ADASS XXXV · 2026

Machine-Learning-Based Event Classification of Terrestrial Gamma-ray Flashes with a Segmented Scintillator Detector

Authors

Antonio Femia[1], Nicolò Parmiggiani[1], Andrea Bulgarelli[1], Valentina Fioretti[1], Luca Castaldini[1], Mattia Maniscalco[1], Gabriele Panebianco[1], Ambra Di Piano[1], Riccardo Falco[1], Ghulam Mustafa[1], Alessandro Rizzo[2], Adriano De Rosa[1], Enrico Virgilli[1], Claudio Labanti[1], Martino Marisaldi[3], Francescopiero Calzolari[4], and Angela Marinoni[4]

[1]Osservatorio di Astrofisica e Scienza dello spazio di Bologna (OAS), Via Piero Gobetti, 93/3, Bologna, 40126, Italy
[2]INAF - Osservatorio Astrofisico di Catania, Via Santa Sofia 78, Catania, 95123, Italy
[3]Birkeland Centre for Space Science, Department of Physics and Technology, University of Bergen, Norway
[4]CNR-ISAC, Via P. Gobetti 101, I-40129 Bologna, Italy

INTRODUCTION

Terrestrial Gamma-ray Flashes (TGFs) are sub-millisecond bursts of atmospheric gamma rays associated with thunderstorms and are of particular interest for space-based gamma-ray observations[1,2,3]. Their detection from orbit requires the identification of brief gamma-ray signals in the presence of background, while operating under the limited computational and power resources typically available on CubeSat-class spacecraft. Within this context, the **Gamma-Sky project**[4] is investigating low-power Edge-AI approaches for the real-time processing of high-energy radiation, using atmospheric X- and gamma-ray detector data. **This work presents a machine-learning-based event-classification methodology** for a CubeSat-class mission, exploiting the spatial and energy information provided by a segmented scintillator detector architecture[5]. **The detector comprises 500 GAGG (Gadolinium Aluminum Gallium Garnet $Gd_3Al_2Ga_3O_{12}$) scintillator crystals ($1\times1\times0.8\ cm^3$ each), with a $0.2\ cm$ Z-axis gap between elements, filling 50% of the total detector volume.** This inter-crystal gap is designed to host the signal readout cables and the associated electronics. The pixelized detector architecture enables the identification of characteristic event patterns related to deposited energy, multiplicity of interactions, and spatial correlations between detector elements. The detector response is used to extract a compact set of event-level features suitable for machine-learning analysis.

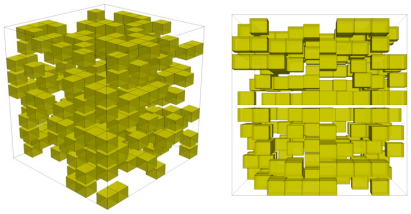


Figure 1: Detector architecture. Left figure: 3D view. Right figure: XZ projection

METHODOLOGY

- All simulations have been conducted using **MEGAlib (Medium Energy Gamma-ray Astronomy Library):** an open-source software framework specifically developed for the simulation and analysis of gamma-ray detectors operating in the keV–GeV energy range.
- Seven background charged-particle populations were considered:
  - ✓ **Electrons (e^-):** primary and secondary populations.
  - ✓ **Positrons (e^+):** primary and secondary populations.
  - ✓ **Alpha particles ($^4He^{2+}$):** primary population.
  - ✓ **Protons (p^+):** primary and secondary populations, with secondary protons distinguished into upward-going (from Earth) and downward-going (from the sky).

  The background simulations are based on the integral particle flux ($1000\ s$) expected in low Earth orbit (LEO) at **$550\ km$ altitude**, assuming an orbital inclination of 0°. Flux values are expressed in $cm^{-2}\ s^{-1}$.

- The TGF spectrum follows an RREA-like exponentially cut-off power law:

$$\frac{dN_\gamma}{dE} \propto E^{-1}\exp\left(-\frac{E}{7.3\ MeV}\right)$$

  Energy range: $10\ keV - 40\ MeV$.

SIGNAL AND BACKGROUND ANALYSIS

- **The simulated charged-particle background** produces a total **event rate of 82.2 Hz** and a **hit rate of 507.1 Hz**, corresponding to an average multiplicity of **6.2 hits/event**.
- **Secondary positrons** dominate the event rate (29.1 Hz), followed by **primary protons** (18.5 Hz). Together, these populations account for approximately **58% of charged-particle events**.
- **Primary protons and secondary positrons** dominate the total hit rate (174 Hz and 169 Hz, respectively), reflecting their different mean hit multiplicities (9.4 and 5.8 hits/event).
- **Alpha particles** exhibit a high multiplicity (~14 hits/event), despite their small contribution to the event rate (2.78 Hz).
- **Primary electrons and positrons** contribute negligibly to the total background rate, although their simulated events exhibit high average hit multiplicities.

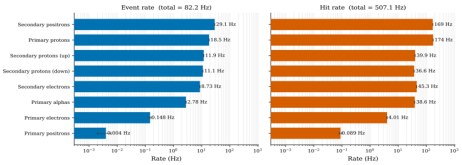


Figure 2: Simulated background rates for the different incident charged particle populations. Left: Event rate, counting each detected event once, independently of its hit multiplicity. Right: Hit rate, accounting for all detector hits produced by each event.

- **TGF photons exhibit predominantly single-pixel interactions**, with a mean hit multiplicity of **1.25 pixels/event**. Approximately 83% of events activate a single pixel, while only 0.24% involve six or more pixels.
- **Charged particles produce significantly higher hit multiplicities**, particularly primary electrons (27.1 pixels/event), alpha particles (13.9) and primary protons (9.4). Secondary electrons, positrons and protons exhibit lower average occupancies.
- **Multi-pixel events are associated with charged-particle backgrounds:** approximately 31% of charged-particle events involve six or more pixels, compared with only 0.24% of TGF photon events.

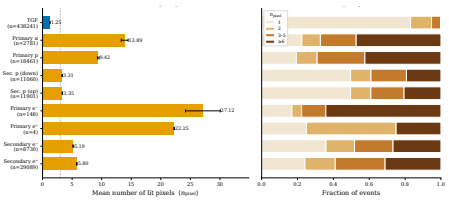


Figure 3: Hit Multiplicity and energy dependence for TGF photons and charged-particle backgrounds. Top panels show the mean hit multiplicity and its distribution across particle populations. The bottom panel illustrates the relationship between primary particle energy and hit multiplicity.

MACHINE LEARNING

To extract informative spatial and topological features, only events activating **at least three pixels ($n_{pixel} \geq 3$)** are selected, retaining **54.83% of charged-particle background events and 5.23% of TGF events**.
A total of **49 features** are obtained:

- **Energy Deposition & Multiplicity:** Number of hit voxels, total deposited energy, statistical moments, percentiles, and energy dispersion.
- **Energy Sharing & Concentration:** Energy entropy, Gini coefficient, active-voxel fractions, and energy concentration in the most energetic voxels.
- **Spatial Geometry & Topology:** Energy-weighted centre of mass, radius of gyration, spatial covariance eigenvalues, linearity, planarity, sphericity, asymmetry, and spatial extent.
- **Hit Statistics:** Total hit count, mean hits per active voxel, and energy–hit correlation.
- **Morphological Complexity:** Fractal dimension, connectivity, maximum path length, branching structure, and skeleton length.

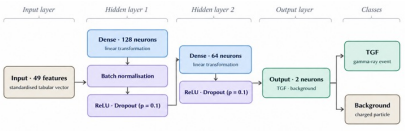


Figure 4: Fully connected neural network adopted for the classification task.

- **ROC AUC of 0.981.**
- **TGF detection efficiency: 97.8%** of simulated TGF photon events are correctly identified (6,725/6,876).
- **Charged-particle rejection: 89.6%** of background events are correctly rejected (12,2105/13,517).
- **Overall classification accuracy of 92.9%,** with a signal precision of **84.1%.**

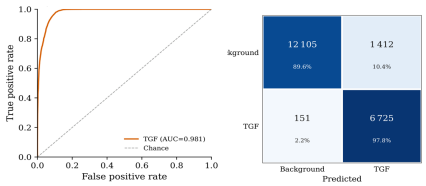


Figure 5: Machine Learning outcomes considering the test dataset. Left: ROC-AUC curves. Right: Confusion matrix.

- **In the 10–30 MeV range** the model reaches a signal recall of **90.9%,** a background rejection of **90.4%**, with a macro-F1 score of **0.897**.
- **99.3%** of background events above 30 MeV are correctly classified, while preserving **86.4%** of TGF events.
- TGF recall remains above **94% throughout the 1–10 MeV range**, demonstrating the classifier's ability to preserve photon-induced events.
- Rejection efficiency increases from **19.5% at 2–5 MeV** to **53.3% at 5–10 MeV,** reaching **90.4% at 10–30 MeV**. This identifies the lower-energy region as the main challenge for background suppression.

E_0 (MeV)	N	Accuracy	$F_{1,\text{macro}}$	Recall		Precision	
				R_{TGF}	R_{bkg}	P_{TGF}	P_{bkg}
0.100–0.300	128	0.977	0.494	1.000	0.000	0.977	0.000
0.300–0.511	597	0.987	0.497	1.000	0.000	0.987	0.000
0.511–1.000	1 309	0.976	0.494	1.000	0.000	0.976	0.000
1–2	1 745	0.777	0.489	0.996	0.057	0.776	0.821
2–5	2 201	0.776	0.593	0.991	0.195	0.769	0.892
5–10	1 753	0.803	0.759	0.949	0.533	0.789	0.850
10–30	2 334	0.906	**0.897**	0.909	0.904	0.827	0.952
> 30	10 326	0.993	0.671	0.864	0.993	0.216	1.000

CONCLUSIONS

The **segmented GAGG detector** combines the **high density (6.63 g/cm³) and high effective atomic number Z_{eff}** of GAGG, providing efficient gamma-ray absorption in a compact volume, with a modular and mass-optimised architecture suitable for CubeSat-class missions. Its **three-dimensional pixelation** enables spatially resolved energy measurements and the extraction of **energy-deposition, multiplicity, and topological features**, supporting machine-learning-based discrimination between TGF photons and charged-particle backgrounds. Furthermore, the modular architecture offers **enhanced fault tolerance**: the failure of individual SiPMs or scintillator channels does not necessarily compromise the entire instrument, allowing continued operation with gradually reduced detection performance. Selecting events with $n_{pixel} \geq 3$ retains **5.23% of TGF events** and **54.83% of charged backgrounds**, preferentially selecting higher-energy photon interactions. On the selected test sample, the **fully connetected neural network classifier achieves an AUC of 0.981**, with **97.8% TGF recall** and **89.6% charged-background rejection**. In the **10–30 MeV range**, signal recall reaches **90.9%,** with **90.4% background rejection** and a macro-F1 score of **0.897**. Above **30 MeV**, charged-particle rejection increases to **99.3%**, demonstrating high-energy background discrimination. These results highlight the **potential of topology-based machine learning for high-energy TGF identification** in compact, CubeSat-class detectors, motivating further studies of end-to-end detection efficiency and onboard implementation.

REFERENCES

[1] Lindanger, A, et al. (2021). "Spectral Analysis of Individual Terrestrial Gamma-Ray Flashes Detected by ASIM". In: Journal of Geophysical Research: Atmospheres 126, e2021JD035347. doi: 10.1029/2021JD035347.
[2] Marisaldi, M., et al. (2010) "Detection of terrestrial gamma ray flashes up to 40 MeV by the AGILE satellite". In: Journal of Geophysical Research: Space Physics 115, A00E13. doi: 10.1029/2009JA014502.
[3] Briggs, M. S., et al. (2010). "First results on terrestrial gamma ray flashes from the Fermi Gamma-ray Burst Monitor". In: Journal of Geophysical Research: Space Physics 115, A07323. doi: 10.1029/2009JA015242.
[4] Castaldini, L., et al., (2026) "GammaSky: real-time edge processing for high-energy events," in Software and Cyberinfrastructure for Astronomy IX, Proc. SPIE, vol. 14155, Art. no. 141551Q, 2026. doi:10.1117/12.3105073.
[5] Saleh-Natur, J., et al., "GALI — A GAmma-ray burst Localizing Instrument," Nuclear Instruments and Methods in Physics Research Section A, vol. 1084, Art. no. 171226, 2026.

INAF/OAS Bologna (Bologna Observatory of Astrophysics and Space Science] · [antonio.femia@inaf.it] · [https://www.oas.inaf.it/en/]