The Astrophysical Journal, 973:63 (10pp), 2024 September 20 https://doi.org/10.3847/1538-4357/ad64cd

A New Deep Learning Model to Detect Gamma-Ray Bursts in the AGILE Anticoincidence System

N. Parmiggiani[1], A. Bulgarelli[1], L. Castaldini[1], A. De Rosa[1], A. Di Piano[1,2], R. Falco[1], V. Fioretti[1], A. Macaluso[3], G. Panebianco[1,4], A. Ursi[5,6], C. Pittori[7,8], M. Tavani[6,9], and D. Beneventano[2]

[1] INAF/OAS Bologna, Via P. Gobetti 93/3, 40129 Bologna, Italy; nicolo.parmiggiani@inaf.it
[2] Università degli Studi di Modena e Reggio Emilia, DIEF, Via Pietro Vivarelli 10, 41125 Modena, Italy
[3] German Research Center for Artificial Intelligence (DFKI), 66123 Saarbruecken, Germany
[4] Department of Physics and Astronomy, University of Bologna, Via Gobetti 93/2, 40129, Bologna, Italy
[5] Italian Space Agency (ASI), Via del Politecnico snc, 00133 Roma, Italy
[6] INAF/IAPS Roma, via del Fosso del Cavaliere 100, I-00133 Roma, Italy
[7] INAF/OAR Roma, Via di Frascati 33, I-00078 Monte Porzio Catone, Italy
[8] ASI Space Science Data Center (SSDC), Via del Politecnico snc, 00133 Roma, Italy
[9] Universitá degli Studi di Roma "Tor Vergata", via della Ricerca Scientifica 1, I-00133 Roma, Italy

Received 2024 April 3; revised 2024 July 16; accepted 2024 July 16; published 2024 September 19

Abstract

The AGILE space mission was launched in 2007 to study X-ray and gamma-ray astrophysics. AGILE operated in spinning mode from 2009 until 2024 February 14, when it re-entered the Earth's atmosphere. This work uses data acquired from the AGILE anticoincidence system (ACS) from 2019 January 1 to 2022 December 31. The ACS is designed to reject charged background particles. It also detects X-ray photons in the 50–200 KeV energy range and saves each panel count rate in the telemetry as ratemeter data, a time series with a resolution of 1.024 s. We developed a method that uses a deep learning model to predict the background count rates of the AGILE ACS top panel (perpendicular to the pointing direction of the payload detectors) using the satellite's orbital parameters as input. Then, we use the difference between predicted and acquired count rates to detect gamma-ray bursts (GRB). We trained the model with a background-only data set. After the training, the model can predict the ACS count rates with a mean reconstruction error of 3.8%. We used the GRBs listed in the GRBweb catalog to search for significant anomalies in the ACS data. We extracted light curves of 140 bins of 1.024 s for each GRB from the AGILE ACS to cover the trigger time of the GRBs. The model detected 39 GRBs with a significance of $\sigma \geqslant 3$. The results contain four GRBs detected for the first time in the AGILE data.

Unified Astronomy Thesaurus concepts: Gamma-ray bursts (629); Convolutional neural networks (1938)

1. Introduction

The Astrorivelatore Gamma ad Immagini LEggero—Light Imager for Gamma-Ray Astrophysics (AGILE) is a space mission of the Italian Space Agency devoted to high-energy astrophysics, launched in 2007 (Tavani et al. 2008; Bulgarelli et al. 2010; Pittori 2019). After almost 17 yr of intense and successful operations, AGILE ceased observations on 2024 January 18 and the satellite re-entered the Earth's atmosphere on 2024 February 14, due to the natural decay of its low Earth orbit (Tavani et al. 2024). This study was carried out when AGILE was still operating nominally. The AGILE payload consists of the Silicon Tracker (ST), the SuperAGILE X-ray detector, the CsI Mini-Calorimeter (MCAL), and an anticoincidence system (ACS; Perotti et al. 2006). The ST, MCAL, and ACS work together to form the Gamma-Ray Imaging Detector (GRID).

The ACS (Figure 1) has five independent plastic scintillation panels (four lateral and one on top) surrounding the AGILE detectors. The ACS 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 ratemeter (RM) data, with 1.024 s resolution. Each ACS panel RM count rate constitutes a time series. This work analyzes the AGILE ACS data from 2019 January 1 to 2022 December 31. We selected this time window to evaluate the model with four years of data and prove that one single model can manage multiple years of data, considering the variations in the background rates due to the solar activity and the descending altitude. In addition, this period has consistent conditions, such as stable payload configurations, mostly uniform telemetry data acquisition flow from the ground station, and not including occasional onboard hardware issues that occurred during the long lifetime of the mission. Once the validity and potential of the method have been demonstrated, we plan to complete the analysis of the entire AGILE archive with an in-depth analysis that requires an ad hoc procedure to take the effect of the different observing conditions into account, including significant orbital decay. This in-depth analysis goes beyond the scope of this paper. We decided to analyze the ACS top panel data perpendicular to the pointing direction of the payload detectors because it is configured to have a wider energy range than other panels (Ursi et al. 2023). In addition, this panel is less influenced by the solar flares because it is perpendicular to the solar panels pointed toward the Sun.

We developed a new method based on a deep learning (DL) model to detect gamma-ray binaries (GRBs) in the AGILE ACS top panel RM. The method is composed of two steps. The first step, similar to the method presented in Crupi et al. (2023), predicts with a DL model the background count rates expected by the AGILE ACS top panel using the satellite's orbital

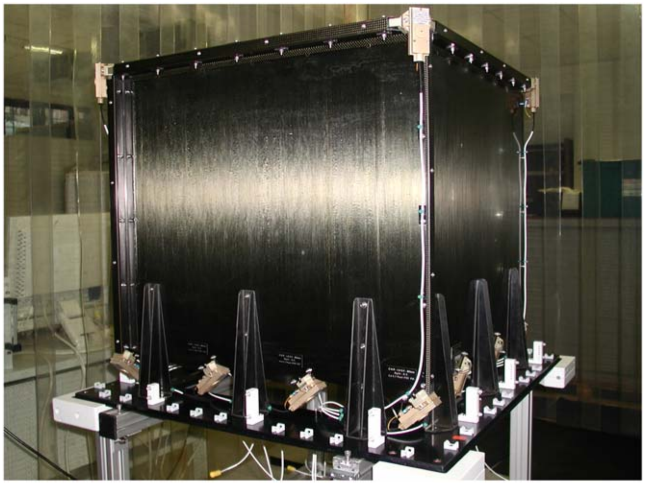

Figure 1. The AGILE Anticoincidence System. Credits: F. Perotti and the AGILE Team.

parameters. Then, the second step of the method detects anomalies using the anomaly score of the light curves with candidate GRBs. The anomaly score, described in detail in Section 4, is calculated using the differences between real and predicted count rates. We calculated the p-value distribution of the anomaly scores obtained on a background-only data set of light curves. The p-value distribution is used to determine the statistical significance of the detected GRBs. Indeed, if the anomaly score is higher than a predefined threshold, we can reject the null hypothesis of having it due to background-only emission and accept the hypothesis of a GRB. We decided to use a threshold of 3σ to classify a light curve as containing a GRB.

In the past, we developed a DL model to detect GRBs in the ACS data Parmiggiani et al. (2023) based on a convolutional neural network autoencoder that we used to perform an anomaly detection analysis. The method presented here improves the number of GRBs detected (seven additional GRBs in the time window overlapping with the previous method). In addition, it has a crucial advantage because it can analyze the raw data acquired by the ACS without needing a detrending algorithm (required by the previous method) that can add artificial anomalies.

Section 2 presents an analysis of the ACS top panel light curves at different timescales. In Section 3, we describe how we defined the data set to train the DL model, its architecture, and the training process. Section 4 describes the method used to detect the GRBs and calculate the statistical significance of the detections. Section 5 reports how we evaluated the model and the results obtained. The conclusions and future works are presented in Section 6.

2. AGILE Orbital and Attitude Patterns

The AGILE orbit was equatorial with an inclination of $2\overset{\circ}{.}5$. After a reaction wheel failure in 2009, AGILE scientific operations were reconfigured and the satellite started operating in a controlled spinning observing mode, around its Sun-pointing axis, with an angular speed of about $0\overset{\circ}{.}8\ \mathrm{s}^{-1}$. This work only uses data acquired during the spinning mode period. Figure 2 shows three days of ACS top data acquired during 2020, while Figure 3 shows the data acquired during two orbits in the same year. The RMs are a time series interrupted at each orbit during the passage over the South Atlantic Anomaly (SAA). These figures show the trends involved during more orbits and a single orbit. In particular, the trends with more impact are: (i) the daily trend, (ii) the orbital trend with a period of $\sim$94 min^{-1}, and (iii) the spinning trend.

Due to the complex AGILE observing pattern, the ACS top RMs show several modulations. We aim to train a DL model that can learn how this observing pattern influences the background and the count rates of the ACS top panel.

3. Deep Learning Model

In recent years, DL models (LeCun et al. 2015) raised popularity thanks to the improvement of computational hardware (e.g., graphical processing unit—GPU), the availability of big data that can be used to train large DL models, and the availability of multiplatform frameworks that are used to implement and train deep neural networks. In the high-energy astrophysical domain, there are several examples of the usage of these techniques. In Zhang et al. (2024), the authors exploit the capabilities of DL to detect GRBs in the Fermi Gamma-ray Burst Monitor (GBM; Meegan et al. 2009) data. I. Sadeh (2020), a new framework based on DL is presented to detect explosive astrophysical transients in the multimessenger context. In Finke et al. (2021), the authors present a new approach based on DL to classify unidentified or unassociated gamma-ray sources in the Fermi Large Area Telescope (LAT) catalog (Abdollahi et al. 2020; Ballet et al. 2020). Many works, such as Shilon et al. (2019) and Jacquemont et al. (2021), use DL models to analyze the Imaging Atmospheric Cherenkov Telescopes (Krennrich 2009) data.

For the AGILE space mission, we already developed two DL models. The first one (Parmiggiani et al. 2021) is used to detect GRBs in the intensity maps of the AGILE/GRID instruments. The second one implements an anomaly detection model to detect GRBs in the ACS data (Parmiggiani et al. 2023) and is the starting point for the work presented in this manuscript. The method presented here aims to detect the GRBs from the ACS data but with a different DL technique, leading to several improvements as described in the next sections.

3.1. Data Set Creation

The main difference with the method presented in Parmiggiani et al. (2023) is that the current method uses as input of the DL model the AGILE orbital parameters' configuration (e.g., detector pointing, altitude), and the output is the predicted background count rate of the ACS top panel. The model performs a regression task and is trained with a supervised technique. The previous DL model was instead trained only on the ACS data using the autoencoder approach.

We generated a labeled data set to train, validate, and test the model. The data set contains more than 80×10^6 orbital parameters' configurations extracted in the four-year time window (defined in Section 1) from the AGILE data archive. Each orbital parameter configuration is associated with the ACS top panel count rate acquired during the same time interval (the labels). The ACS top panel has a resolution of 1.024 s, and we can extract from the AGILE data archive the AGILE orbital parameters for each bin of the ACS top panel light curve. The AGILE orbital configuration used to train the model contains 17 values:

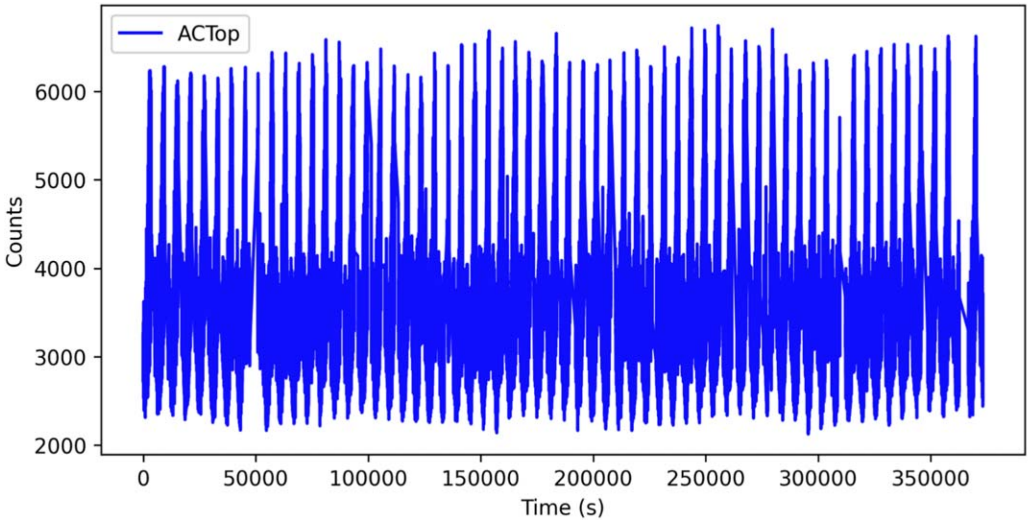

Figure 2. Raw data extracted from the ACS top data, corresponding to a three-day time interval, acquired in 2020 with a 1.024 s time resolution.

1. Earth latitude and longitude: the AGILE position projected on the Earth's surface.
2. AGILE attitude: the pointing directions of the three axes of the satellite expressed in R.A.–decl.
3. Earth R.A.–decl.: the position of the Earth.
4. Sun R.A.–decl.: position of the Sun.
5. Distance between the ACS top pointing direction with respect to the satellite–Earth vector: this value is calculated using the R.A.–decl. coordinates of the Earth and the ACS top panel pointing direction.
6. Distance between satellite–Earth vector and satellite–Sun vector: This value is calculated using the R.A.–decl. positions of the Sun and the Earth. The model uses it to know when the Earth covers the Sun.
7. Seconds from the start of the day: This value starts from 0 and reaches 86,399. This parameter is used to learn trends within the day.
8. Day of the year: This value ranges from 1 to 366. The model uses it to learn trends during one year of data.
9. Year: This value indicates the year in which the data have been acquired. Indeed, we can have different background values that change over the years for several reasons, such as solar activity.

We excluded from the data set the time windows containing the SAA passages and known GRBs to train the model with background-only data. The list of known GRBs is taken from the public GRBweb catalog[10] (Coppin et al. 2020), which collects GRBs detected from several facilities. We also considered the GRBs detected by other AGILE instruments.

We also removed the time windows containing bright solar flares (class M or X) from the data set. We decided to remove solar flares of class M and X because only very bright solar flares impact the ACS top panel. The detection of solar flares is out of the scope of this work. They have very different temporal behavior than GRBs, so a model trained to detect GRBs cannot be efficiently used to detect solar flares. In addition, we are analyzing the data of the ACS top panel, while AGILE solar flares are detected by the ACS lateral panel (LAT4) that is constrained to always point toward the Sun. Our team plans to develop a DL model dedicated to solar flare detection using LAT4 panel data in the future. The information to remove these time windows is extracted from the first AGILE solar flare catalog (Ursi et al. 2023), which collects solar flares detected with the AGILE ACS fourth panel that points toward the Sun. We removed 89 time windows related to solar flares. The AGILE solar flare catalog coverage ends on 2022 August 31. We have also manually removed the time windows for three bright solar flares that occurred in the last months of 2022.

Finally, we excluded from the data set the bins with count rates that exceed a threshold calculated from the already filtered data set using the percentile value greater than the 4σ probability. These outliers usually are cosmic rays that generate a peak in a single bin and must be removed to train the model with background-only data.

After applying these filters, we extracted 4×10^7 samples from the data set obtained to create the p-value data set used in Section 4 for the p-value analysis.

We split the remaining data set into training, test, and validation data sets with respective percentages of 80%, 10%, and 10%. Thus, the training data set contains 3.7×10^7 samples, while the test and validation data sets contain both 4.7×10^6 samples. The orbital parameters and the ACS count rates are normalized between 0 and 1 to improve the training process.

3.2. Deep Learning Architecture

We implemented the DL model to perform a regression task. The model is implemented with a dense neural network (Goodfellow et al. 2016). To find the best hyperparameter configuration, we evaluated 60 different configurations using the validation data set and by combining different numbers of layers, neurons, and dropout levels. The final model has one input layer that receives the 17 configuration parameters in input. The model has three hidden layers with 512, 256, and 128 neurons, respectively. The three hidden layers use the ReLU activation function. The output layer has one neuron and uses a linear activation function to provide a continuous value as output. We used dropout layers with a dropout rate of 1% as a regularization mechanism. The model is implemented using two open-source frameworks: Keras[11] and Tensorflow.[12]

---

[10] https://heasarc.gsfc.nasa.gov/FTP/fermi/data/gbm/bursts/

[11] https://keras.io/

[12] https://www.tensorflow.org/

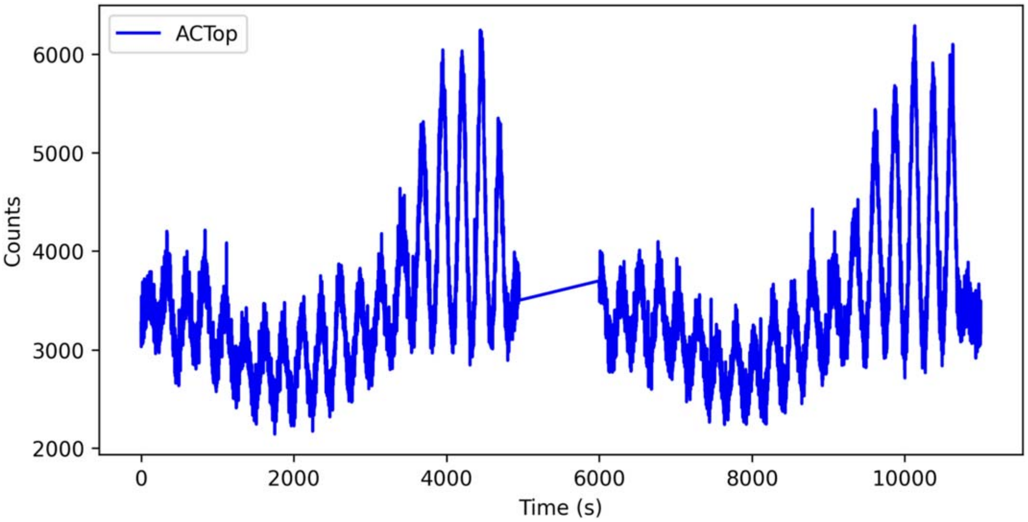

Figure 3. Raw data extracted from the ACS top data acquired during two orbits in 2020 with a 1.024 s time resolution. The data interruption in the middle of the light curve corresponds to the satellite passages into the SAA, where all detectors are put in idle mode.

3.3. Model Training

We trained the model using the Adam optimizer (Kingma & Ba 2015) with a variable learning rate: 0.005 for the first two epochs and 0.001 for other epochs. We used the mean absolute error (MAE) loss function defined as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|, \quad (1)$$

where n is the number of elements evaluated, y_i is the count rate observed, and $\hat{y}_i$ is the count rate predicted by the model.

We trained the model with a batch size of 131,072 samples using an NVIDIA Tesla V100 GPU on an IBM Power 9 processor. We selected a batch size with a power of 2 because this optimizes memory usage and computation efficiency. During the training, the model is validated using the validation data set described in Section 3.1.

Figure 4 shows the training results. The model was trained for 10 epochs, and then we decided to stop the training because the model converged, and we avoided the risk of overfitting. The validation loss is lower than the training loss; this effect can be caused by several reasons, such as the dropout layers for regularization used only during the training and not during the validation. However, the two curves converge, and there is no evidence of overfitting or underfitting.

We evaluated the trained model using the test data set, which contains more than 4.7×10^6 orbital parameters' configurations, and we compared the predicted ACS top panel count rates with the real ones. After the training, the MAE value calculated on the test data set (after applying the inverse normalization to have count units) is 136. We can calculate the reconstruction error percentage as the ratio between the MAE and the mean count rate value. The mean reconstruction error ratio is 3.8%.

Figure 5 shows an example of two ACS top light curves with the acquired and predicted values. The blue light curves are acquired, while the model predicts red light curves. The bottom panels show the differences between acquired and predicted data. The left plots contain the data of a single AGILE orbit, while the right plots show the data of two orbits.

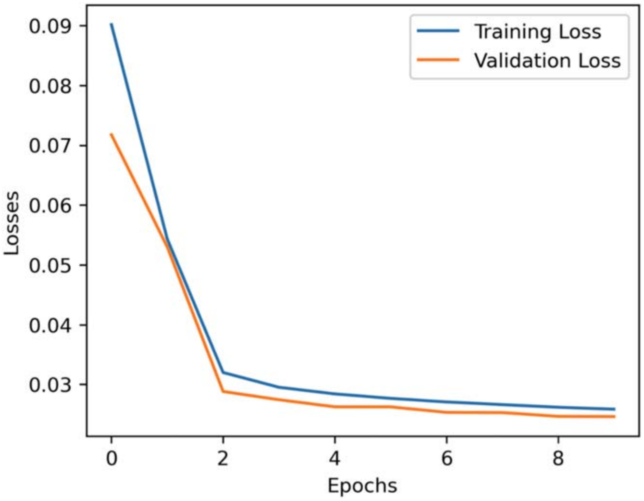

Figure 4. Training loss and validation loss during the 10 training epochs.

The model learned the patterns in the data, such as the orbital and spinning modulations. Indeed, it can accurately predict the background count rates of the ACS top panel in each orbital and spinning phase.

Figure 6 shows the histogram of the differences between acquired and predicted count rates in the test data set. The mean is 1.3, and the standard deviation is 171. The mean is close to zero, indicating that the deviation between real and predicted count rates is due to background fluctuations.

4. GRB Detection Method

The trained DL model is used to detect GRBs in the ACS top panel data. We can extract light curves, in the same period used for the training, with 140 bins of 1.024 s length from the ACS data archive and then evaluate all the bins in the light curves with the DL model. The method used to extract the light curves is described in Parmiggiani et al. (2023) but in this work, we avoided the detrending procedure by using directly the count rates acquired by the ACS top panel. We decided to evaluate the GRBs acquired during the same period used to train the model because the conditions of the AGILE orbit change over

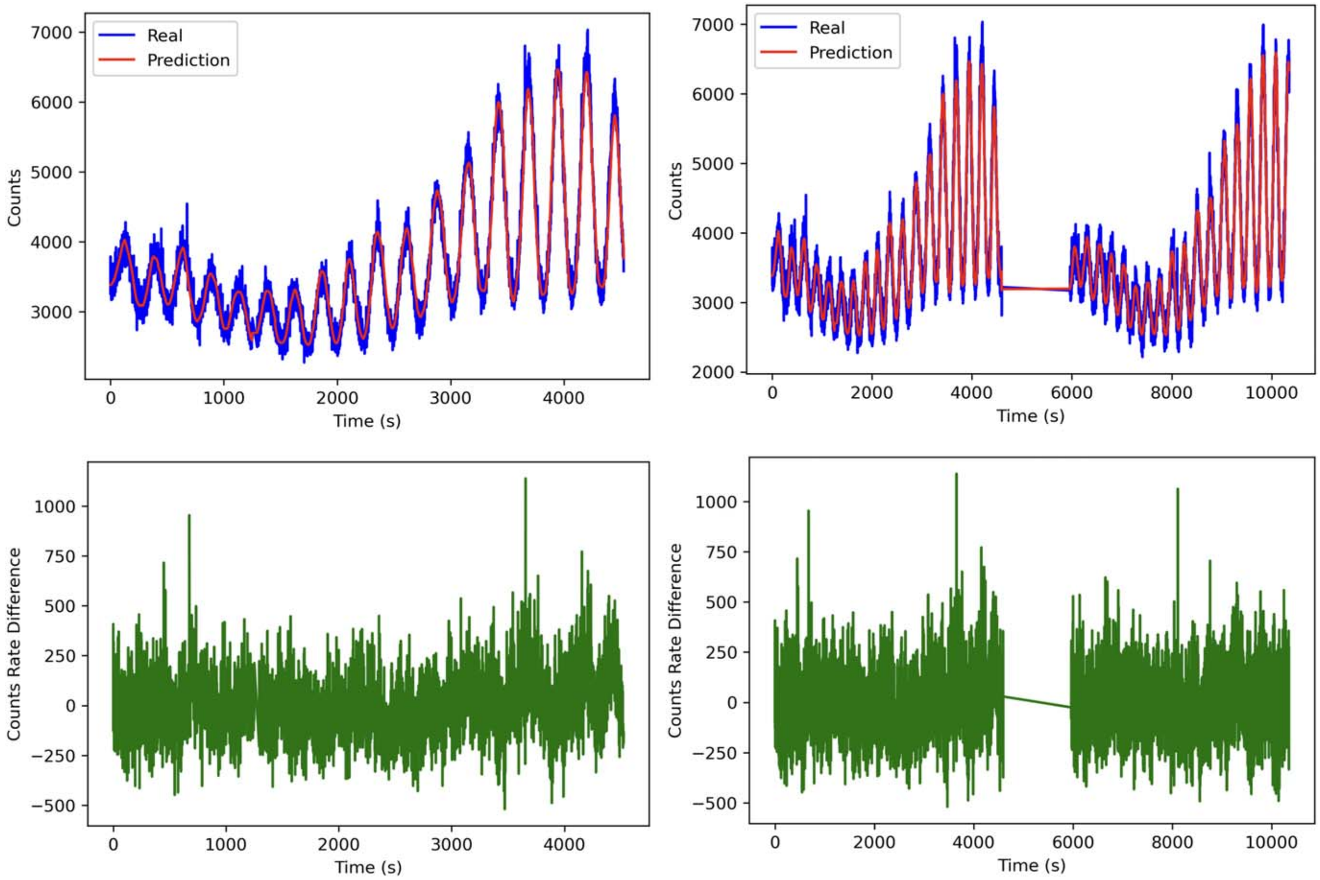

Figure 5. Examples of ACS top panel acquired count rates (blue) and predicted count rates (red). The difference between the two light curves is shown in the bottom panels (green). The left panels show the data acquired during one orbit, while the right panels show the data acquired during two orbits. Both images are generated with data acquired on 2019 January 11.

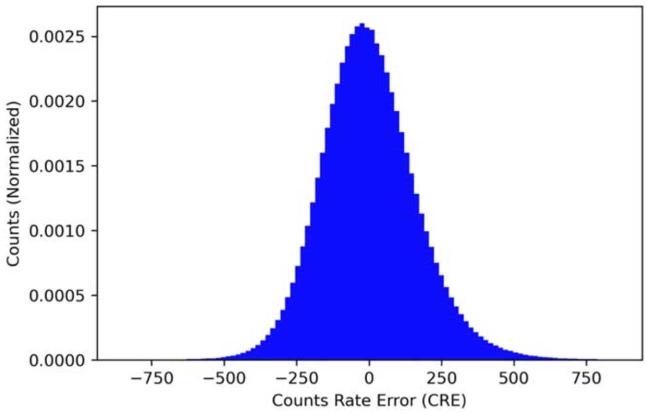

Figure 6. Histogram of the differences between acquired and real count rates in the test data set.

the years and to have a high precision in the prediction of the background count rate we need the same environmental conditions.

In order to use the DL model to detect GRBs, we must be able to calculate the statistical significance of detections. We decided to calculate the *p*-value distribution on the anomaly score (AS) of each light curve to define the statistical significance of the GRBs detected in the light curves. The AS of a light curve is calculated as the mean count rate differences of the light curves using only the count rate differences that exceed the mean value of the differences of the light curve plus 2 times the standard deviation. In this way, if there is a GRB, we increase the impact of the bins with high differences that are related to the transient event.

To calculate the *p*-value distribution we evaluated the *p*-value data set (defined in Section 3.1), which contains 4×10^7 orbital parameter configurations acquired during background-only time windows. To obtain a threshold for a 5σ detection we have to determine the *p*-value distribution using 2×10^7 light curves but the ACS archive does not contain enough data. For this reason, we used a bootstrapping data augmentation technique, described in Parmiggiani et al. (2023). We evaluated the *p*-value data set using the trained model to obtain the array of differences between acquired and predicted count rates. Then, we randomly sampled from this array, a data set of 2×10^7 light curves with 140 bins. We applied the bootstrap technique on the array of differences to optimize the computation of these values that is performed only once. We can use this method because the *p*-value data set contains background-only data and the order of the bin is not important. In addition, this random sampling procedure reduces the bias that can be present if we evaluate a light curve of continued data due to external factors such as solar flares.

Given φ, the distribution of the ASs obtained evaluating the *p*-value data set with the trained model, the probability that a

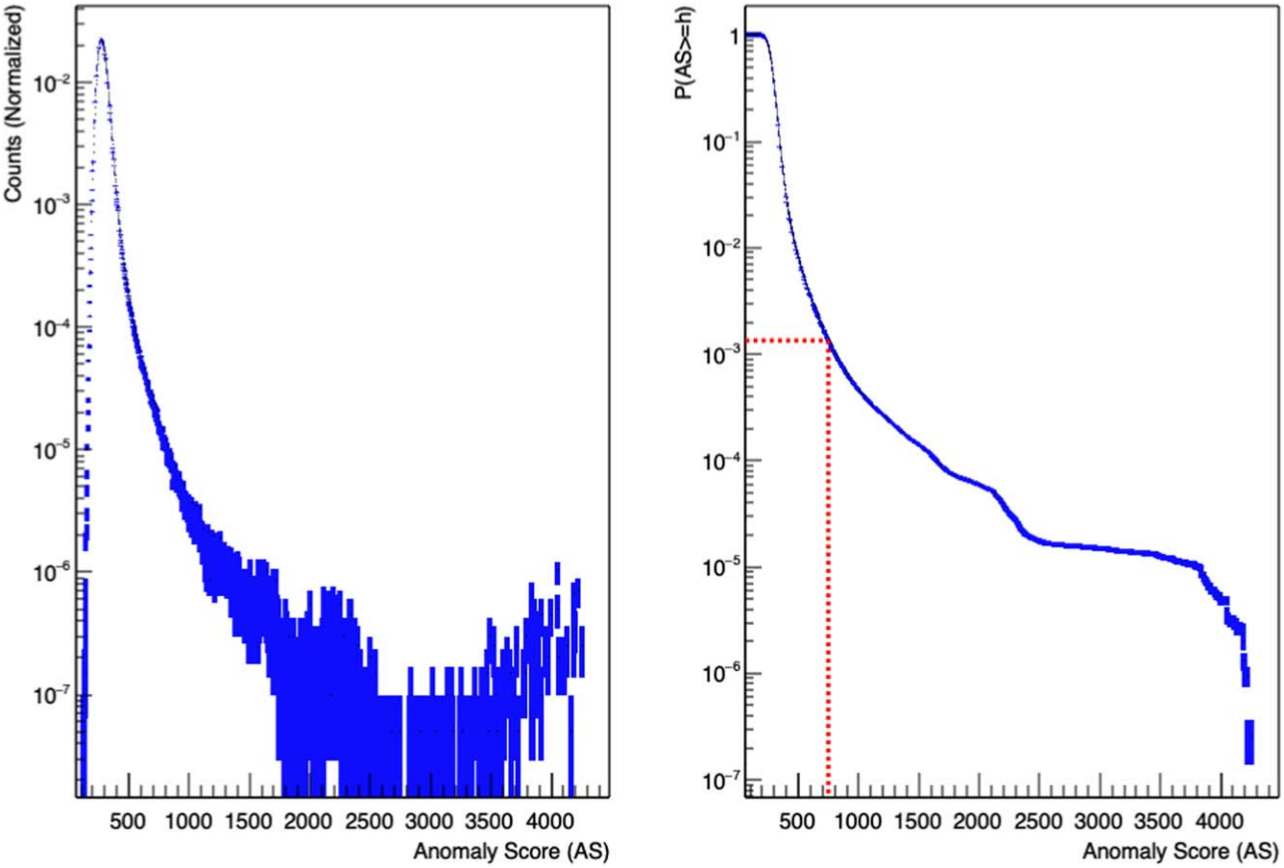

Figure 7. Distribution of the AS obtained with the trained model and the p-value data set (left panel). The p-value distribution was used to determine the thresholds at different sigma levels on the AS (right panel). The red dotted line shows the threshold for 3σ.

light curve of 140 bins of 1.024 s in a background-only time window has $AS \geqslant h$ (the cumulative distribution function complement) is:

$$P(AS \geqslant h) = \int_{h}^{+\infty} \varphi(x)dx, \tag{2}$$

which is also called p-value $p = P(AS \geqslant h)$ and defines the probability of obtaining that AS value or greater when the null hypothesis is true. We can reject the null hypothesis to state that we have a GRB signal with a probability p to have a false positive.

This p-value is used to calculate the corresponding number of standard deviations in a one-tailed test of a Gaussian (normal) variate. This value is the Z-value, called S or sigma, in high-energy astrophysics.

We can evaluate the p-value thresholds to calculate the statistical significance (S) with the following equation:

$$\sigma = \varphi^{-1}(1 - p), \tag{3}$$

where p is the p-value, and Φ is the standard normal cumulative distribution function. We classify as GRBs the light curves with $\sigma \geqslant 3$.

Figure 7 shows the normalized distribution of the AS and the associated p-value distribution. The plot shows the 3σ threshold on the AS value with a dotted red line.

Table 1 shows the thresholds calculated at different sigma levels.

Table 1
Relation between the σ Levels and Thresholds on the AS of the Light Curves

Nσ	p-value	AS Threshold
3	1.35×10^{-3}	747.9
3.5	2.32×10^{-4}	1269.9
4	3.17×10^{-5}	2250.5
4.5	3.40×10^{-6}	4047.8
5	2.86×10^{-7}	4220.4

5. DL Model Evaluation

We evaluated the trained model using the GRBs of the GRBweb catalog, described in Section 3.1. The time window of this analysis is the same as the time window widows used to extract the training data set, from 2019 January 1 to 2022 December 31. We analyzed four years of data covering two years not covered by the already presented analysis of ACS data (Ursi et al. 2022; Parmiggiani et al. 2023).

5.1. Gamma-Ray Bursts Data Set

We generated the candidate GRB data set by extracting a light curve of 140 bins of 1.024 s from the ACS top data for each GRB selected from the GRBweb catalog. The light curves start 20 s before the GRB trigger time including a background time window before the GRBs because it is possible to have some difference in the trigger times from AGILE and other space missions due to different energy ranges or visibility

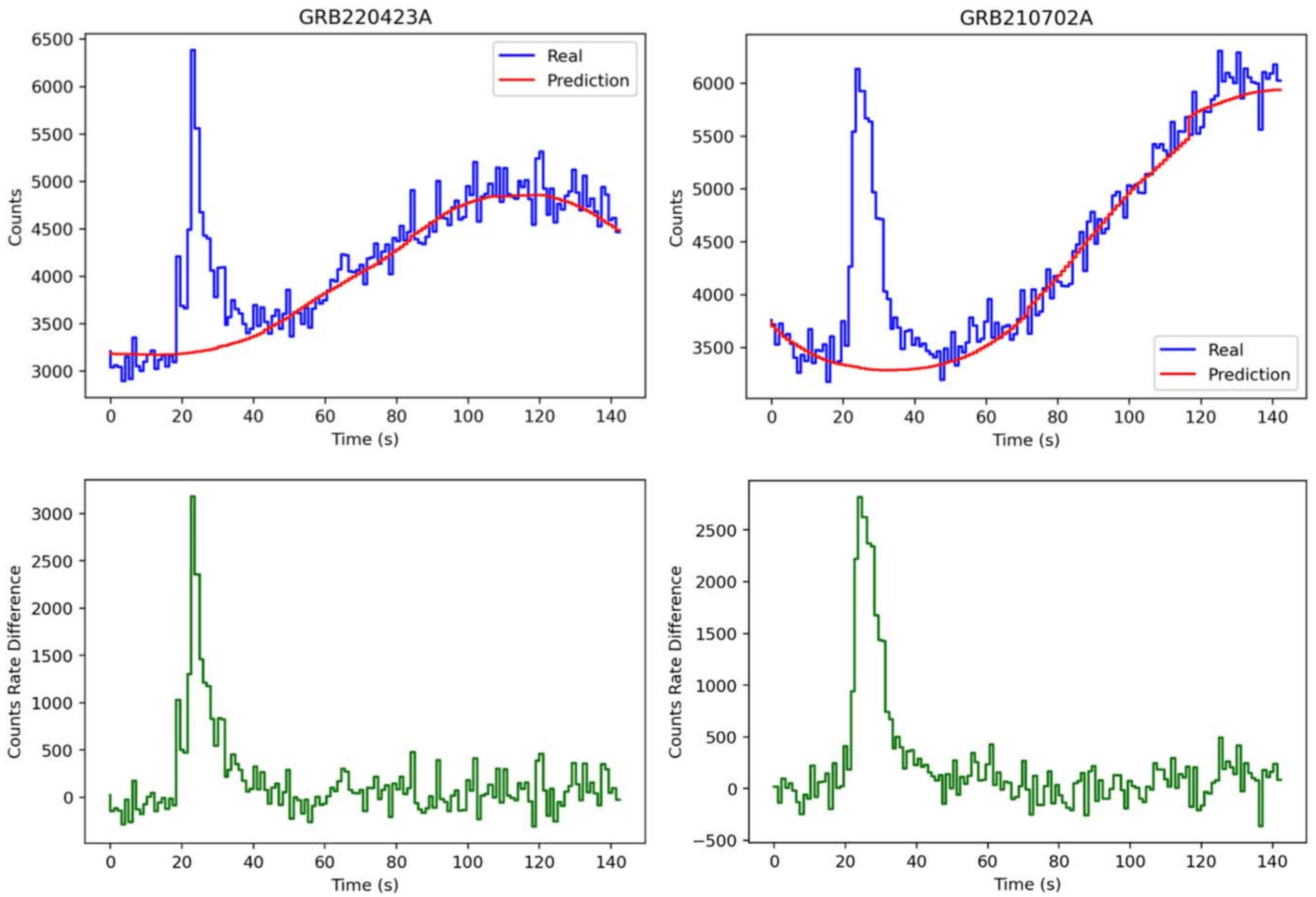

Figure 8. Example of light curves related to GRB220423A and GRB210702A. The top panels show the acquired count rates (blue) and the predicted ones (red). The bottom panels show the difference between the acquired and predicted count rates.

constraints. Not all GRBs have coverage in the AGILE data because the satellite can be in the SAA region or due to other issues during the data acquisition. In addition, we excluded from the list the GRBs that fall inside the time windows of a solar flare. We extracted 760 light curves related to the GRBs presented in the catalog with ACS top data coverage.

5.2. *Results*

We evaluated the GRB data set with the DL model and calculated the AS of each light curve. The calculated AS is compared with the thresholds defined in Table 1 to check if there is a detection and its sigma level. Figure 8 shows two light curves where the DL model detected two GRBs. The top panel shows the acquired ACS top count rates (blue) and the predicted ones (red) related to GRB220423A and GRB210702A. The difference between the acquired and predicted count rates is shown in the bottom panels. The peaks in the count rates highlight the presence of GRBs. The presence of GRBs yields an increase in the differences between the acquired and predicted count rates.

The model detected 41 anomalies with a $\sigma \geqslant 3$. We evaluated them manually and removed two anomalies unrelated to GRBs. The first removed anomaly is caused because the model underestimate the background for part of the light curve, so the detection algorithm calculated a high anomaly score associated with a 3σ detection. The second anomaly is related to a burst with a duration lower than 1 s but distant from the trigger time of the GRB. As a future improvement, we could develop an additional algorithm to classify anomalies and discard those unrelated to GRBs.

The remaining 39 anomalies are associated with a GRB. Table 2 shows the detected GRBs with the associated σ level. In the period from 2019 to 2022, the model detected 22 GRBs. From this list, 15 have already been detected by Parmiggiani et al. (2023) while 7 are new detections with DL models. Two of these seven GRBs have a GCN circular published by the AGILE Team, while five have no GCN circulars. For this reason, this is the first detection made with the ACS system of 5 GRBs. These results prove that the new DL model presented here improves the detection capabilities of GRBs in the AGILE ACS data.

From the 22 GRBs detected by this new model in the same period covered by Parmiggiani et al. (2023) and Ursi et al. (2022), 16 have been detected by MCAL and are present in Ursi et al. (2022) while 6 GRBs are not present in the second MCAL GRB catalog.

Considering both Parmiggiani et al. (2023) and Ursi et al. (2022), this new method detected five new GRBs. Only one GRB from this list has a GCN circular published by the AGILE Team. The AGILE team has never detected the remaining four GRBs before with the ACS or other instruments. These GRBs are GRB190129A, GRB190415A, GRB190505A, and GRB191111B.

During the time period not covered by Parmiggiani et al. (2023) and Ursi et al. (2022), from 2021 to 2022, the new

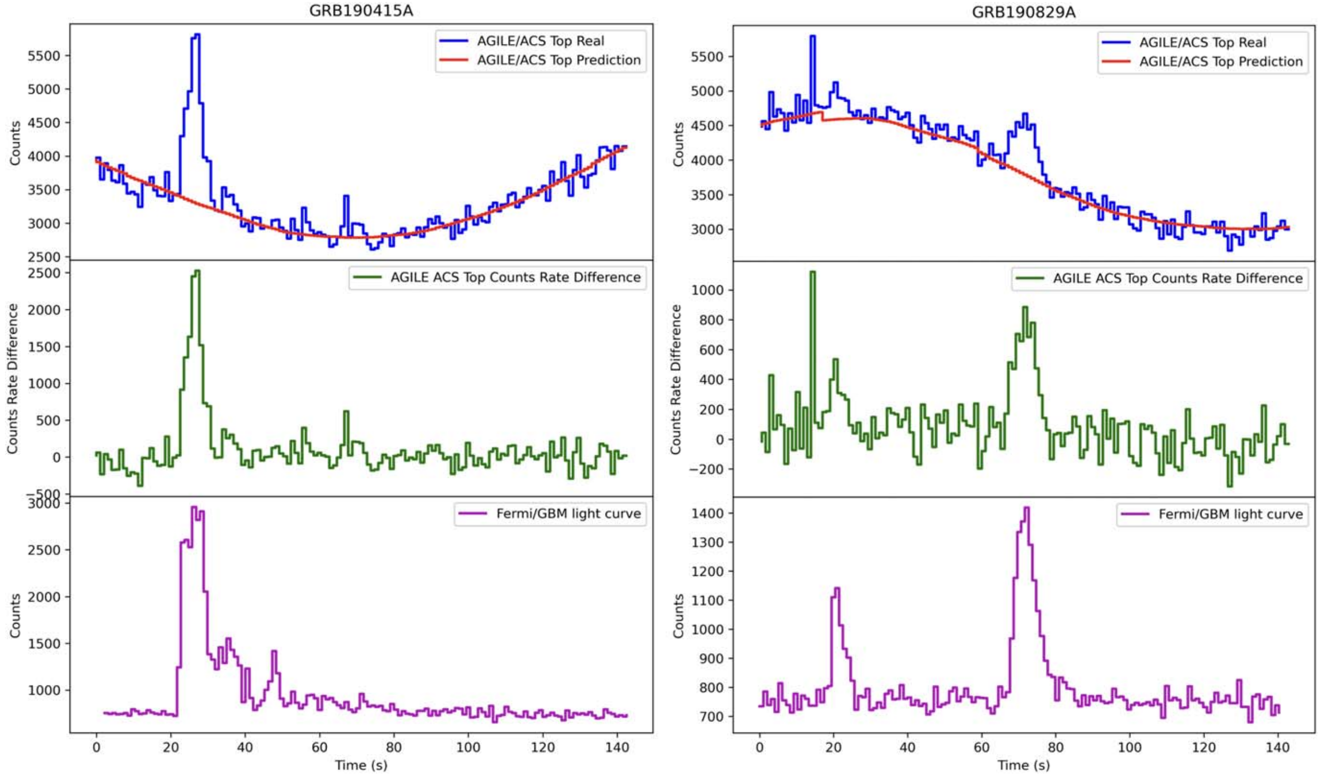

Figure 9. Comparison between light curves of the GRB190415A and GRB190829A acquired by the AGILE/ACS top panel (50–200 keV) and the Fermi/GBM (50–300 keV). The trigger time of the GRBs is 20 s from the start of the light curves.

Table 2
GRB Detected with $\sigma \geqslant 3$

GRB	$N\sigma$	GRB	$N\sigma$	GRB	$N\sigma$	GRB	$N\sigma$
190103A	3.5	190727B	3.5	200919C	3.5	211120A	3
190114C	5	190731A	4	201020B	3.5	211226A	3
190129A*	3	190829A	3	210410A	3	220209A	3
190329A	3.5	190928A	5	210511B	3	220308A	5
190415A*	3.5	191111B*	3	210606B	3.5	220423A	3.5
190501A	4	191221B	4	210610B	3	220623A	3
190505A*	3	200131A	3	210702A	3.5	220624A	4
190604A	3	200521A	3.5	210704A	4	220910A	4
190611B	3	200829A	5	210818A	4	221023A	5
190706C	3	200903E	4	211023A	3	⋯	⋯

Note. List of the 39 GRBs detected with a significance $\sigma \geqslant 3$ by using our DL model to analyze the list of GRBs obtained from the GRBweb catalog. The GRBs marked with * are GRBs detected in the AGILE data for the first time.

method detected 17 GRBs. All these GRBs are confirmed by GCN circulars published by the AGILE Team. A dedicated MCAL analysis can be executed to verify if there are detections in the MCAL data, but this analysis is out of the scope of this work.

With this new method, we detected four new GRBs never detected with ACS data and without a related GCN communication. For this reason, we can consider them as new AGILE detections.

We compared the results obtained with the AGILE/MCAL detector, bearing in mind, however, that we cannot expect to obtain the same number of detections with ACS data due to differences in the energy ranges and sensitivity of the instruments. As described in Section 5.1 of Parmiggiani et al. (2023), the ACS data can be considered complementary to other AGILE instruments. In particular, the results that we obtain with the ACS panels can add information about the GRBs detected by the AGILE/MCAL instrument due to the different time resolutions and even allow us to detect new GRBs.

In Figure 9, we compare the light curves acquired by the AGILE ACS top panel and the light curve acquired by the Fermi/GBM (Goldstein et al. 2022) for the GRB190415A and the GRB190829A. The previous DL model presented in Parmiggiani et al. (2023) has not detected both GRBs. The two light curves show the same peaks and temporal patterns that clearly correspond to the same transient event.

6. Conclusions

This manuscript presents a new method to detect GRBs using the AGILE ACS top panel data. This method is based on DL technologies to predict the count rates of the AGILE ACS and uses the difference between the acquired and predicted count rates to detect GRBs. The main advantage of this method with respect to the previous DL model used to detect GRBs in the ACS data is that it can be applied directly to the raw data without needing a detrending algorithm that can introduce artificial anomalies.

We trained the DL model with a data set of more than 3.7×10^7 satellite parameter configurations associated with the count rates of the ACS top panel acquired simultaneously. Each satellite configuration contains several parameters, such as the satellite attitudes of the three axes or the longitude and latitude of AGILE on the Earth's surface. The trained model predicts the ACS top count rates with a mean error value of 3.8% and can learn the complex observing pattern of the AGILE satellite.

In order to detect the GRBs with an associated statistical significance, we calculated the *p*-value distribution of the anomaly scores calculated with the data predicted by the DL model.

We used a list of 760 GRBs presented by the GRBweb catalog to test the model. The DL model detected 39 GRBs with a $\sigma \geqslant 3$. Considering the time window that overlaps with previous works, this method detects four additional GRBs not detected before by AGILE. These GRBs are not present neither in the Second AGILE/MCAL GRB catalog nor in the previous DL model developed to analyze the AGILE ACS data. These results confirm that this new model improves the GRB detection capability on the ACS data with respect to the previous DL model. From the time window that is not covered by previous analyses, the DL model detects 17 GRBs. These detections are confirmed by the GCN circulars published by the AGILE Team.

We plan to use this method to analyze the data of other AGILE instruments. We think that the presented method can also be used for other types of analysis, such as detecting different transient events or studying the background levels in the satellite's instruments. Our work can be a starting point for future research for the AGILE space mission but also for the next generation of high-energy facilities.

Acknowledgments

The AGILE Mission is funded by the Italian Space Agency (ASI) with scientific and programmatic participation by the Italian National Institute for Astrophysics (INAF) and the Italian National Institute for Nuclear Physics (INFN). The investigation is supported by the ASI grant I/028/12.7-2022 and also partially funded with an INAF "Mini-Grant" 2022. We thank the ASI management for unfailing support during AGILE operations. We acknowledge the effort of ASI and industry personnel in operating the ASI ground station in Malindi (Kenya), the Telespazio Mission Control Center at Fucino, and the data processing done at the ASI/SSDC in Rome: the success of AGILE scientific operations depends on the effectiveness of the data flow from Kenya to SSDC and the data analysis and software management. We thank IBM Italy and Var Group S.p.A. for the opportunity to use a Power 9 with NVIDIA Tesla V100.

Appendix A
Training and Inference Workflows

Figure 10 presents the workflow of the training and inference phases. The deep learning model is trained on a data set obtained by extracting orbital and attitude parameters from the AGILE archive of the ACS top data. The trained model can then be used for the inference phase to detect GRBs, starting with the extraction of the orbital and attitude configurations of the desired time window from the AGILE data archive. From these input configurations, the model outputs the light curve of expected count rates for the desired time window. We then can extract the acquired light curve for the same time interval from the AGILE data archive and calculate the difference between the two, which is used by the anomaly detection algorithm (described in Section 4) to evaluate the presence of a candidate GRB.

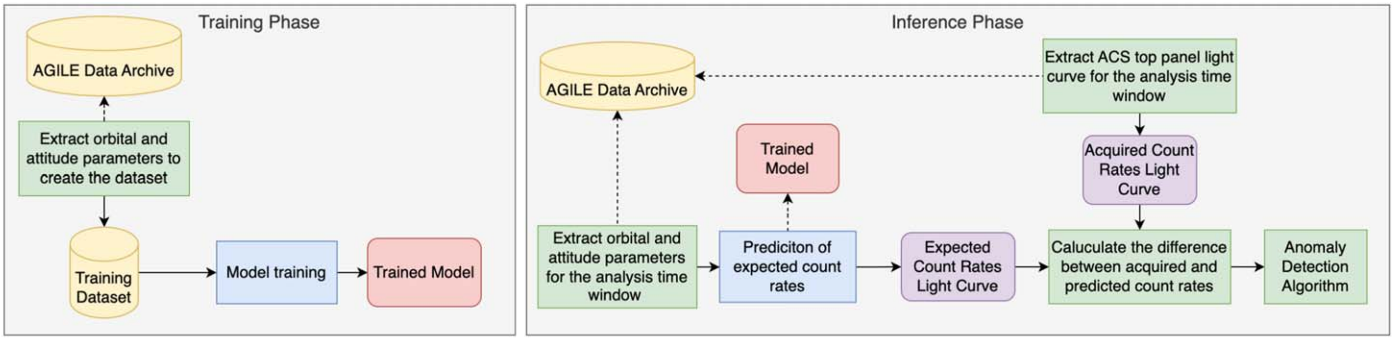

Figure 10. Workflows for the training and interference phase. The former is used to obtain a trained model while in the latter we use the trained model to predict the expected count rates for the ACS top panel in a time window.

Appendix B
Deep Learning Glossary

In this section, we describe the terminology used for DL projects.

1. Activation function: This function calculates the output of the nodes in a neural network using the inputs and weights associated with the nodes. There are several activation functions for different purposes. A neural network can have different activation functions used in different layers to obtain the desired behavior.
2. Deep learning: It is a subset of machine learning. The adjective "deep" is used because the model developed in this context has multiple layers that are able to extract features from the data set at different levels of abstraction during the training phase.
3. Dense neural network (DNN): a machine learning model (deep if it has more than two hidden layers) that is a subset of the more general neural network (NN) inspired by the structure and function of the human brain. The NN has several nodes (neurons) connected with other nodes (synapses). In the structure of the DNN, the nodes of each layer are connected with all nodes of the previous layer. This structure creates a large number of interconnections and parameters to be trained.
4. Dropout layer: This layer is used to prevent the overfitting of the model during training. The dropout layer randomly turns off certain neurons of the model during the training. This technique allows to improve the generalization of the model on unseen data.
5. Epoch and batch size: The training procedure completes an epoch when the entire training data set is used to train the model. The training lasts for several epochs but this number is highly variable. Usually for computation reasons the data set is divided into mini-batches and the training algorithm performs the update of the model weights using one mini-batch for each iteration. The batch size refers to the number of samples in the mini-batch. This parameter can be changed to optimize the training results.
6. Input, output, and hidden layers: The input layer consists of nodes of the NN that are used to represent the input data. Each node is related to a feature of the input data. The output layer produces the final results of the model that we expect in output. The hidden layers are the layers between the input and output layers. These nodes can have several structures and perform the analysis of the input data.
7. Hyperparameters: These parameters define the model structure and the behavior of the training algorithm. They are not updated during the training phase.
8. Machine learning inference: This process uses a trained model to calculate a numerical output based on the input data.
9. Learning rate: It is one of the hyperparameters and indicates the step size used during training for updating the model weight at each iteration.
10. Overfitting: The model is overfitted when it fits too closely with the training data and cannot generalize on data variances not present in the training data set. This behavior is the opposite of the purpose of DL models, which is being able to generalize the analysis on new data.
11. Rectified linear unit (ReLU) activation function: The ReLU activation function is largely used in DL models, mainly due to its efficient computation. It is defined as $f(x) = \max(0, x)$.
12. Training test and validation data sets: The training data set contains the data used during the training procedure. The test data set is only used to evaluate the trained model. The validation data set is used to evaluate the impact of different hyperparameters on the training data set. It is also used during the training procedure to evaluate at each iteration the model improvements. The idea behind the usage of the test and validation data set is to evaluate the model with data that are not used for its training. This evaluation checks the generalization capabilities of the model.

ORCID iDs

N. Parmiggiani https://orcid.org/0000-0002-4535-5329
A. Bulgarelli https://orcid.org/0000-0001-6347-0649
L. Castaldini https://orcid.org/0009-0000-5501-4328
A. De Rosa https://orcid.org/0009-0001-8367-9359
A. Di Piano https://orcid.org/0000-0002-9894-7491
R. Falco https://orcid.org/0009-0004-1676-7596
V. Fioretti https://orcid.org/0000-0002-6082-5384
A. Macaluso https://orcid.org/0000-0002-1348-250X
G. Panebianco https://orcid.org/0000-0002-3410-8613
A. Ursi https://orcid.org/0000-0002-7253-9721
C. Pittori https://orcid.org/0000-0001-6661-9779
M. Tavani https://orcid.org/0000-0003-2893-1459
D. Beneventano https://orcid.org/0000-0001-6616-1753

References

Abdollahi, S., Acero, F., Ackermann, M., et al. 2020, ApJS, 247, 33
Ballet, J., Burnett, T. H., Digel, S. W., & Lott, B. 2020, arXiv:2005.11208
Bulgarelli, A., Argan, A., Barbiellini, G., et al. 2010, NIMPA, 614, 213
Coppin, P., de Vries, K. D., & van Eijndhoven, N. 2020, PhRvD, 102, 103014
Crupi, R., Dilillo, G., Bissaldi, E., et al. 2023, ExA, 56, 421
Finke, T., Krämer, M., & Manconi, S. 2021, MNRAS, 507, 4061
Goldstein, A., Cleveland, W. H., & Kocevski, D. 2022, Fermi GBM Data Tools: v1.1.1, https://fermi.gsfc.nasa.gov/ssc/data/analysis/gbm
Goodfellow, I., Bengio, Y., & Courville, A. 2016, Deep Learning (Cambridge, MA: The MIT Press)
Jacquemont, M., Vuillaume, T., Benoit, A., et al. 2021, 2021 Int. Conf. Content-Based Multimedia Indexing (CBMI) (New York: IEEE), 1
Kingma, D. P., & Ba, J. 2015, in Int. Conf. Learning Representations (San Diego, CA: ICLR)
Krennrich, F. 2009, NJPh, 11, 115008
LeCun, Y., Bengio, Y., & Hinton, G. 2015, Natur, 521, 436
Meegan, C., Lichti, G., Bhat, P. N., et al. 2009, ApJ, 702, 791
Parmiggiani, N., Bulgarelli, A., Fioretti, V., et al. 2021, ApJ, 914, 67
Parmiggiani, N., Bulgarelli, A., Ursi, A., et al. 2023, ApJ, 945, 106
Perotti, F., Fiorini, M., Incorvaia, S., Mattaini, E., & Sant'Ambrogio, E. 2006, NIMPA, 556, 228
Pittori, C. 2019, RLSFN, 30, 217
Sadeh, I. 2020, ApJL, 894, L25
Shilon, I., Kraus, M., Büchele, M., et al. 2019, APh, 105, 44
Tavani, M., Addis, A., Argan, A., et al. 2024, ATel, 16450, 1
Tavani, M., Barbiellini, G., Argan, A., et al. 2008, NIMPA, 588, 52
Ursi, A., Parmiggiani, N., Messerotti, M., et al. 2023, ApJS, 267, 9
Ursi, A., Romani, M., Verrecchia, F., et al. 2022, ApJ, 925, 152
Zhang, P., Li, B., Gui, R., et al. 2024, ApJS, 272, 4