MACHINE LEARNING FOR ASTROPHYSICS
2ND EDITION CATANIA, 8-12 JULY, 2024

Contribution ID: **218** Type: **Oral Presentation**

A new Deep Learning Model for Gamma-Ray Bursts' light curves simulation

Tuesday, 9 July 2024 16:50 (20 minutes)

R. Falco, N. Parmiggiani, A. Bulgarelli, G. Panebianco, M. Lombardi, L. Castaldini, A. Di Piano, V. Fioretti, C. Pittori, M. Tavani

In Multi-Messanger astronomy the rapid and precise detection of transient events such as Gamma-Ray Bursts (GRBs) is fundamental to alert the scientific community of new transients and allow them to perform follow-up observations. For this reason, many space mission teams are developing new detection algorithms for GRBs using classical and machine-learning technologies. To train or test these algorithms, it is necessary to have a large GRB dataset, but lack of data is a challenge. GRBs are rare events, and the available datasets are usually very limited in number. In addition, the existing datasets often only cover some of the different characteristics of GRBs with a balanced population. It therefore becomes essential to have a system to simulate GRBs.

This work aims to develop a Deep Learning (DL) based model for generating synthetic GRBs that closely replicate the properties and distribution of real ones. The DL model was trained to provide a larger set of artificial data that can be used for data augmentation and to develop detection algorithms both with classic techniques and machine learning methods.

We propose a new method to generate GRBs based on a 1D version of VAEGAN (Variational Autoencoder Generative Adversarial Network) based on Convolutional Neural Networks. To create the training dataset we extracted the light curves (LCs) of GRBs presented in the Fourth Fermi-GBM Catalog using only long GRBs (formal separation of T90 at about 2 seconds). To create our dataset, an LC was selected for each detector that detected the GRB as a data augmentation technique. Subsequently, a pre-processing phase was applied filtering only the best LCs resulting in 5964 LCs. After evaluating the distribution of the GRB duration (through the t90 parameter), we set the time series length at 220.

A quantitative analysis comparing the histogram of real LC rates (i.e. 1000 LCs never seen by the model) with the histogram of the synthetic rates (i.e. 1000 generated LCs) resulted in an Euclidean norm distance of 3.6 and a Wasserstein distance value of 0.1. The results show that the synthetic LCs are generated with similar properties to the real ones. Furthermore, we are also working on a conditional version of our model using some physical parameters of the GRBs for a more precise generation. This method can generate synthetic LCs for high-energy astronomy projects such as AGILE, CTA and COSI.

Presenter: FALCO, Riccardo

Session Classification: Time domain