







### Deep Learning Models to Analyze Gamma-ray sky maps and time series

N. Parmiggiani, A. Bulgarelli, A. Macaluso, L. Baroncelli, L. Castaldini, A. Di Piano, R. Falco, F. Farsian, V. Fioretti, G. Panebianco, A. Rizzo INAF/IASF Milano 18/12/2024







## My Background



- Master's Degree in Management Engineering
- PhD in Computer Engineering and Science. Main Focus: Astronomical data analysis, and AI
- Technology Researcher at Istituto Nazionale di Astrofisica Bologna
- Main research activities: development of real-time analysis software and machine learning algorithms to analyze gamma-ray data
- Space projects: AGILE (ASI) and COSI (NASA)
- Ground projects: Cherenkov Telescope Array Observatory and ASTRI Mini-Array
- nicolo.parmiggiani@inaf.it
- https://orcid.org/0000-0002-4535-5329

## Summary

- Introduction
- Deep Learning models for the AGILE space mission
- Deep Learning models for the COSI space mission
- Deep Learning models for the Cherenkov Telescope Array Observatory



## **Context and Research Goals**



- This research aims to develop **Deep Learning** (DL) and **Quantum Deep Learning** models to analyze the data acquired by the **AGILE** instruments to detect and localize **Gamma-Ray Bursts**.
- We are developing a **DL model for COSI to localize the GRBs** using simulated data acquired by its detectors.
- We developed DL models to analyze the **simulated data of the Cherenkov Telescope Array Observatory**.
- We developed DL models to analyze **sky maps** as 2D images and **time series** and we approached different **classes of problems**:
  - Binary Classification
  - Anomaly Detection
  - Regression





0.00067 0.00135 0.00203 0.00270 0.00338 0.00406 0.00473 0.00541 0.00608

## AGILE satellite



• AGILE is an ASI space mission launched in 2007, designed to study X-ray and gamma-ray astronomy. AGILE terminated the in-orbit operations on February 2024, after almost 17 years of successful scientific observations.



## **COSI** satellite

- The Compton Spectrometer and Imager (COSI) is a NASA Astrophysics Small Explorer satellite mission.
- COSI is a soft gamma-ray survey telescope (0.2-5 MeV) planned for launch in 2027.
- It is designed to probe the origins of Galactic positrons, uncover the sites of nucleosynthesis in the Galaxy, perform pioneering studies of gamma-ray polarization, and find counterparts to multi-messenger sources.
- COSI's compact Compton telescope combines improvement in sensitivity, spectral resolution, angular resolution, and sky coverage to facilitate groundbreaking science.



## Cherenkov Telescope Array Observatory



Photos: Gabriel Pérez Dia and Marc-André Besez

#### **Cherenkov Telescope Array Observatory**

With tens of telescopes among two sites **CTAO** will have unprecedented **sensitivity**, high **angular resolution**, broad **sky coverage** and a wide **energy range (20 GeV - 300 TeV)**.

**CTAO** will optimise its scientific return with a **real-time analysis system**, enabling the scientific community to become a key player in the **multiwavelength and multimessenger** landscape of modern astrophysics.

## **Deep Learning Models for AGILE**

### Detection of GRBs in the AGILE/GRID data

- We developed a Convolutional Neural Network (CNN) to detect GRBs in AGILE/GRID sky maps (0.1–10 GeV) and compared the results obtained with the standard aperture photometry (Li&Ma) methods.
- The maps have a size of **100** × **100 pixels** and a bin size of 0.5°. The integration time used to generate these maps is **200** seconds.
- We simulated three datasets of 40 000 maps for the training, testing, and validation phases. The CNN is trained with a supervised learning technique, so the datasets are labeled.
- Half of the maps are background-only maps, and the other half have a simulated GRB in a radius of 1° from the center.
- The **GRB fluxes used for the simulations are randomly generated** from a distribution obtained from the Second Fermi/LAT GRB Catalog, rescaling the fluxes to fit the AGILE/GRID energy range.

Figure 8. Smoothed intensity maps from the simulated dataset used to train the CNN. The top two images are background-only. The bottom two images contain a simulated GRB. No other  $\gamma$ -ray sources are present inside the maps. The maps are represented in ARC projection and Galactic coordinates, with a bin size =  $0.5^{\circ}$ .





### Model training and evaluation

- We developed this **binary classification** model using 2D convolutional layers, max polling and dropout layers.
- The CNN was trained for 5 epochs using a batch size of 200 maps.
- The model reached on the test set an accuracy of 98.2 %
- We calculate the **p-value distribution** using datasets of **ten million** simulated maps for three different background levels.
- We obtained different p-value distributions that are used to calculate the **statistical significance** of a detection obtained with the CNN model.







## Detection of GRBs in the AGILE/GRID data



- We **evaluated the CNN** using the Fermi-LAT, Fermi-GBM, and Swift-BAT GRB catalogs using data from 2010 to 2020.
- We analyzed AGILE/GRID data covering the trigger times and positions of cataloged GRBs with **both the CNN and the standard Aperture Photometry** methods.
- The CNN detected **21 GRBs with a sigma > 3** from the list of GRBs obtained with Fermi and Swift catalogs. Using the same parameters, the Aperture Photometry detected **only two GRBs** from that list.

Parmiggiani N. et al. A Deep Learning Method for AGILE/GRID Gamma-ray Bursts detection, <u>Astrophysical Journal</u>, Volume 914, Issue 1, id.67, 12 pp (2021)

Parmiggiani, N., National Prize for Artificial Intelligence and Big Data research, WMF and IFAB 2021. Media INAF

## AGILE/GRID - Improvement

INAF ISTUTO NAZIONALE DAS DASTROPISICA DASTROPISICA

- To cover science alerts with larger localization error (i.e. Gravitational Waves), we trained a new DL model to detect GRBs in a maximum radius of 20 degrees from the center of the map.
- In addition, we implemented a new DL model to localize the GRB position in maps classified from the first model as containing a source. This model performs a regression on the coordinates of the GRBs.
- The classification model achieves an accuracy of 97 %, while the localization model has a mean error of 0.7 degrees.

Parmiggiani N. et al. Preliminary Results of a New Deep Learning Method to Detect and Localize GRBs in the AGILE/GRID Sky Maps. proceedings of the ADASS XXXII (2022) conference <u>arXiv</u>





## Anomaly detection in the AGILE Anticoincidence system ratemeters

- We developed a Deep Learning model to detect GRBs in the ratemeters of the AGILE Anticoincidence System.
- The Anticoincidence System comprises five independent panels surrounding the AGILE detectors
- This system aims to reject charged background particles.
- It can also **detect hard X-ray** photons in the energy range of 50 200 keV.
- The ACS continuously records each panel count rate in telemetry as ratemeters (RM) data, with 1.024 seconds resolution. Each ACS panel RM count rate constitutes a time series.





Figure 1. The AGILE Anticoincidence System. Credits: F. Perotti and the AGILE Team.

## Neural Network Design and Training

- We developed a Deep Learning model to detect GRBs in the ratemeters of the AGILE Anticoincidence System.
- The ACS can detect GRBs and solar flares. We want to detect GRBs so we avoided using the panel oriented toward the Sun due to its sensitivity to solar flares, which interferes with GRBs detection.
- We have to apply a **detrending procedure** to remove orbital and spinning modulations from the data.
- we decided to analyze time windows of 140 seconds with bins of 1.024 seconds, by analyzing the T\_50 and T\_90 of third Swift/BAT GRB catalog (Lien et al. 2016).







## Neural Network Design and Training

- We implemented the model with a **1D Convolutional Neural** Network autoencoder.
- The autoencoder aims to encode the input data in a representation with reduced dimension and then decode this representation to the original input object, **minimizing the reconstruction error**.
- An autoencoder can be used for anomaly detection because when the input is different from the usual (e.g. a GRB is present), the reconstruction error is higher.
- We trained the model with an **unsupervised technique** using 5000 background-only time series randomly extracted from the ACS data archive, excluding time windows with anomalies (e.g., South Atlantic Anomaly passages and known GRBs)





## **Results on real GRBs**

- To define the thresholds at different sigma levels, we calculated the p-value distribution using 15 million background-only time series generated using the bootstrap data augmentation technique
- We evaluated the trained model using the list of GRBs present in the **GRBweb catalog** (2010-2020), which collects GRBs from several observatories.
- We extracted from the AGILE archive the time series covering the trigger time of the catalog.
- The model **detected 72 GRBs**, **15 of which** were detected for the first time in the AGILE data

INAF ISTUTO NAZIONALE DI ASTROPISCA DI ASTROPISCA DOLOGNA



Parmiggiani N., Bulgarelli A., Ursi A. et al. "A Deep-learning Anomaly-detection Method to Identify Gamma-Ray Bursts in the Ratemeters of the AGILE Anticoincidence System", <u>Astrophysical Journal</u>, Volume 945, (2023)

INAF mini-grant (2022) to carry on these researches based on Deep Learning techniques.

## Anomaly detection on MCAL data

- We trained a DL model following the method used with the anticoincidence ratemeters to detect GRBs in the MiniCALorimeter (MCAL) ratemeters.
- We compared the preliminary results obtained with this method analyzing the MCAL ratemeters of 2020 and 2019 with the second MCAL GRB catalog, confirming the detection of 26 GRBs.
- We plan to analyze the full AGILE data archive.





0.0

0

### Predict the ACS background

- **Goal**: predict the background count rates of the AGILE ACS system using the satellite **orbital and attitude parameters.**
- We can use the predicted counts of the background to detect
  GRBs where the differences with the acquired counts are higher than a predefined threshold.
- We used the ACS top panel for this work because it is less influenced by solar flares.
- There are three main trends that impact the AGILE ACS data:
  - i) the daily trend
  - $\circ$  ii) the orbital trend with a period of ~ 94 min
  - iii) the spinning trend.





Similar approach used by R. Crupi et al. Searching for long faint astronomical high energy transients: a data driven approach. EA 2023

### Deep Learning model

- We trained a Feedforward Neural Network model to perform a regression task using 20 million orbital configuration parameters.
- We calculated the **difference** between the **real** and **predicted** counts of the test dataset to check the accuracy of the model. The model has a mean prediction error of 3.8%.
- We can apply this detection method to raw data without applying the detrending algorithm that can introduce artificial anomalies.
- We calculated the **p-value distribution** using 20 million light curves to define the thresholds at different sigma levels.





### Detect GRBs using the predicted values

- We tested this detection method using the **GRB web catalog** and extracting light curves from the ACS archive (2019-2022).
- The method detects 39 GRBs with sigma > 3. Four GRBs are new detections that were not detected in previous analyses.
- We also compared the results obtained with the light curve of the Fermi/GBM detector because they have a similar energy range. ACS (50-200 keV) and Fermi/GBM (50-300 keV)
- We are investigating other possible applications of this kind of Deep Learning model to predict the background level of the AGILE detectors.



Parmiggiani N., et al. "A New Deep Learning Model to Detect Gamma-Ray Bursts in the AGILE Anticoincidence System", <u>Astrophysical Journal</u>, Volume 973, (2024)

DAS

## Simulation of GRB light curves

### Deep Learning to simulate GRB light curves

- Goal: simulate the light curves of GRBs using Deep Learning generative architectures such as Generative Adversial Network (GAN) and Variational Autoencoder (VAE).
- The training of the DL model is done using the light curves of the **fourth Fermi-GBM GRB** catalog after applying filters to remove light curves not suitable for this study.
- There is a work in progress to use **conditional/controlled GAN** to generate dataset with specific GRB properties (e.g. duration.). In addition, we are evaluating the **Physics-Informed Neural Network**.



## **Quantum Deep Learning**

### Quantum Computing and Deep Learning



- INAF is working on Quantum Computing inside the Spoke 10. We provide scientific use cases to test Quantum Computing technologies.
- At INAF/OAS Bologna we are working in the context of **Quantum Machine Learning**.
- We are developing Quantum Deep Learning models to detect Gamma-Ray bursts from sky maps and time series.



INAF Spoke 10 (quantum computing) members: A. Bulgarelli, C. Burigana, V. Cardone, F. Farsian, M. Meneghetti, G. Murante, A. Rizzo, R. Scaramella, F. Schillirò, V. Testa, T. Trombetti.

### Why Quantum Computer?



- **Goal**: Develop Quantum Deep Learning models to **exploit the features of Quantum Computers** for the analysis of the data acquired by the AGILE satellite.
- **Compare the results** obtained with Quantum Deep Learning models with those obtained with classical Deep Learning models to check for improvements (e.g., Quantum speed-up, fewer parameters etc)
- We used as benchmark the models already developed with classical Deep Learning models. Also the dataset of AGILE maps and time series are the same.
- At the moment we are working also on Cherenkov Telescope Array Observatory and COSI simulated data.

### **Embedding Techniques**



- To use classical data with quantum computers we have to represent it as a **quantum state**, so that it can be processed by a quantum computer.
- The classical data can be encoded into a quantum state by using a quantum circuit.
- A quantum embedding technique takes a classical datapoint x and translates it into a set of gate parameters in a quantum circuit, creating a quantum state.
- We evaluated three embedding techniques:
  - Angle embedding
  - Amplitude embedding
  - Data re-uploading



### Frameworks

- We used three different frameworks for the implementation of the quantum models:
- **Tensorflow-Quantum** is an open-source framework developed by Google for rapid prototyping of hybrid-quantum classical models. It combines TensorFlow for classical machine learning with quantum computing.
- **Qiskit** is an open-source software development kit, developed by IBM Research. It is possible to run the circuits on real quantum devices with the IBM Quantum Experience.
- PennyLane is an open-source software framework for quantum machine learning. It can be used with different quantum hardware and simulators. It provides interfaces with classical machine learning frameworks.







### **Implemented Architectures**



- We used both **Hybrid and Fully quantum** approach to implement models that can analyze the time series and the sky maps:
- Hybrid: The feature extraction is made using a quantum convolutional neural network, and the final classification is done with a classical neural network.
  - Pros
    - Reduced decoherence effect (on real QC) because the quantum circuit is shorter
    - Flexibility to combine the benefits from quantum and classical models







#### Time Series GRB classification

Approach	Framework	Accuracy on Training Dataset	Accuracy on Test Dataset	Parameters	Qubits
Classical	Keras & TensorFlow	96.5%	93.9%	56	//
Fully Quantum	Qiskit	91.1%	90.7%	14	7

- The results shows that the Quantum DL models can achieve a **comparable accuracy** with the classical models but using **less parameters** in the model.
- The next step is to test these models with **real quantum computers**.

A. Rizzo et al.: "Quantum Convolutional Neural Networks for the detection of Gamma-Ray Bursts in the AGILE space mission data", ADASS 2023, <u>https://arxiv.org/pdf/2404.14133</u>.

F. Farsian et al.: "Benchmarking Quantum Convolutional Neural Networks for Signal Classification in Simulated Gamma-Ray Burst Detection", 2025 submitted to IEEE Computer Society, proceeding of "Astrophysics and Cosmos Observation: HPC and Big Data Management" conference.

## **Deep Learning for COSI GRB localization**

#### 31

### Localization of GRBs using BGO and GeDs data

- The model aims to localize the GRBs using the count rates of the BGO shield composed of five panels.
- In addition, to improve the results, we used the counts detected by the Germanium detectors (GeDs).
- We simulated 50 000 GRBs (without background) at different sky coordinates to create our labeled dataset.
- We calculated the ratios between the integral of counts detected by different panels to have a measure independent from the flux of the GRB.
- These ratios are the input of the DL model and the GRB positions are the labels.



Time (s)



### Model Training



- The model used is a **feedforward neural network** with **four hidden fully connected layers**. The model has 43k trainable parameters.
- We use dropout layers for the regularization and to prevent overfitting.
- The activation functions are LeakyReLU and the cost function is the Mean Square Error.
- The model is trained for 185 epochs using the early stopping feature from Keras.



### First results



- Mean localization error of **12.2**° for the full sky and 7.2° for Theta > 90.
- The error is higher when Theta < 60° (in the aitoff plot >30)



### Conversion of angles in sine and cosine



- Instead of using theta, phi as target coordinates for the models we converted the two angles expressed in degrees to sine and cosine values.
- This avoid the "wrap around problem" -> jump from 360 to 0
- High localization error when theta is around 45°.

#### before:

12.2° of mean localization error. 7.2° if theta > 90°



#### after:





### Additional data: GeDs counts

- We can add the counts collected by the GeDs to **improve the** localization when the angle Theta < 60° (see figure).
- We use the four columns of GeDs divided into two layers -> 8 count rates and we calculate the ratios between these count rates



ti-coincidence Shield Modules (6x)





### Final Results

- We trained again the model using both the count rates from BGO and GeDs
- The results with BGO data only have an issue at Theta near 45 ° but by adding the GeDs data this issue is solved.

Only BGO: mean error is 7.2° Loc. error at 40-50° is 31°

With GeDs: mean error is 4.8° Loc. error at 40-50° is 8.4°













### Summary of results

- With the current simulations (without background) the model can localize a GRB with a **mean localization error of 4.8°** using the data of BGO and GeDs.
  - These position determinations will complement **COSI's Compton localizations**.
- When Theta ≈ 45° the BGO data cannot localize the source with a low error. We introduced the GeDs counts to help the model reduce the localization error from 31 to 8.4°.
- We are now evaluating the **impact of the background** noise and developing an additional method (used by GBM) based on the **chi-squared fitting**.

N. Parmiggiani et al. "Deep learning techniques to detect and localize Gamma-ray Bursts in sky maps and time series acquired by the AGILE and COSI space missions." Seventeenth Marcel Grossman Meeting, Pescara, 7-12 July 2024.

# Deep Learning to analyzie CTAO sky maps

## Credits by Ambra Di Piano

### **Deep Learning Enhancements**

#### CONTEXT:

- high level Cherenkov data analysis (DL3) in real-time
  - <sup>1</sup>ACADA <sup>2</sup>Science Alert Generation
- machine learning to overcome limitations of real-time
  - > variability of conditions, degraded sensitivity, lack of knowledge

#### **GOALS**:

- perform background-subtraction without requirements on
  - knowledge on background
  - knowledge on target coordinates
- perform candidates localisation without requirements on
  - knowledge on background

#### MODELS:

- Cleaner: CNN autoencoder
  - encoding with Conv2D and AvgPooling2D
  - decoding with Conv2D and UpSampling2D
  - trained with two sets of data: noisy and clean counts map

#### Regressor: CNN regressor

- combination of Conv2D and MaxPooling2D
- dense neurons and dropout layers
- > trained with a set of data and a set of known coordinates as labels



### Simulations

#### **Fixed parameters:**

- One source in the FoV
- Flux and spectral model (crab-like)
- Array (4LST) from prod5-v0.1
- Binning (200x200) and pixel (0.025 deg)
- Exposure (100 s) and smoothing (5 $\sigma$ )
- IRF background-subtraction

#### **Random parameters:**

- Background/IRFs (zenith and NSB)
- Pointing and source coordinates
- Source offset (within FoV)

#### Planned variations (coming next):

- Vary exposure and flux
- Multiple sources in FoV
- Ring background-subtraction (known targets)



### **Background Subtraction**

#### **EVALUATION METRIC - SOURCE EXCESS COUNTS**

For each map in the sample we compute the source excess counts, by integrating the number of photons within the source region after applying the CNN denoising.

#### **CNN EXCESS =** $\sum$ **DENOISED SOURCE COUNTS**

We use the standard aperture photometry method as a reference, from the photometric excess counts we subtract the CNN excess counts for each map in the sample.

ERR = (CNN - PHOTOMETRIC) / PHOTOMETRIC

We then compute the error percentage E(%), where  $N_s$  is the photometric excess and  $N_s^{cnn}$  is the CNN excess.



### **Candidate Localization**



A. Di Piano et al., "Machine Learning Enhancements for Real-Time Scientific Analysis of Cherenkov Telescope Data", ADASS, Nov 2024, Valletta (Malta) Proceeding to be published as part of ASP Conference Proceedings

## Anomaly detection to detect GRBs in the CTAO light curves

## Credits Leonardo Baroncelli

### Context

- This work introduces an anomaly detection technique for identifying GRB events from light-curves in real-time analysis scenarios with short-term exposures.
- > Anomaly detection setting:
  - We define **normal data** as the signal received from a celestial region devoid of any sources but the background.
  - Conversely, **anomalous data** represents the signal emanating from the same celestial region in which an astrophysical source appears in addition to the background.

#### Use case:

- We **simulate** the data acquisition of a sub-array configuration of CTAO (4 LSTs in the North site, pointing at the zenith angle of °40 degrees (extra-galactic observations).
- We compute time series of flux measurements.

#### ➤ Model:

- Autoencoder architecture (CNN or RNN) trained with unsupervised learning.
- Trained to reconstruct the multivariate time series of flux points of normal data.
- We do not input the whole time series to the model, but we extract sub-windows.

### Inference

- > During inference, the model's reconstruction error is used as the anomaly score.
- The anomaly score serves as the test statistic for the p-value analysis, where the null hypothesis represents the absence of a GRB event in the data.



## Testing

- > The GRB templates used in the simulations come from the **POSyTIVE catalog** (M.G Bernardini, 2019).
  - Peak flux criteria (selecting 419 templates).
  - Simulating 500s of observation. GRB starts at tonset=250s. Only the GRB afterglow model.
  - Short-term analysis scenario: **1 sec integration time**.
  - Flux sub-windows of 5 points.
- We compute significance values for each subsequence within the light curves: a detection was considered successful if the significance value reached a threshold of 5.



46

### Results

- We then compared the cumulative number of detections for both our method and Li&Ma's technique, considering only the earliest detection event for each light curve.
- > Our technique achieves a higher number of GRB detections.



Paper in preparation. PhD Thesis: Real-Time Anomaly Detection of Gamma-Ray Bursts for the Cherenkov Telescope Array using Deep Learning, University of Bologna, <u>http://amsdottorato.unibo.it/10991/</u>

## Conclusions

## **Conclusions and Future Works**



- We developed **Deep Learning and Quantum Deep Learning models to detect GRBs in the sky maps and time series** generated with the data acquired by the detectors onboard the AGILE space mission.
- The results obtained prove the **capability of neural networks to analyze high-energy astrophysical data,** and in the analyzed context, they outperform classical analysis methods.
- We developed a Deep Learning model to localize the GRBs using the COSI BGO and GeDs simulated data. We still have to evaluate the impact of the background noise.
- We developed several Deep Learning models to detect and localize GRBs in CTAO simulated data.
- Our goal is to use the knowledge acquired during the development of Deep Learning models for **AGILE** for the next generation of high-energy projects such as **CTAO and COSI**.

## Thank You!





