Boletim da Sociedade Astronômica Brasileira, **37**, no. 1, 98-100

Multitask learning for gamma-ray event reconstruction in IACTs

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. The Imaging Atmospheric Cherenkov Technique (IACT) is essential for the study of extreme astrophysical phenomena such as active galactic nuclei (AGNs), black holes, and supernova remnants. Also, it plays an important role in the search for dark matter signatures. This technique involves using large telescopes that detect Cherenkov light emitted when high-energy particles, such as cosmic and gamma rays, interact with the Earth's atmosphere. Gamma rays, being electrically neutral, do not undergo deflection along their path and thus point directly to their sources. In contrast, cosmic rays, composed mainly of protons and atomic nuclei, are deflected by magnetic fields. Therefore, separating gamma-ray events from the cosmic-ray background is crucial in data analysis. The reconstruction of key event parameters, such as the energy and direction of the primary particle, is essential for source identification. Traditionally, the reconstruction relies on Hillas parametrization of Cherenkov images and geometric triangulation. However, with recent advances in computing, machine learning-based approaches have become increasingly promising. This work presents a multitask learning architecture based on convolutional neural networks (CNNs), capable of performing both classification of gamma and hadronic events and energy and direction reconstruction of the primary particle. The model was trained and tested on Cherenkov images generated through Monte Carlo simulations using the CORSIKA software and the sim_telarray package, ensuring complete control over simulation parameters and producing a highly reliable dataset. Preliminary results indicate that this method achieves superior angular and energy resolution compared to traditional reconstruction techniques and significant gains in processing time. This approach represents a significant advancement in the analysis of gamma-ray events. It is especially relevant in the context of the next-generation Cherenkov telescopes, such as the CTAO, which is expected to deliver unprecedented volumes of data.

Resumo. A Técnica de Imageamento Cherenkov Atmosférico (IACT) é essencial para o estudo de fenômenos astrofísicos extremos, como núcleos ativos de galáxias (AGNs), buracos negros e remanescentes de supernovas. Além disso, desempenha um papel importante na busca por sinais de matéria escura. Essa técnica envolve o uso de grandes telescópios que detectam a luz Cherenkov emitida quando partículas de alta energia, como raios cósmicos e raios gama, interagem com a atmosfera terrestre.
Os raios gama, sendo eletricamente neutros, não sofrem deflexão ao longo de sua trajetória e, portanto, apontam diretamente para suas fontes. Em contraste, os raios cósmicos, compostos principalmente por prótons e núcleos atômicos, são desviados por campos magnéticos. Portanto, a separação de eventos de raios gama do fundo de raios cósmicos é crucial na análise de dados. A reconstrução de parâmetros-chave do evento, como a energia e a direção da partícula primária, é essencial para a identificação da fonte.
Tradicionalmente, a reconstrução baseia-se na parametrização de Hillas das imagens Cherenkov e na triangulação geométrica. No entanto, com os recentes avanços na computação, abordagens baseadas em aprendizado de máquina tornaram-se cada vez mais promissoras. Este trabalho apresenta uma arquitetura de aprendizado multitarefa baseada em redes neurais convolucionais (CNNs), capaz de realizar tanto a classificação de eventos gama e hadrônicos quanto a reconstrução de energia e direção da partícula primária. O modelo foi treinado e testado em imagens Cherenkov geradas por meio de simulações de Monte Carlo utilizando o software CORSIKA e o pacote sim_telarray, garantindo controle total sobre os parâmetros de simulação e produzindo um conjunto de dados altamente confiável. Resultados preliminares indicam que este método atinge resolução angular e energética superior em comparação com técnicas de reconstrução tradicionais, além de ganhos significativos no tempo de processamento. Esta abordagem representa um avanço significativo na análise de eventos de raios gama. É especialmente relevante no contexto dos telescópios Cherenkov de próxima geração, como o CTAO, que deve fornecer volumes de dados sem precedentes.

Keywords. astroparticle physics – data analysis – gamma rays

1. Introduction

Very High Energy (VHE) gamma-ray astronomy explores the most extreme environments in the universe, probing photons with energies ranging from a few GeV to hundreds of TeV. Unlike charged cosmic rays, which are deflected by galactic and extragalactic magnetic fields, gamma rays travel in straight lines, pointing directly back to their sources. This characteristic makes them unique messengers for identifying particle acceleration sites, such as Supernova Remnants (SNRs), Active Galactic Nuclei (AGNs), and Pulsar Wind Nebulae (PWNe) (Hinton, Ong & Torres (2018)). The detection of VHE gamma rays on the ground is achieved indirectly via Imaging Atmospheric Cherenkov Telescopes (IACTs). When a high-energy gamma ray enters the atmosphere, it initiates an electromagnetic cascade (Extensive Air Shower - EAS), emitting faint blue Cherenkov light. Mirrors on the ground reflect this light into high-speed cameras, recording an image of the shower. A major challenge in IACT data analysis is the suppression of the hadronic background. The flux of charged cosmic rays (protons and nuclei) is orders of magnitude higher than that of gamma rays. Traditionally, the separation between gamma-ray events (signal) and hadronic events (noise) has been performed using parameterized methods, such as Hillas parameters, coupled with machine learning algorithms like Random Forests. However, with the advent of the next-generation Cherenkov Telescope Array Observatory (CTAO), which will feature roughly 70 telescopes and produce approximately 2 PB of raw data per day, more efficient and precise reconstruction methods are required. This work explores the application of Deep Learning (DL), specifically Convolutional Neural Networks (CNNs), to analyze data from IACTs. We propose a multitask learning architecture designed to perform simultaneous event classification (gamma vs. hadron) and parameter reconstruction (energy and direction), leveraging

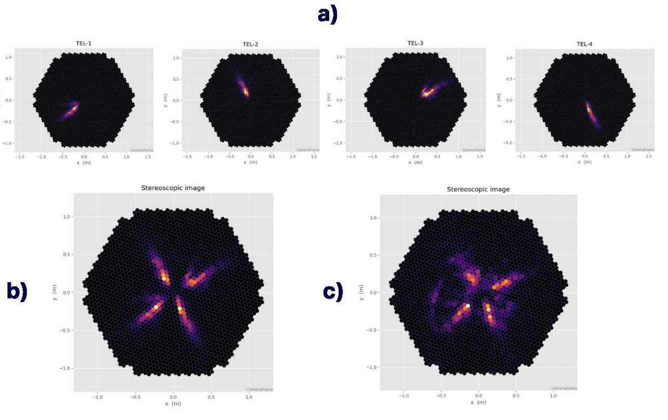

Figure 1. **a)** 4 individual telescope gamma-ray images and **b)** stereoscopic image from a gamma-ray event vs **c)** stereoscopic image from cosmic-ray. It's possible to see the distinct cherenkov profile of gamma-rays and cosmic rays showers

stereoscopic data from the 4 Large-Sized Telescopes (LSTs) of the CTAO.

2. Methodology

To develop and validate the deep learning models, a robust pipeline of simulation and pre-processing was established.

2.1. Simulations

A reliable dataset was generated using Monte Carlo simulations to ensure complete control over ground truth parameters. The interaction of primary particles (gamma rays and protons) with the atmosphere was simulated using the CORSIKA (COsmic Ray SImulations for KAscade) software (Heck et al (1998)). The simulation output was then processed by the sim_telarray package, which simulates the telescope response, including the optics and electronics of the camera (Bernlohr (2008)). The dataset includes stereoscopic images of events (Fig. 1), focusing on the configuration of the Large-Sized Telescopes (LSTs) of the CTAO, which are optimized for the low-energy range of the observatory. Specifically, gamma-ray and cosmic-ray (proton) events were simulated with energies matching the energy range of the LSTs, from 20 GeV to 150 GeV. The simulations utilized the VIEWCONE function within the CORSIKA software to model arrival directions distributed within a cone centered around zenith angles of 20º, 40º, and 60º. In total, approximately 10^6 events were simulated.

2.2. Image pre-processing

Standard CNN architectures are designed for rectangular grids of pixels. However, IACT cameras, such as those used in the LSTs, typically employ hexagonal pixel geometries to maximize light collection efficiency over an area. To make the data compatible with standard DL libraries, we applied an image mapping technique that maps the hexagonal lattice onto a square grid, preserving the spatial correlations of the shower image without significant loss of information (Nieto et al (2019)).

2.3. Neural network architecture

We implemented a multitask CNN architecture designed to process stereoscopic inputs from multiple telescopes triggered by the same event (e.g., 4 individual telescope images combined). The architecture consists of a shared convolutional backbone responsible for feature extraction, which then splits into separate "heads" for specific tasks (Fig. 2).

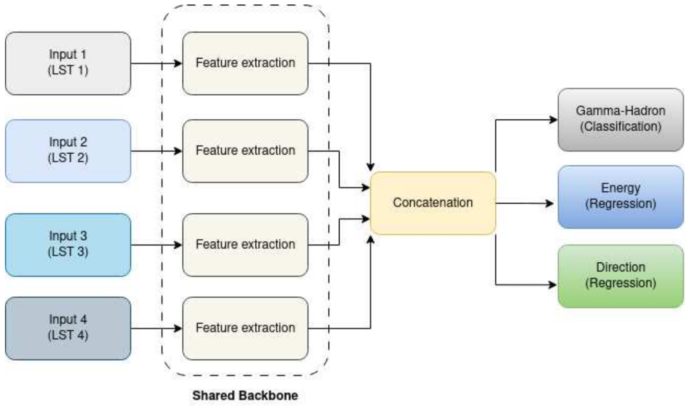

Figure 2. Architecture ofof the Multitask Neural Network. Images from four telescopes are processed by a shared backbone for feature extraction. The resulting vectors are concatenated, feeding three independent heads: Gamma-Hadron classification, Energy and Direction reconstruction.

2.3.1. Gamma-Hadron segregation

Performs a classification task and outputs a probability score distinguishing between gamma-ray and hadronic events.

2.3.2. Energy and direction reconstruction

The architecture includes two heads that perform a regression task and output continuous values for the reconstruction of the primary particle's energy and direction of arrival.

3. Training

The training process was implemented using the Keras API within the TensorFlow framework. The dataset utilized for the training phase consisted of pre-processed images, balanced equally between gamma-ray events (signal) and proton events (background) to prevent class imbalance bias. This dataset was randomly partitioned into two subsets: 80% for training and 20% for validation, ensuring the model's ability to generalize to unseen data.

4. Results

The proposed model was evaluated using the simulated dataset.

4.1. Gamma-Hadron segregation

The classification capabilities of the model were tested on a validation set containing both gamma and proton events. Quantitative evaluation metrics indicate a high performance for the stereoscopic model, as shown in Fig. 3:

- *Accuracy*: The model achieved an accuracy of 98% in distinguishing gamma rays from the hadronic background.

- *Loss*: The Binary Cross-Entropy (BCE) loss converged to 0.014.

- *AUC*: The Area Under the ROC Curve (AUC) reached 0.99, demonstrating excellent discrimination power across different threshold settings (Fig. 4).

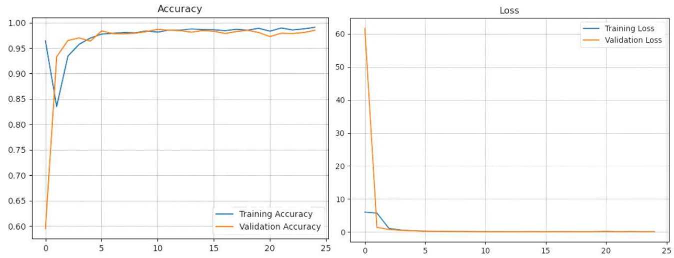

FIGURE 3. Left: Training (blue) and validation (orange) accuracy across 25 epochs. **Right**: The corresponding binary cross-entropy loss curves.

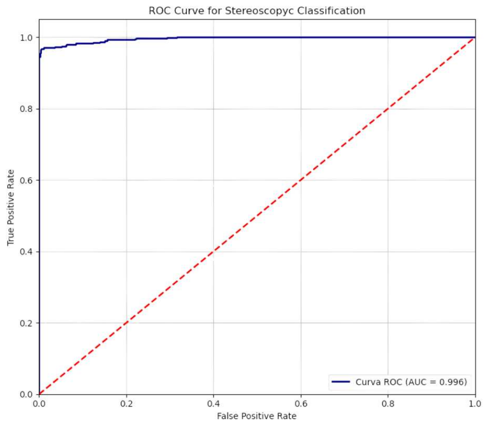

FIGURE 4. The model achieves an Area Under the Curve (AUC) of 0.996, indicating excellent discriminatory power between Gamma-ray and Hadron events. The dashed diagonal line represents the performance of a random classifier.

The Confusion Matrix and ROC curves derived from the testing phase confirm that the false positive rate (hadrons classified as gammas) is kept minimal, which is crucial for maximizing the sensitivity of the observatory to faint gamma-ray sources (Fig. 5).

4.2. Energy and Direction Reconstruction

Beyond classification, the multitask network simultaneously performed regression tasks to reconstruct the primary particle's energy and arrival direction. Both tasks were evaluated using the 68% containment metric, which minimizes the impact of outliers. The model demonstrated an angular resolution of $\Psi_{68} \leq 0.7°$. With an energy resolution of $(\Delta E/E)_{68} \leq 30\%$, at the LST energy range. These results, illustrated in Fig. 6, confirm that the shared feature representation effectively benefits regression tasks without compromising the classification performance.

5. Conclusion

This work demonstrates the efficacy of Multitask Deep Learning for the analysis of IACT data. By utilizing CNNs trained on CORSIKA and sim_telarray simulations, we achieved a classification accuracy of 99% and an AUC of 0.99 for gamma-hadron separation using stereoscopic LST data. Furthermore, the proposed architecture achieved robust regression capabilities, with

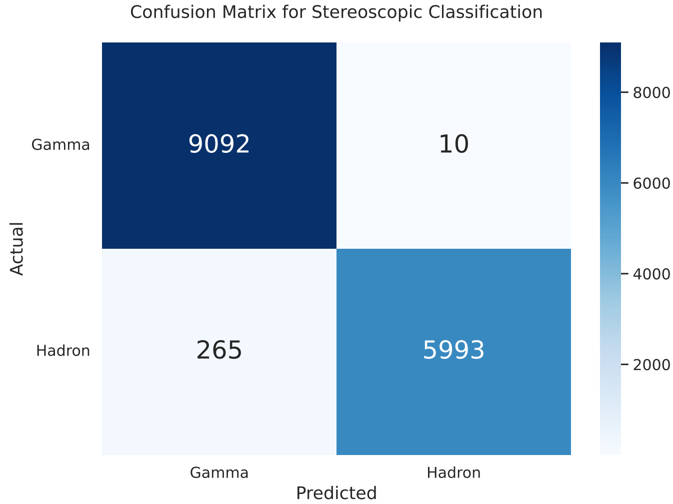

FIGURE 5. Confusion Matrix of the Stereoscopic Classification Model. The matrix displays the performance of the classification network on the test set, distinguishing between Gamma-ray events (positive class) and Hadron events (negative class).

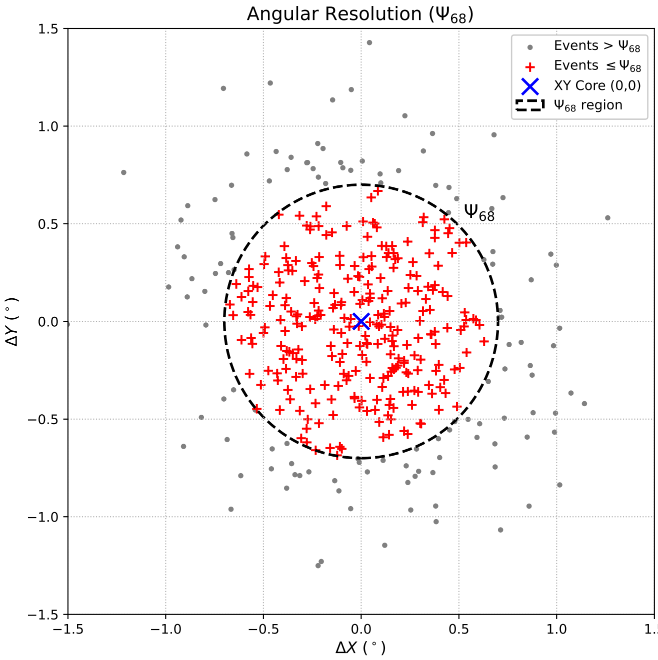

FIGURE 6. The plot shows the angular reconstruction error relative to the true source position. The dashed circle marks the 68% containment region.

an arrival direction angular resolution of $\Psi_{68} \leq 0.7°$ and energy resolution of $\Delta E/E \leq 30\%$ at low energies.

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

Bernlohr, K. , 2008, Astroparticle Physics, 39, 149
Heck, D. F, et al., 1998, Tech. Rep. FZKA-6019.
Hinton, J.A., Ong, R.A., Torres, D., 2018, in Science with Cherenkov Telescope Array, The Cherenkov Telescope Array Consortium, 12
Nieto, D, et al., 2019, PoS(ICRC2019)753