Rapid Gamma-Ray Burst Localization aboard the e-Astrogam Satellite using a 3D Convolutional Neural Network

Ruoxi Shang
University of California at Berkeley

Andreas Zoglauer
Space Sciences Laboratory & Berkeley Institute for Data Science, University of California at Berkeley

Abstract

The discovery of the association of a short gamma-ray burst (GRB) with gravitational waves from the neutron-star-neutron-star merger GW170817 gave rise to the requirement to localize GRBs in close to real time even with the most sensitive but also data-analysis-wise most complex gamma-ray space telescopes, the Compton telescopes. Here we report on the implementation and testing of a 3D convolution neural network trained to localize the origin of Compton-scattered gamma-rays from GRB's on the sky. The ultimate goal of this project is the implementation of a neural network for GRB localization in a FPGA aboard a satellite such as the envisioned e-Astrogam.

Background

The gravitational wave event GW170817 [1] was the dawn of multi-messenger astronomy: For the first time a gravitational wave originating from a neutron-star-neutron-star merger could be associated with an electro-magnetic outburst, a short gamma-ray burst from the same merger event [2].

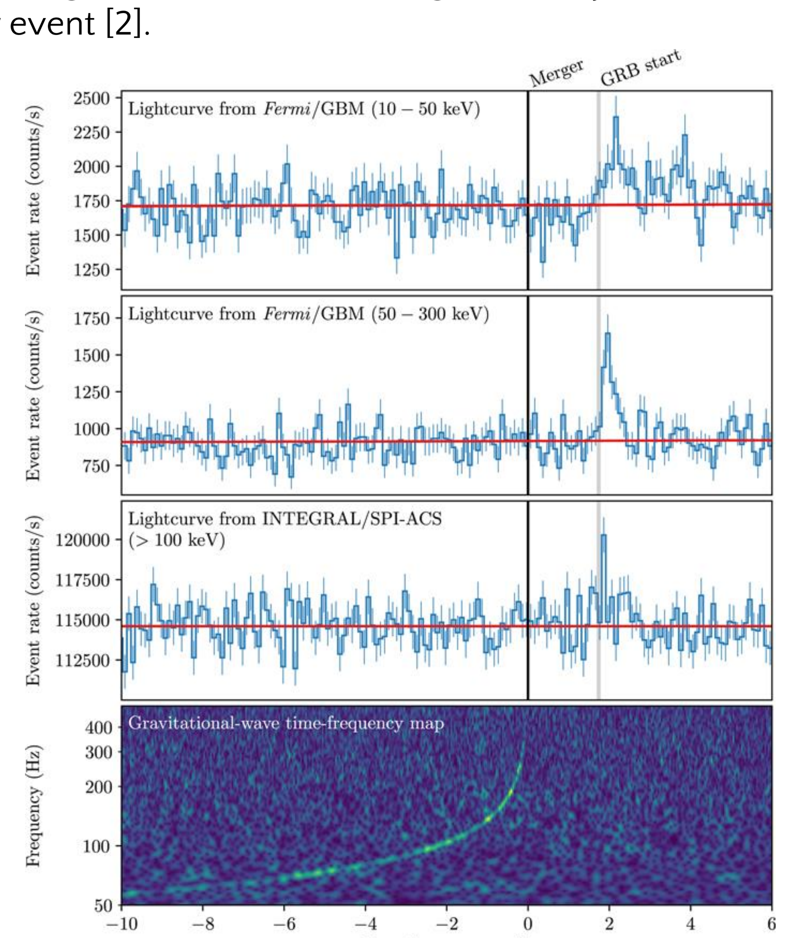

Fig. 1. The top 3 plots show the light curve of the gamma-ray burst as measured by the FERMI/GBM and INTEGRAL/SPI-ACS instruments, the lower plot shows the gravitational wave "chirp" as measured by LIGO (from [2])

In order to allow immediate follow-ups of these events on the ground, the superior localization capabilities of gamma-ray detectors (compared to gravitational wave detectors) needs to be fully utilized. The best way to achieve this goal is to perform the localization on board a satellite such as the envisioned e-ASTROGAM (picture from [3]):

Fig 2: Artist's concept of the envisioned e-Astrogam satellite

However, the data from the most sensitive type of these gamma-ray detectors, Compton telescopes, are inherently difficult to analyze within the constraints aboard a space satellite.

Methodology

A Compton telescope measures individual gamma rays as a sequence of hits (position and energy) in its detectors (see Figure 3). The data of each individual gamma ray can be reduced to 3 main components which are necessary to localize the origin of GRBs on the sky: The Compton scatter angle, φ, and the direction of the scattered gamma ray in Galactic coordinates, ξ and ψ (for details see e.g. [4]).

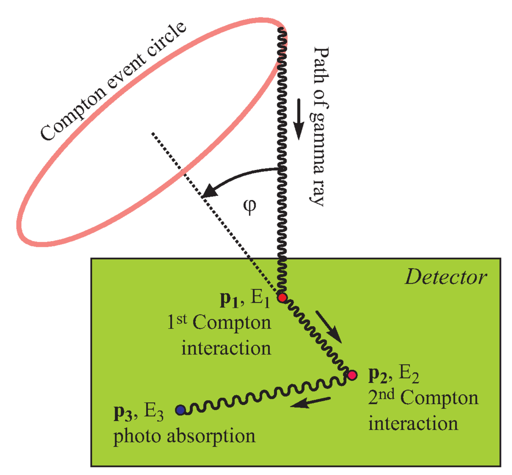

Fig. 3: Measurement principle

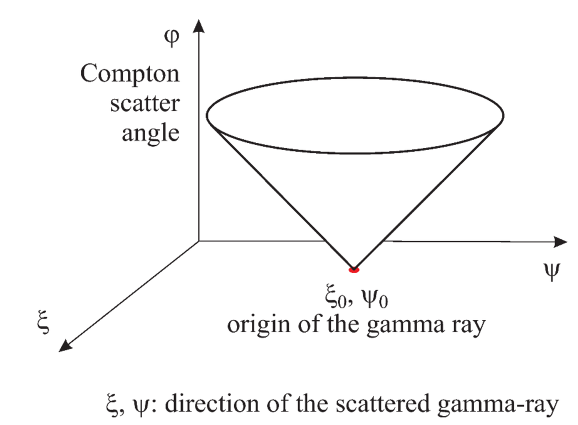

Fig. 4: The data space

In this 3D data space spanned by φ, ξ, and ψ, the point-spread function of a GRB describes a cone with 90 degree opening angle pointing at the origin location of the gamma rays, ξ_0 and ψ_0 (see Figure 4). The cone is smeared out by the energy and position resolution of the detector as well as Doppler-broadening.

Finding the cone is further complicated by the high background in the detector and the low count rate of these gamma-ray burst events (i.e. we might just have, e.g., 50 counts occupying the cone surface). On ground the localization can be performed using iterative image deconvolution techniques (such as Richardson-Lucy, Maximum-Entropy), however, the computational requirements for this approach are too high (e.g. [5]) to be implemented in a FPGA aboard a satellite. In addition, image deconvolution can not be infinitely parallelized, and thus will be always significantly slower than a neural network.

Therefore we investigated using a convolutional neural network to find the Compton cone of a GRB in the data space and thus determine the origin of the gamma-ray burst. The input for the network are the events entered into the 3D data space spanned by φ, ξ, and ψ. The data is feed into a modified VoxNet-like [6] neural network with 4 convolutional layers followed by 2 fully connected layers. The base model has the structure: C(64, 5, 2)–C(64, 3, 1)–FC(128)–FC(2), where C(f,d,s) indicates a 3d convolutional layer with f filters of size d and at stride s, Pool(m) indicates pooling with area m, and FC(n) indicates fully connected layer with n outputs. The output of the network is the location of the GRB (χ_0 and ψ_0), hence the output layer has 2 outputs.

We implemented our network using Python and TensorFlow, and optimized the network by trying different architectures of the network. Since the data cloud is in a geometric structure, identifying the location of the point of the cone doesn't require a complicated neural network. After comparing the performance with different architectures, we found out that the best structure to be used in our case has the structure:

C(64, 5, 2)–C(64, 3, 1)–P(2)–FC(128)–FC(2)

The additional pooling layer down-samples the feature space, enabling the network to generalize better with the dataset, since it doesn't need every point from the 3D data space to locate the position of the end point of cone (location of gamma ray). Hence, the network would be able to cover more variations of the dataset and the reduction in size would increase efficiency.

Results

For the training of the neural network we simulated 131,072 gamma-ray bursts with a random location on the sky and evaluated its performance with 2048 additional test GRBs. Each GRB contained 2000 gamma-ray events. The training took 48 hours on a Titan V, which was donated to us by the NVIDIA Corporation.

After training, the model determined a localization of the GRB in the sky in sky coordinates. These were compared to the true origin and the mean angular deviation as well as its RMS where determined.

Figure 5 shows the current results of the baseline model without pooling layers. We can achieve a median location accuracy of roughly 2 degree, which is within the expectations given the statistics of the GRBs. The training took 48 hours on a Titan V, which was donated to us by the NVIDIA Corporation.

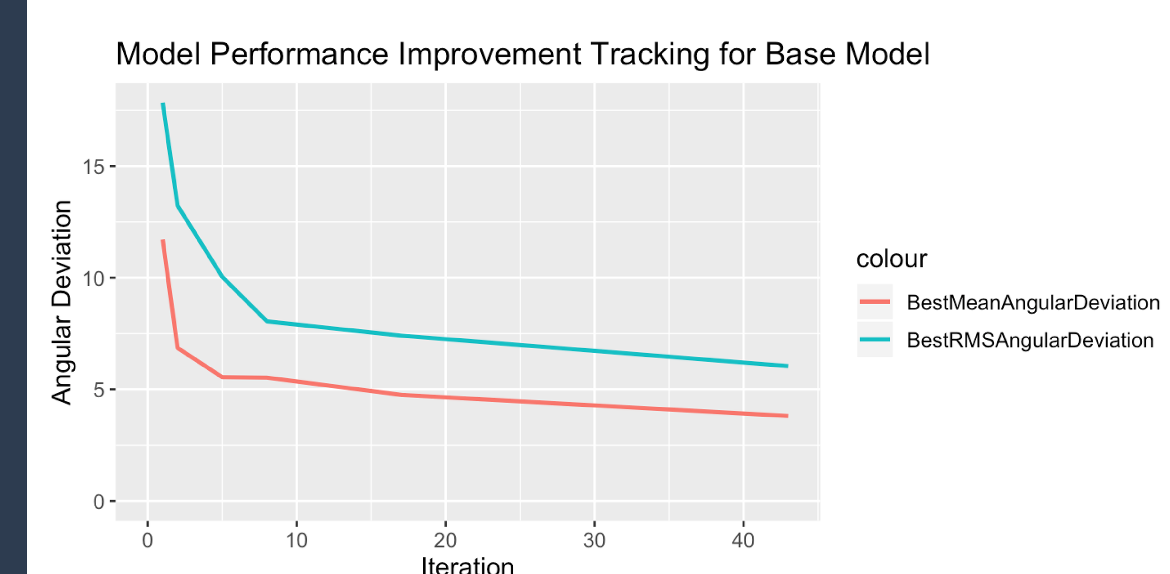

Figure 5: GRB localization performance of the base model

Figure 6 below shows the results after optimizations and adding a pooling layer. We can see that there's indeed a huge improvement in identification accuracy.

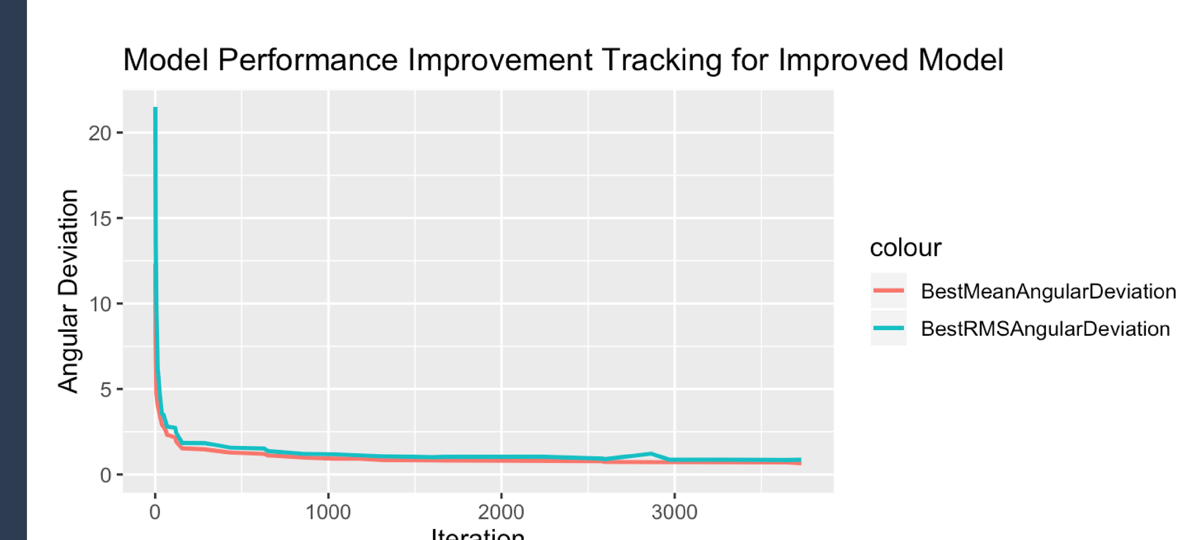

Figure 6: GRB localization performance of the improved model

Table 1 below show the resulting localization accuracy and iteration time as numbers along with the time spent per iteration.

	Baseline Model	Improved Model
Number of Iterations	43.000	3731.000
Best Mean Angular Deviation	3.809	0.655
Best RMS Angular Deviation	6.042	0.869
Total time training per Iteration	37.136	35.295
Total time testing per Iteration	2.699	2.486

Table 1: Comparison of base model and improved model's performance summary. The results are collected after there's no improvement for 1000 iterations.

Conclusions and Outlook

We have presented an approach which for the first time used a neural network to successfully localize sources measured with a Compton telescope.

One of the key limiting factors of the currently achieved location accuracy is the grid into which the events are binned. It is currently fixed at 5 degrees, while the angular resolution of the telescope can reach 1–2 degrees. This limit is a result of the network size which can fit into an FPGA aboard a satellite. Ideally the resolution should be below the angular resolution of the telescope, i.e. it should be around 1 degree. However, this would lead to 5^3 = 125 more data cells, which is beyond what we can handle at the moment. Therefore, one of the key future developments is to implement a sparse 3D convolution approach, since the number of entries into the grid (the number of gamma-rays) is much, much smaller than the number of bins in the grid.

Further next steps include to add more realism into the simulations. The current setup uses a simplified toy-model detector simulation and identical gamma-ray bursts. Therefore, the next steps include:

- Perform full Monte-Carlo simulations of GRBs with the e-ASTROGAM mass model
- Add realistic in-orbit background to the simulations
- Add the natural variation of the GRB's into the simulation (varying flux, spectral shape, duration, and light curve)

Acknowledgments

This work was performed as part of UC Berkeley's 2019 undergraduate research apprenticeship program (URAP) at the Berkeley Institute for Data Science (BIDS).
Development of the machine learning approaches was in part supported by the Gordon and Betty Moore Foundation through Grant GBMF3834 and by the Alfred P. Sloan Foundation through Grant 2013-10-27 to the University of California, Berkeley.
We gratefully acknowledge the support of the NVIDIA corporation with the donation of the Titan V GPU used for this research.

References

1. B. P. Abbott et al. "Multi-messenger Observations of a Binary Neutron Star Merger". *The Astrophysical Journal*, 848(2):L12, Oct. 2017.
2. B. P. Abbott et al. "Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A". The Astrophysical Journal Letters, 848(2), Oct. 2017
3. A. De Angelis et al. "Science with e-ASTROGAM", Journal of High Energy Astrophysics, 19:1-106, 2018
4. C. Kierans, "Detection of the 511 keV positron annihilation line with the Compton Spectrometer and Imager", PhD Thesis, UC Berkeley, Aug. 2018
5. A. Zoglauer et al. "Design, implementation, and optimization of MEGAlib's image reconstruction tool Mimrec", *Nuclear Instruments and Methods in Physics Research A*, 652(1):568–571, Oct. 2011
6. Daniel Maturana and Sebastian Scherer. "Voxnet: A 3d convolutional neural network for real-time object recognition", in *IEEE/RSJ International Conference on Intelligent Robots and Systems,* Sept. 2015.