PROCEEDINGS OF SCIENCE

ONLINE ICRC 2021
THE ASTROPARTICLE PHYSICS CONFERENCE
Berlin | Germany
37th International Cosmic Ray Conference
12–23 July 2021

Localisation and classification of gamma ray sources using neural networks

Chris van den Oetelaar,[a,*] **Saptashwa Bhattacharyya,**[b] **Boris Panes, Sascha Caron, Gabrijela Zaharijas,**[b] **Roberto Ruiz de Austri and Guðlaugur Jóhannesson**

[a] *Radboud University, Department of High Energy Physics, Heyendaalseweg 135, Nijmegen, The Netherlands*

[b] *University of Nova Gorica, Centre for Astrophysics and Cosmology, Vipavska 11c, Ajdovščina, Slovenia*

E-mail: c.oetelaar@student.ru.nl, sbhattacharyya@ung.si

With limited statistics and spatial resolution of current detectors, accurately localising and separating gamma-ray point sources from the dominating interstellar emission in the GeV energy range is challenging. Motivated by the challenges of the traditional methods used for the gamma-ray source detection, here we demonstrate the application of deep learning based algorithms to automatically detect and classify point sources, which can be applied directly to the binned Fermi-LAT data and potentially be generalised to other wavelengths. For the point source detection task, we use popular deep neural network structure U-NET, together with image segmentation, for precise localisation of sources, various clustering algorithms were tested on the segmented images. The training samples are based on the source properties of AGNs and PSRs from the latest Fermi-LAT source catalog, in addition to the background interstellar emission. Finally, we have created a more complex but robust training data generation exploiting full detector potential, increasing spatial resolution at the highest energies.

37th International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany

*Presenter

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PoS(ICRC2021)663

1. Introduction

Since its launch in 2008, Large Area Telescope (LAT) on board the Fermi Gamma-Ray Space Telescope has been surveying the whole sky in the energy range $\lesssim$ 30 MeV to more than 300 GeV. Integrating the data over the years, the Fermi-LAT collaboration has produced several generations of high-energy γ-ray source catalogs. The 4FGL catalogue[1], after a first release corresponding to 8 years of data, has been recently updated with an incremental 10 year version named 4FGL-DR2, which contains 5787 sources. While this large number of detected sources present an improvement by more than an order of magnitude with respect to the pre-LAT gamma-ray catalogs, the detection of faint sources remains a challenging task in the Fermi-LAT analyses. This is mainly due to the bright interstellar emission (IE), which is produced via interactions of the galactic cosmic-ray population with the interstellar gas and the interstellar radiation field. Detection and classification of faint sources is important for a range of scientific questions, from explaining the galactic center γ-ray excess to finding undetected sources that could belong to exotic possibilities such as Dark Matter sub-halos. Deep Learning (DL) is a subset of Machine Learning (ML) where artificial neural networks learn from large amounts of data and, the most prominent factors that contributed to the huge boost of DL are the appearance of large, high-quality, publicly available labelled data-sets, along with the empowerment of parallel GPU computing. DL methods have been shown to outperform previous state-of-the-art ML techniques in several fields, with computer vision being one of the most prominent cases and, Convolutional Neural Network (CNN) stand out as one of the most significant DL schemes. The U-Net architecture[2] is a popular CNN model which is shown to be powerful in addressing the task of semantic segmentation, in which, for each pixel in the image, a class-label is predicted. Motivated by the challenges of the traditional methods used for the gamma-ray source detection and classification, we discuss a CNN based pipeline that can be applied directly to the binned Fermi-LAT data and together with image segmentation network, we apply clustering algorithms to the segmented images to precisely localise sources.

This work is a continuation of a previous publication [3], that primarily focused on the network development. In the current work we are building a more robust data analysis pipeline using full detector potential and including more source classes, time variability of Blazars type sources (BLL and FSRQ) which we also intend to apply and test on the real data itself. In this proceedings we focus on deep learning and clustering 'source localization' algorithms, while the classification part will only be briefly introduced. This project is part of a large scale effort that aims to apply the same algorithm scheme to different wave-lengths, starting from Meerlichts optical data [4].

PoS(ICRC2021)663

2. Synthetic catalogue generation

500 simulated synthetic catalogues are generated in order to have enough data to train the networks, based on the Fermi-LAT 4FGL-DR2 catalogue. The simulations consist of 4 source classes, two AGN types (BLL and FSRQ), that are the most numerous source classes, and two galactic source classes (pulsars and a joint class of PWN and SNR). In addition we include a two component background (isotropic and interstellar emission), moddeled using the latest Fermi-LAT background model (gll_fermi_v07.fits and iso.fits). The background dominates the lower flux sources in the galactic disk.

2.1 Flux spectra

Two types of flux spectra are used for the simulation of the point sources, a logarithmic parabolic (LogPar) and an exponentially cut-off power law spectrum (PLEC). BLL, FSRQ and PWN+SNRs are simulated with the LogPar spectrum, while pulsars are simulated using PLEC. The LogPar function is shown in equation 1.

$$\frac{dF}{dE} = F_{0,class}\left(\frac{E}{E_0}\right)^{\alpha-\beta\, log(E/E_0)} \tag{1}$$

The choice of the parameters flux density ($F_{0,class}$), pivot energy (E_0), spectral slope (α) and curvature (β) then determine the spectrum of the source. The PLEC function of the pulsars is shown in equation 2.

$$\frac{dF}{dE} = F_{0,PSR}\left(\frac{E}{E_0}\right)^{-\Gamma} exp\left(a(E_0^b - E^b)\right) \tag{2}$$

For the pulsars, the spectrum of the source is determined by the choice of the flux density ($F_{0,PSR}$), pivot energy (E_0), low energy spectral slope (Γ), exponential factor (a) and the exponential index (b). For each point source, the parameters are drawn from different distributions. The F_0 from a normal distribution, E_0 from a log normal, β from a Gumbel, Γ from a normal, a from log normal and b from a Cauchy. The values for these distributions were determined from the sources of the class found in the 4FGL catalogue.

2.2 Spatial distribution and luminosity function

We assume two different spatial distributions for our source classes. As BLL and FSRQ are an extraglactic population, we assume a uniform distribution in the sky. The PSR and PWN+SNR are galactic sources, thus they are concentrated around a latitude of 0 in the galactic disk, simulated through a double Gaussian. The first Gaussian ($A = 0.11, \mu = 0°, \sigma = 1.39°$) and the second Gaussian ($A = 0.012, \mu = 0°, \sigma = 19.2°$) where A is a normalization factor between the two.
The synthetic catalogues follow the 4FGL catalogue for the sources with a flux above the Fermi-LAT detection threshold of $3.4 \cdot 10^{-13}\ erg\ cm^{-2}\ s^{-1}$ ($F_1^{E,TH}$). The 4FGL number of sources drop for flux below the threshold in the 4FGL catalogue. In order to be able to test the sensitivity of the network below $F_1^{E,TH}$, the counts are level with the peak for fluxes below the threshold in the synthetic catalogues.

PoS(ICRC2021)663

2.3 Skymap generation

The sky maps of the synthetic catalogues have been created using the Fermi-LAT science tools from the Fermi science support center. P8R2.ULTRACLEANVETO_V6 was used as the instrument response function. Depending on the location the γ-rays are tracked in the instrument, they are categorised as Front and Back type events. The spatial resolution for Front events are roughly a factor 2 better than for Back events at a given energy. Due to the difference in Front and Back resolutions and the limit of storage space and computing time, we only generate sky maps for Front events in the current work. In the future, Back event sky maps should be generated and incorporated into the training data structure as to make use of all the information available in the original Fermi-LAT data. The data is based on 10 years of the Fermi-LAT DR2 dataset from 4/08/2008 to 4/08/2018. Yearly binned exposure maps were derived from the corresponding spacecraft files in a RING ordered HEALPix format with the gtmodel from science tools. To remove contribution from the bright Earth limb, we apply a cut on the photon zenith angle, allowing only photons below the cut. Higher energy γ-rays (>1 GeV) have a zenith cut of 105° applied. Lower energy γ-rays (<1 GeV) have a zenith cut of 100°. This creates all-sky maps for each source and background type (6 in total) for each of the 6 energy bins for each of the 10 yearly bins. This results in 360 (6x6x10) HEALPix maps per synthetic catalogue. The data is split into individual years to generate images of the yearly variations of sources for use in classification. For the localisation, the 10 yearly maps are merged together during patch generation, described in the next subsection. Due to the different spatial resolutions for the different energy bins, the HEALPix maps have different Nside sizes for each bin. Nside determines the number of HEALpixels in the sky, $n_{pix} = 12 \cdot Nside^2$. The sizes of the all-sky maps are shown in table 1 showing the values for the Front maps. The resulting sky maps contain photon counts representing the average over infinitely many Poisson realisations of the sky, the infinite statistics scenario, also called the Asimov dataset. To create actual training and validation images, these Asimov values are then Poisson sampled, to reflect the Poisson statistics of the real data.

Table 1: The sizes of the all-sky Front maps and patches are shown per energy bin.

Energy bin (GeV)	0.3-0.5	0.5-1	1-2	2-7	7-20	20+
approximate resolution	0.8°	0.8°	0.4°	0.2°	0.1°	0.1°
HEALPix Nside	7	7	8	9	10	10
number of pixels over the sky	588	588	768	972	1200	1200
image width and height of patch (pixels)	16	16	32	64	128	128

3. Localisation

3.1 Patch generation

The contributions of all classes and backgrounds of all 10 years are combined into whole-sky images in each energy bin, still in HEALPix format. The combined whole-sky images are then cut into 10°by 10°patches and projected into Cartesian coordinates. The spatial resolution of the telescope depends on the energy of the gamma ray. Thus the energy bins were chosen such that

PoS(ICRC2021)663

there was a factor two for the resolution between the bands. This allows us to sample the higher energy bins at a higher resolution, resulting in more pixels per patch, also specified in Table 1. This means each patch is made up out of 6 images with different resolutions. These patches are the input of the network, the target of the network is a masked image. The mask is a 128 by 128 pixel image where a 5 pixel disk is drawn over all point sources present in the patch. This mask is what the network tries to recreate from the patch, thus creating disks where it locates a point source. An example of a patch with its mask is shown in Figure 1.

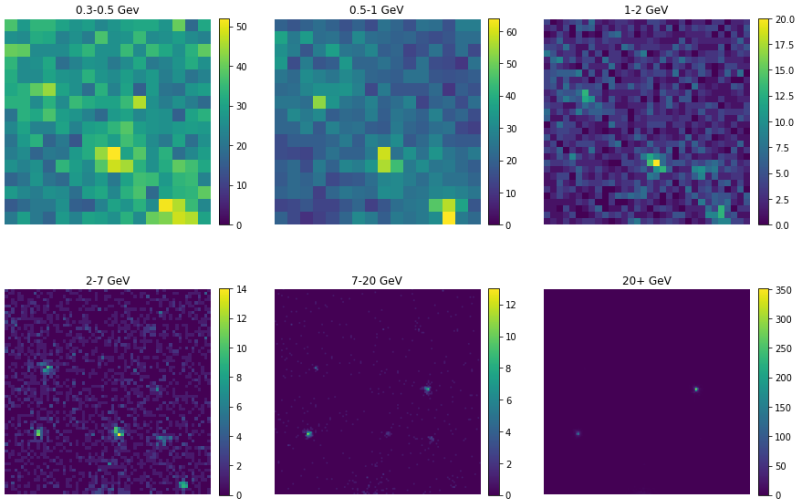

Figure 1: The Poisson sampled 6 energy bins of a patch with its corresponding mask.

768 patches are generated per catalogue according to the healpix pixelization sceme. With the patch size of 15° this creates overlap between all the patches, thus assuring that every part of the sky and every source in it is featured on at least one of the patches. The origin coordinates for the HEALPix pixelization scheme are randomly re-positioned on the sky between each catalogue to assure that every patch from one catalogue to another does not feature exactly the same background. The counts generated in these patches represent the infinite-statistics scenario, meaning the expected averages of infinite Poisson realisations of the patch, also called the Asimov dataset. The counts in these Asimov patches or dataset is then Poisson sampled during training to produce the actual images that the network trains and are evaluated on. Currently with 169 synthetic catalogues created and processed, there are a total of 129,792 patches. At the moment, a total 1000 catalogues have been generated and are being set up for processing. The patches from 10 catalogues were used as a test set, 30 for a validation set and the remaining 119 were used as a training set for the network. The training set will be used to train the network, the validation set will be used to score the performance of the networks during training.

PoS(ICRC2021)663

3.2 Localisation network architecture

The network used to predict the location of sources in the sky is based on the U-net architecture. This provides a per pixel classification of background or point source. This however means that clusters of pixels still need to be clustered together to get a prediction of point source locations. In this project we use a method based on OpenCV's blob-detector instead of the more resource expensive k-means clustering used in the previous iteration. In the classical U-net architecture, the input image goes through convolutional layers before the dimensions of the image is halved while increasing the number of channels. This is repeated several times until the desired lowest resolution is reached, the so called bottleneck. Here the information of the input images is compressed into a very small image with a high number of channels.

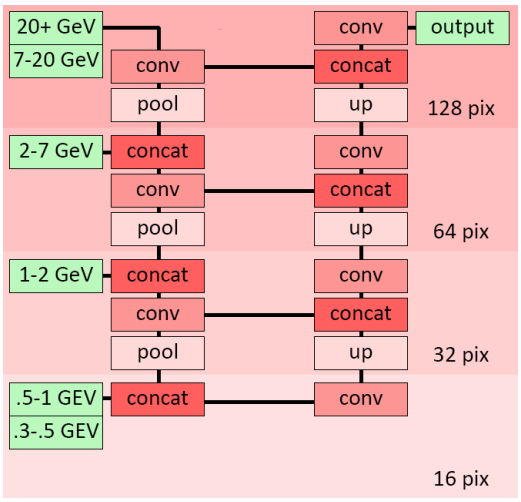

Figure 2: Modified U-net for multi resolution input shown in the left side column. Each tier has convolutional layers (conv), concatenate layer for merging multiple inputs (concat) and maxpooling (pool) or upsampling (up) to halve or double resolutions.

In steps, the resolution of the image is then doubled again and put through convolutional layers. Meanwhile, convolutional layers before the bottleneck are linked with the equal sized convolutional layers located after the bottleneck. This combines information of the large and small scale structures in the image in order to make accurate small scale predictions on large images. In this case however, the images in the patch are different sizes, thus they can not be input in the same layer of the network. The multi version input suited for the different resolutions for each energy bin is shown in figure 2. The 20+ and 7-20 GeV images with a size of 128 pixels are input in the first layer, then after halving the resolution the 2-7 GeV image of size 64 is input. The 1-2 GeV image of size 32 is merged into the network in the next step, and lastly the 16 pixel sized 0.5-1 and 0.3-0.5 GeV images are inserted.

3.3 Clustering

The target of the network is a 128 size image with 5 pixel radius disks of value 1 on each point source and 0 for all other locations. To translate the output of the network to point source locations, we use a method based on OpenCV's blob detection. First, the output has a disk shaped filter or kernel applied to it. The kernel slides over the output and sums all values located within a 5 pixel radius, the same shape disk as the targets have in the training data. This gives a high value for locations in the image that have high value pixels surrounding it in a disk shape, lower values at edges of high value pixel clumps and low values in areas of low value pixels, where the network predicts background. This essentially grades how much each pixel is at the center or a disk in the prediction. The resulting filtered images then has a threshold (initialised at 30) applied to it. Only pixels above the threshold are selected, and neighbouring pixels are grouped together. Per group, the pixel with the highest value is taken as the prediction of the location of a point source. If there

PoS(ICRC2021)663

are groups with more then 5 pixels, the threshold is increased by 10 and the process is run until no groups larger than 5 remain. Through this iterative process, sources which are close enough together for their disks to overlap are discerned as separate sources, rather than a single source. This is especially important in the more densely populated regions in the galactic disk.

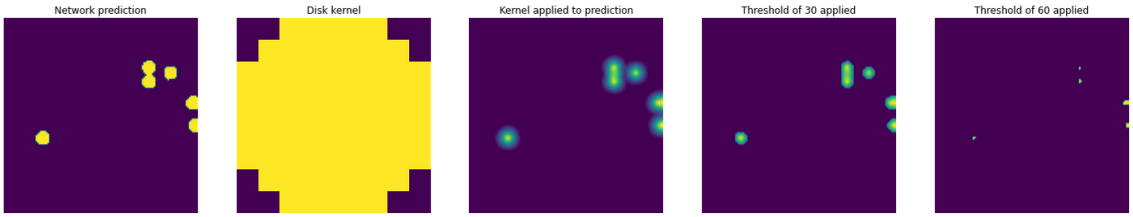

Figure 3: Aspects of the clustering process from network prediction to the disk kernel and application of thresholds.

3.4 Results localisation

The output layer of the network has a sigmoid activation function, thus outputs values between 0 and 1. For evaluation, a threshold of 0.5 is applied to assign a 0 or 1 to each pixel. This results in a binary accuracy (Acc_b) of the network of 0.982 . However, this is a misleading metric for the accuracy of the network, due to 96% of the pixels consisting of background. A prediction of only background would score an accuracy of 0.965 (Acc_0). Thus the accuracy is adjusted (Acc_a) according to equation 3.

$$Acc_a = \frac{Acc_b - Acc_0}{1 - Acc_0} \tag{3}$$

This results in an adjusted accuracy of 0.48 for the test dataset. The predictions of the test set are then run through the clustering method to get the location predictions of the point sources. There are two performance metrics for the location prediction. The completeness is a function of the number of sources located correctly (TP) and sources not located by the network (FN), it measures the fraction of sources present found. The completeness is a function of the number of sources located correctly (TP) and sources not located by the network (FN), it measures the fraction of true sources in the sources found.

$$completeness = \frac{TP}{TP + FN} \tag{4}$$

$$purity = \frac{TP}{TP + FP} \tag{5}$$

Tuning the sensitivity of the clustering algorithm, a higher completeness can be exchanged for a lower purity and vice versa. A Tuning was chosen that had a completeness of 0.58 and a purity of 0.90. Examples of the results of localisation are shown in figure 4. The localisation predictions of the network are then used to produce 7 by 7 pixel images of each source for classification. This classification step then determines the class of object (BLL, FSRQ, PSR, PWN+AGN and Fake), filtering out the incorrect localisation predictions as Fake sources.

PoS(ICRC2021)663

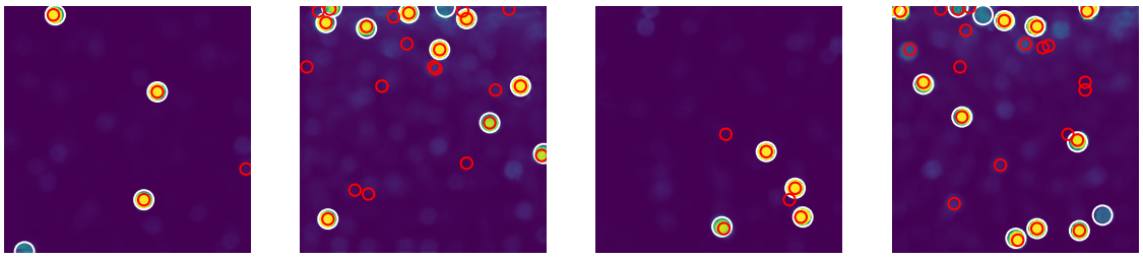

Figure 4: Mask predictions as output of the localisation network with dark purple as predicted background and yellow as predicted point sources, the red circles mark the true position of sources, the white circles mark the positions predicted by the network.

4. Classification

For image generation, instead of different resolutions for different energy bins, we adjust the size of the image per energy bin. The images are 7 by 7 pixels with a width of (4°,4°,2°,1°,0.5°,0.5°) for the energy bins (0.3-0.5,0.5-1,1-2,2-7,7-20,20+)GeV. In addition to the 6 energy bins, 10 yearly bins are also added. In the yearly bins, all energy bins are summed and binned pear year of observation time. This captures yearly fluctuations in the source, which we hope to use to distinguish BLL from FSRQ sources. The dataset generated this way is once again an Asimov dataset, which is Poisson sampled to train and evaluate on. The training target for each image is the class of the object depicted from the list (BLL, FSRQ, PSR, PWN+AGN, FAKE). Where FAKE are a combination of random points of background in the sky and false positives from the localisation network. The input of the network is not only the image of 16 bins but also the location in the sky (lon,lat) of the image.
The network for classification uses convolutional and maxpooling layers to reduces the size of the input image before flattening the tensor to a 1d array. After flattening, the position of the image is provided as input to the network as well. The activation function of the output layer is Softmax, forcing the outcome to the categories (BLL, FSRQ, PSR, PWN+AGN, FAKE) to sum to 1. The highest valued category is then taken for the evaluation of the networks accuracy.
Once sufficient catalogues have been processed into images, the classification network can be trained and evaluated properly.

5. Outlook

The previous paper [3] has already illustrated the power of this method for localisation and classification. The new iteration is significantly more ambitions, amongst other tings, increasing the number of source classes from two to four. The next steps include seeing if and how the addition of the time variability parameter to the classification algorithm through the addition of the yearly images will affect classification capabilities. We also plan to apply the localisation and classification method to the real Fermi data. This will indicate the methods capability of automatising catalogue generation in gamma-rays. The efforts to test the same method on optical data through Meerlicht could open the door to a method fit for multiple wavelengths.

PoS(ICRC2021)663

References

[1] Abdollahi, S. etal. Fermi Large Area Telescope Fourth Source Catalog. ApJs247, 33 (2020) 1902.10045

[2] Ronneberger, O. etal. U-Net: Convolutional Networks for Biomedical Image Segmentation. (2015) 1505.04597

[3] Caron, S. etal. Identification of point sources in gamma rays using U-shape convolutional neural networks and a data challenge. (2021) 2103.11068

[4] Hosenie, Z. etal. MeerCRAB: MeerLicht Classification of Real and Bogus Transients using Deep Learning, Experimental Astronomy. (2020) 2104.13950

PoS(ICRC2021)663