzing data from these telescopes, separating events initiated by gamma rays from those initiated by hadrons. The training dataset was obtained through computational simulations using the CORSIKA software, generating cosmic showers initiated by protons and high-energy photons. Subsequently, the sim_telarray package was used to simulate the bunch of photons captured by the IACT cameras. These images were used to train a convolutional neural network with the objective of distinguishing showers initiated by gamma rays from hadronic ones. The model achieved an accuracy of over 90% in distinguishing between these two types of events, demonstrating its great potential applicability in the analysis of data from these kind of events.

Resumo. A cada segundo, inúmeras partículas vindas do espaço chegam ao nosso planeta, algumas delas interagem com a atmosfera terrestre, que funciona como um escudo. Esse é o caso dos raios cósmicos e dos raios gama que, ao interagirem e penetrarem na atmosfera, geram uma cascata de partículas mais leves e menos energéticas. Devido à opacidade da atmosfera aos raios gama, telescópios de imageamento Cherenkov atmosférico, como o HESS, VERITAS, MAGIC e o futuro CTAO, são essenciais para a observação desses eventos. Uma característica dos raios gama é que, ao contrário das partículas carregadas, eles viajam pelo Universo sem sofrer deflexão, mantendo informações sobre o local em que foram gerados e permitindo o estudo detalhado dessas regiões. O presente trabalho explora o uso de técnicas de *deep learning* , especificamente redes neurais convolucionais, para analisar os dados desses telescópios, diferenciando os eventos iniciados por raios gama daqueles iniciados por hádrons. O conjunto de dados de treinamento foi obtido através de simulação computacional utilizando o *software* CORSIKA, Subsequentemente, o pacote sim_telarray foi usado para simular o pacote de fótons capturado pelas câmeras dos telescópios. Essas imagens foram usada para treinar uma rede neural artificial com o objetivo de classificá-las entre eventos iniciados por raios gama e eventos iniciados por hádrons. O modelo alcançou uma acurácia de mais de 90% na classificação desses dois eventos, demonstrando potencial de aplicação na análise de dados desses eventos.

Keywords. astroparticle physics – data analysis – Gamma rays

1. Introduction

High-energy particles such as gamma rays and cosmic rays provide crucial insights into some of the most extreme phenomena in the universe, such as pulsars, supernovae, and active galactic nuclei. Gamma rays, which are the highest-energy photons in the electromagnetic spectrum, travel through space without suffering deflections by magnetic fields, carrying valuable information about their sources. In contrast, cosmic rays consist primarily of charged particles, such as protons and heavier nuclei, that are deflected by magnetic fields, which obscure their origins. Despite these differences, both types of particles produce extensive air showers (EAS) upon interacting with Earth's atmosphere (Fig.1), leading to the emission of Cherenkov radiation (Grupen 2020).

The detection of Cherenkov radiation has been widely adopted in observatories like HESS, MAGIC, and VERITAS, with the Cherenkov Telescope Array Observatory (CTAO) set to become the most advanced facility for such studies. The CTAO will feature an array of telescopes with unprecedented sensitivity and resolution (Acharya 2018), enabling detailed investigations into the origins and properties of high-energy particles. However, the volume of data generated by these instruments poses a significant challenge. To address this, artificial intelligence techniques, particularly convolutional neural networks (CNNs), offer a powerful solution for classifying events with speed and precision. By automating the analysis of Cherenkov imaging data, CNNs enhance the scientific output of these ob-

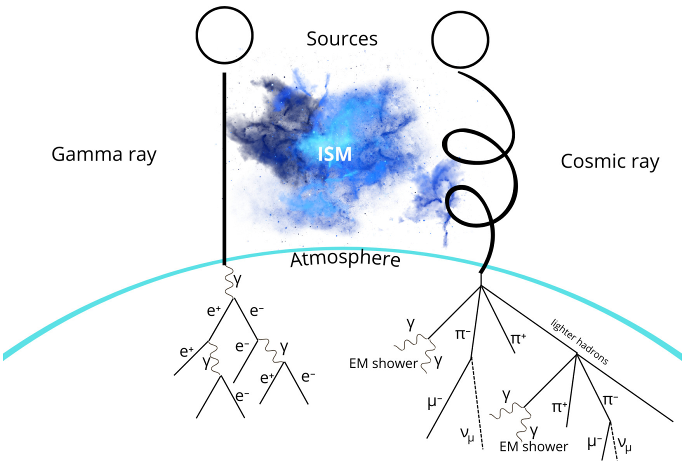

Figure 1. Trajectories of gamma rays and cosmic rays through space and their interactions with Earth's atmosphere. Gamma rays initiate electromagnetic showers (left), while cosmic rays produce more complex extensive air showers (right).

servatories, reducing human effort and improving classification accuracy.

Computational simulations are great allies in the generation of labeled datasets for training CNNs. In this work, the CORSIKA software (Heck 1998) was used to simulate the development of extensive air showers in the Earth atmosphere, modeling interactions between primary particles and atmospheric nuclei. The sim_telarray package (Bernlohr 2008) was employed to simulate the Cherenkov light emitted by these showers and captured by telescopes, considering the optical and electronic characteristics of instruments like those used in the HESS-I array. These simulations provide large labeled datasets that are essential for the effective training and validation of deep learning models.

2. Objectives

The main objective of this work is to train a CNN model capable of classifying gamma-ray and hadronic events. To achieve this, simulated images generated through Monte Carlo techniques are utilized, followed by preprocessing steps to ensure compatibility with CNNs.

3. Methods

This work was divided into three main stages: simulation, preprocessing, and model training.

3.1. Simulation

The simulation process was performed using the CORSIKA software to model EAS initiated by gamma rays and protons, which comprise approximately 90% of the total cosmic rays at energies below 10^{19} eV (Ventura 2021). This simulation captured the interactions between primary particles and the Earth's atmosphere, including particle cascades and the production of Cherenkov light, allowing for configurations that closely match reality. The sim_telarray package was employed to simulate the detection of Cherenkov light by a telescope array. The parameters were tuned to replicate the HESS-I telescope configuration, which consists of four 12-meter diameter telescopes located in Namibia. Each simulated event could be detected by none or up to four telescopes, resulting in 0 to 4 images per event. This methodology produced a dataset of approximately 130 000 images, covering a primary particle energy range between 20 GeV and 30 TeV.

In total, 10^5 showers initiated by photons and 5×10^5 showers initiated by protons were generated. The resulting dataset of about 1.5×10^5 images, like ones showed in the Fig. 2, provides a reliable source of labeled input for CNN training.

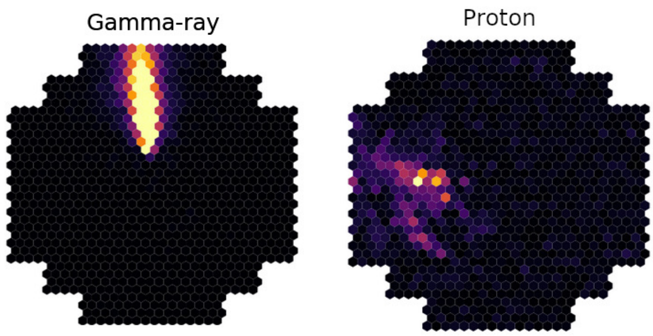

Figure 2. Simulated gamma-ray and cosmic ray events as detected by telescopes. The typical profile of a gamma-ray event is shown on the left, while a cosmic ray event is displayed on the right.

3.2. Preprocessing

Due to the format and arrangement of the photomultiplier cells in the HESS-I telescope, and in most IACTs, the pixels that form an image are hexagonal. This configuration offers certain advantages over other formats, such as greater precision in detecting higher-energy events, more efficient arrangements on curved surfaces, and more uniform angular resolution, resulting in improved event reconstruction.

However, this hexagonal pixel geometry presents challenges when the images are used with machine learning algorithms, which are typically designed to work with square-pixel formats. To surpass this issue, the solution used is to transform the data into square-pixel formats. We used the dl1-data-handler with the oversampling method (Fig.3), where each hexagonal pixel is divided into a grid of smaller square pixels, preserving spatial resolution an intensity (Nieto 2019).

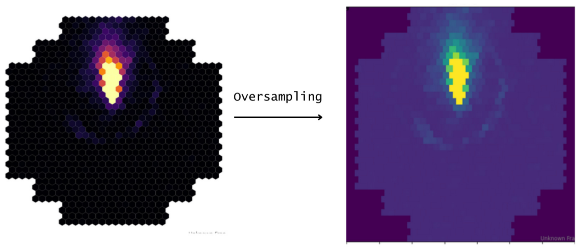

Figure 3. Comparison of input image geometries: (left) original hexagonal pixel structure captured by IACTs and (right) transformed square-pixel format obtained through the oversampling method.

3.3. Training

The dataset used in model training utilized the first 50 000 images of events generated by gamma rays and the first 50 000 images of events generated by cosmic rays. This resulted in a training dataset of 100 000 that was divided into training (80%) and validation (20%) subsets to ensure model generalization.

The CNN architecture used for this task employed multiple convolutional and pooling layers in sequence to extract features from the IACT images, followed by dense layers for final classification. The hidden layers used ReLU activation function, while output layer used a sigmoid activation function to facilitate binary classification between gamma-ray and cosmic ray events.

To enhance model performance and mitigate overfitting, the hyperparameters were adjusted throughout the training process. L2 regularization was applied to penalize excessively large weights, enhancing the model's generalization capacity.

The training process spanned 30 epochs, with metrics such as accuracy, precision, and loss was monitored to ensure continuous improvement (Fig. 4). The final parameters of the CNN are summarized in Table 1

Table 1. Training Parameters for the CNN Model

Parameter	Value
Epochs	30
Batch Size	64
Optimizer	Adam
Learning Rate	0.001
Regularization Technique	L2 Regularization
Loss Function	Binary Cross-Entropy

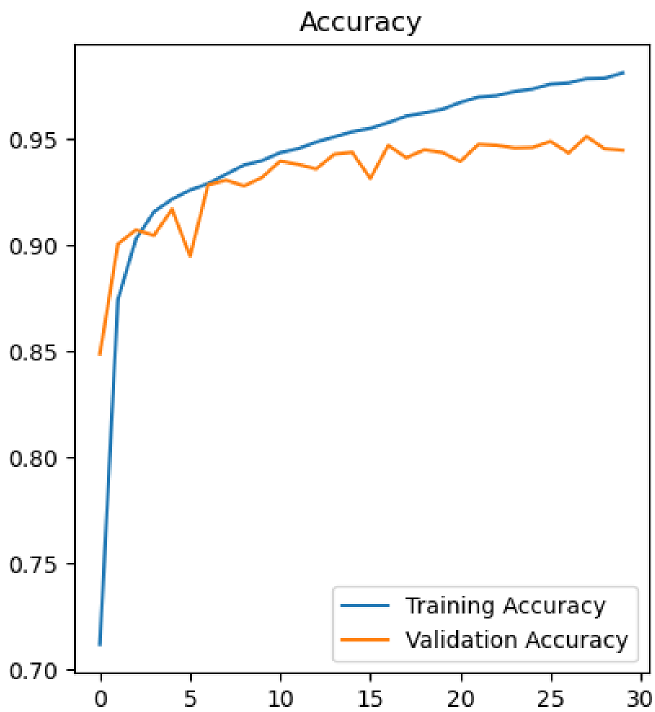

Figure 4. Training and validation accuracy of the CNN model over 30 epochs.

4. Results

The trained CNN demonstrated high performance in classifying gamma-ray and cosmic-rays events. Metrics derived from the confusion matrix (Fig. 5) revealed an accuracy of 92.1%, indicating its robustness in handling unseen data. Precision and recall metrics further highlighted the model's reliability, with values of 96.38% and 93.3%, respectively. The F1-score, a harmonic mean of precision and recall, was 94.8%.

5. Conclusions

This study demonstrated the potential of integrating simulated data with machine learning techniques for classifying extensive air shower events. Using convolutional neural networks, the trained model achieved over 90% accuracy in distinguishing between gamma-ray and hadronic events, underscoring the effectiveness of deep learning in astroparticle physics.

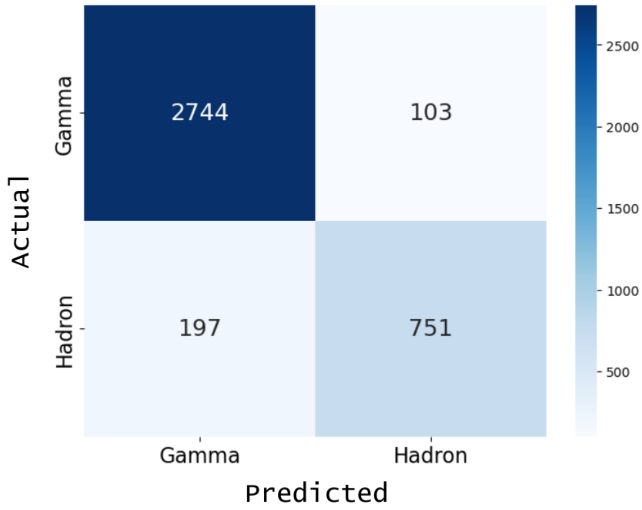

Figure 5. Confusion matrix showing the performance of the CNN model on previously unseen data.

Acknowledgements. This study was financed in part by Fundação Araucária and Secretaria de Estado da Ciência, Tecnologia e Ensino Superior through grant PDI 346/2024

References

Acharya, B. et al., 2018, Science with the Cherenkov Telescope Array, CTA Consortium (ed.), World Scientific Publishing Co.
Bernlohr, K. 2008, Astroparticle Physics, 30, 149
Grupen, C., 2020, Astroparticle Physics, Springer
Heck, D. F, et al., 1998, CORSIKA: a Monte Carlo code to simulate extensive air showers, Report number: FZKA-6019
Nieto, D. et al., 2019, Studying deep convolutional neural networks with hexagonal lattices for imaging atmospheric Cherenkov telescope event reconstruction, PoS ICRC2019, 753, doi: 10.22323/1.358.0753
Ventura, C. A. V., 2021, Study of Anisotropy, Energy and Chemical Composition of Ultra-High Energy Cosmic Rays Measured by the Pierre Auger Observatory, JCAP, 6, id. 022