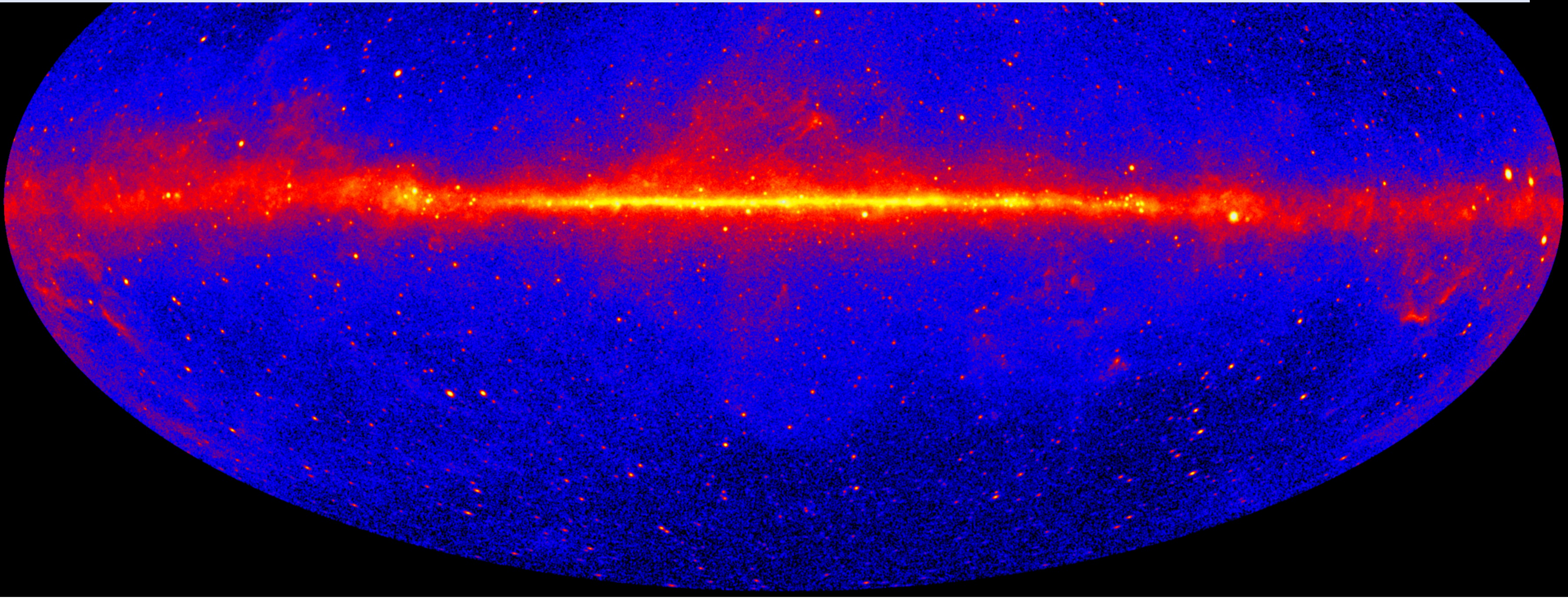
Towards foundation model for astrophysical source detection: An End-to-End Gamma-Ray Data Analysis Pipeline Using Deep Learning



S. Bhattacharyya, S. Caron, D. Malyshev, R. Nicolas, G. Principe, Z. Rokavec, R. Ruiz de Austri, D. Skočaj, F. Stoppa, D. Tabernik, G. Zaharijas

Judit Pérez Romero

judit.perez@ung.si

SMASH machine learning for science and humanities postdoctoral program

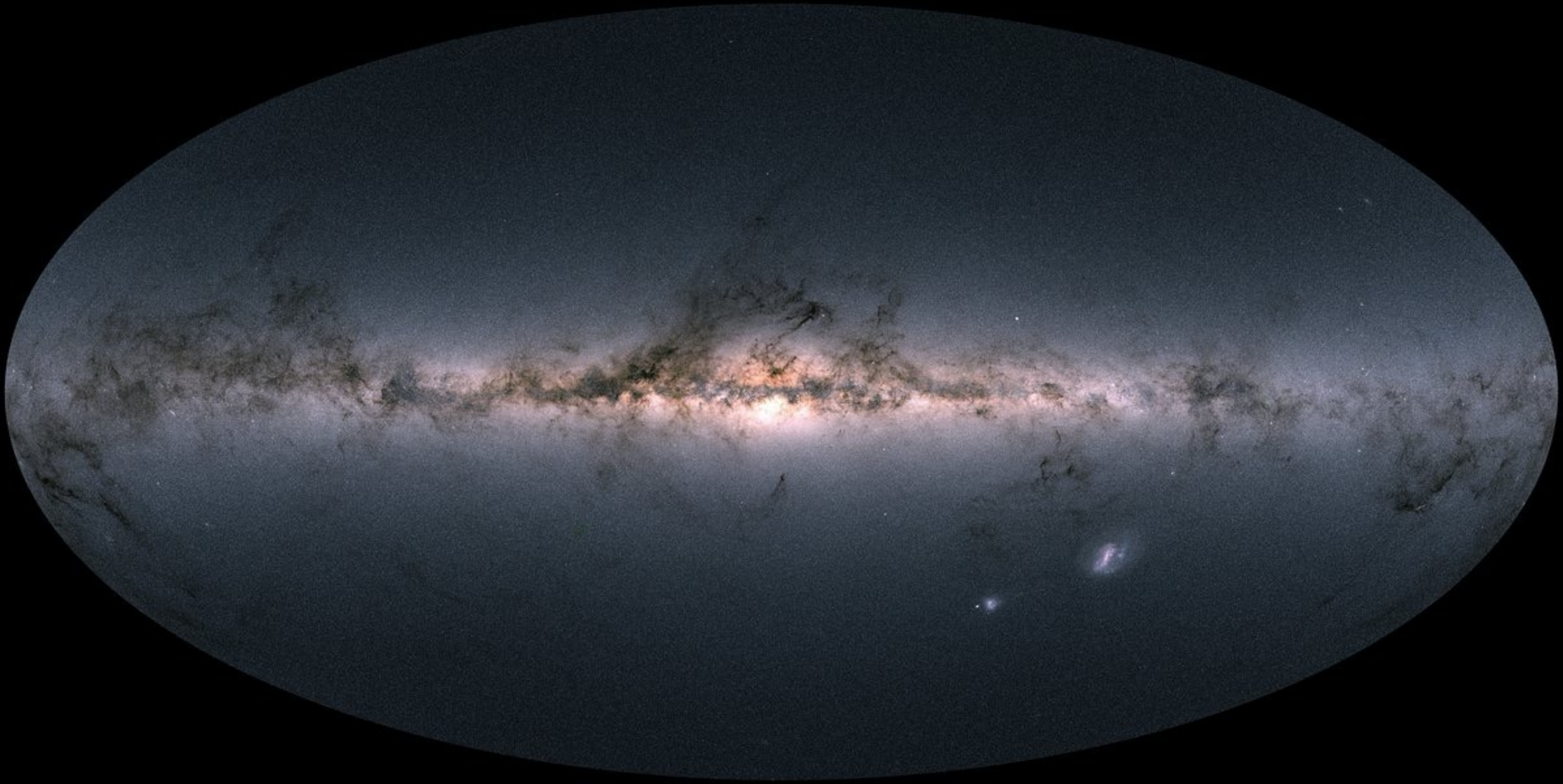
Co-funded by the European Union

I FEEL SLOVENIA

UNIVERZA V NOVI GORICI SCIENTIA VINCES 1995

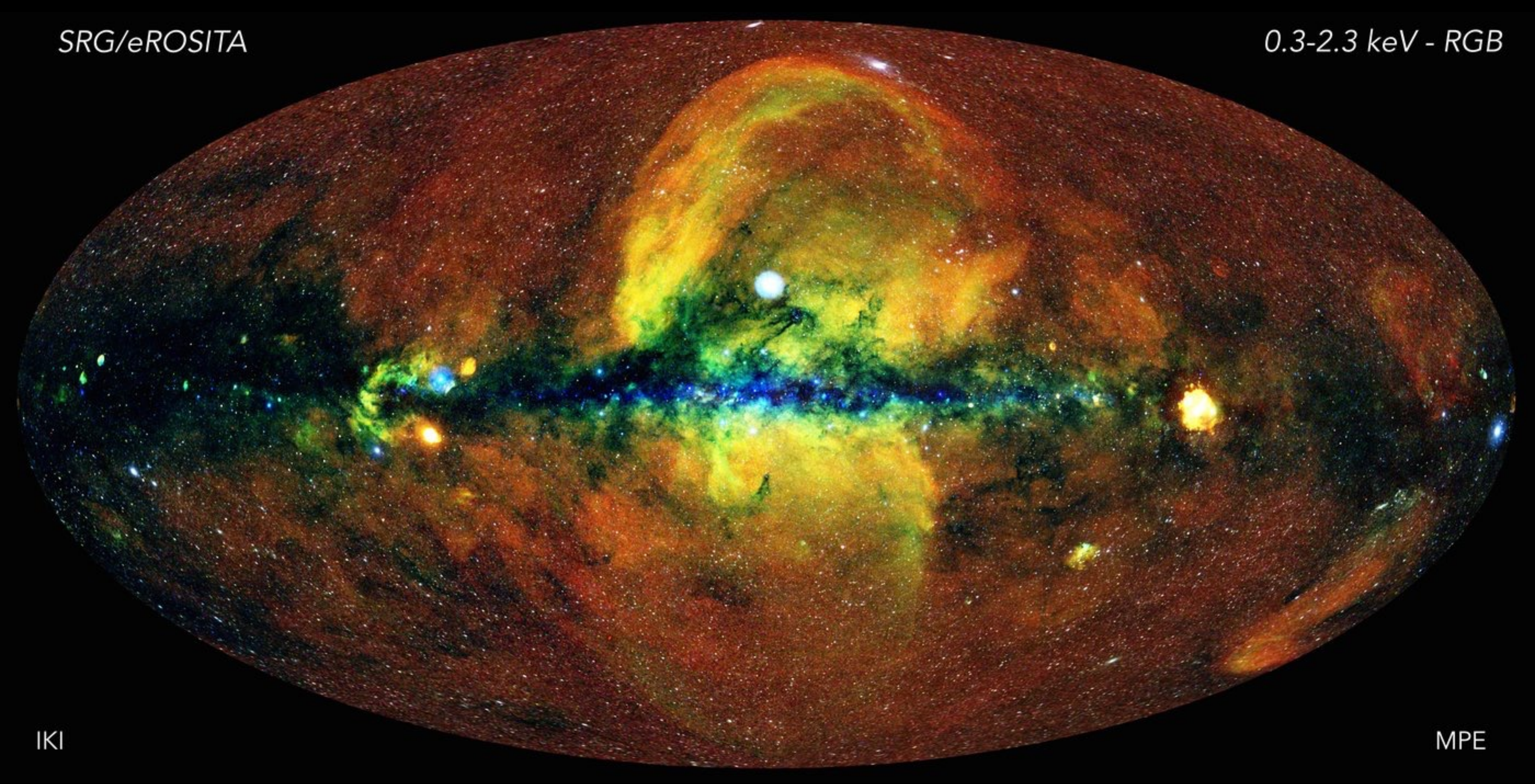
EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE EuCAIFCon 2025

Cagliari, Italy
17/06/2025

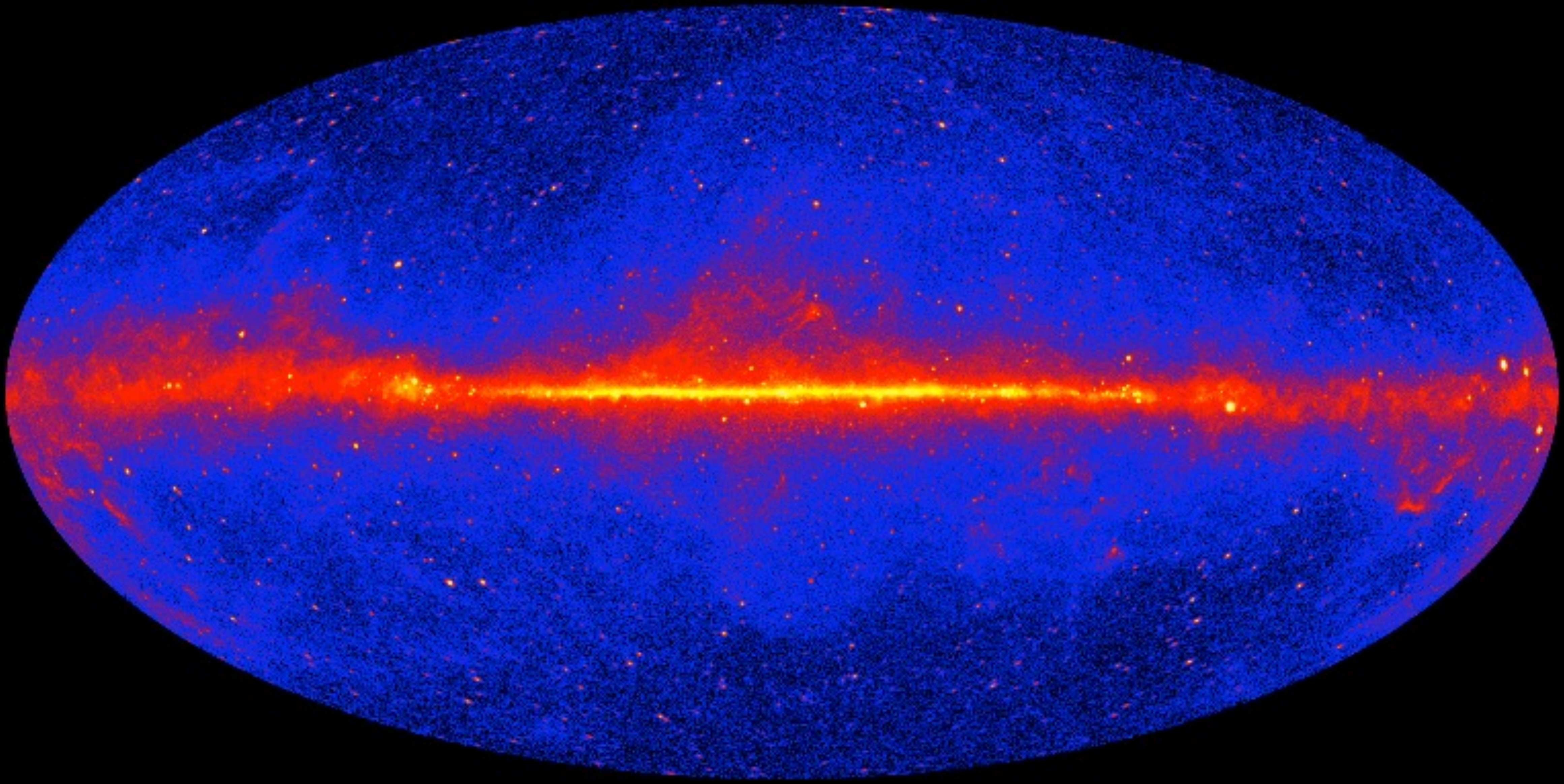


Total brightness and colour of stars seen by Gaia between July 2014 and May 2016

https://www.eso.org/public/images/ESA_Gaia_DR2_AllSky_Brightness_Colour_black_bg_8k/



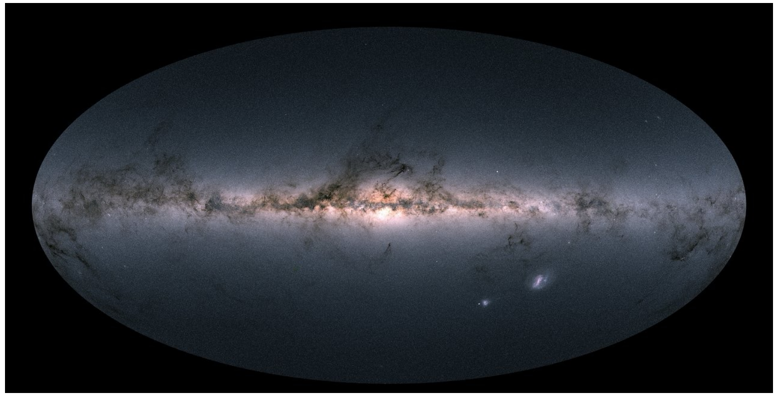
https://www.mpe.mpg.de/7461950/erass1-presskit

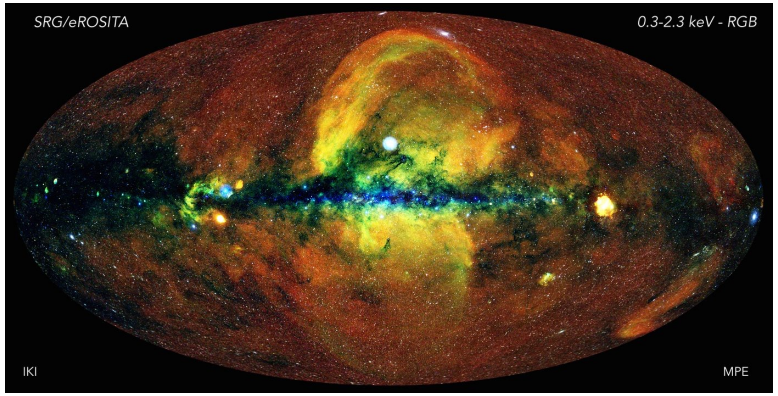


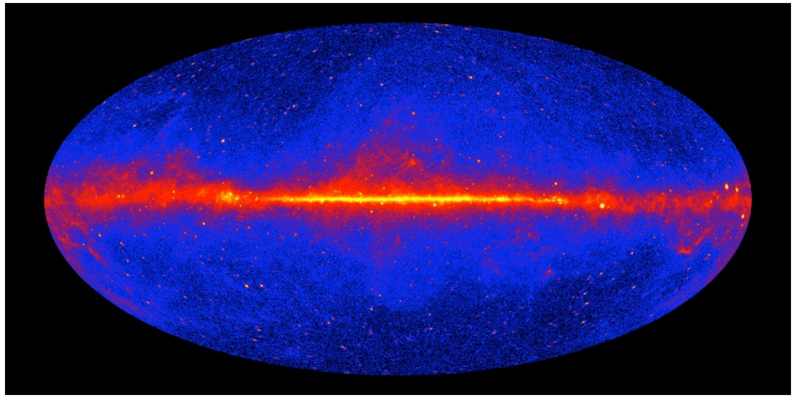
Gamma-ray emission > 1 GeV from 12 years of *Fermi*-LAT data

https://fermi.gsfc.nasa.gov/

DETECTION OF ASTROPHYSICAL SOURCES ACCROSS WAVELENGTHS







Final goal

Given a sky-map, can a DNN-based pipeline detect localized sources with catalogue properties?

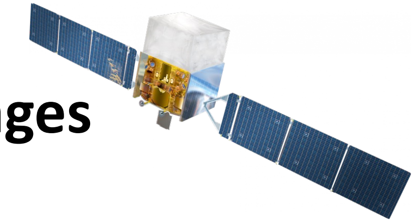
- locations [longitude, latitude - deg]
- extension [deg]
- flux [above certain energy/wavelength]

Can these methodologies be applicable across different wavelengths?

OUTLINE

1. **Gamma-ray analysis pipeline:**

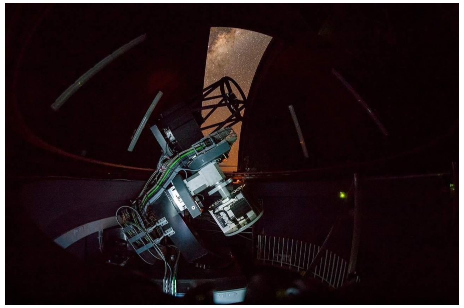
    1. **Current data: *Fermi*-LAT all-sky images**

    2. **Future data: CTAO simulations of the galactic plane survey (GPs)**

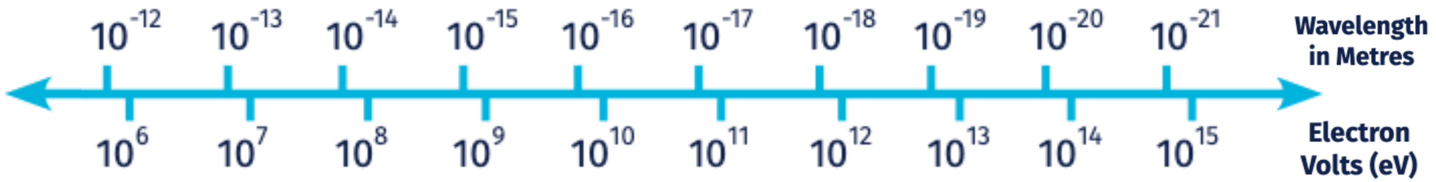
2. **Optical analysis pipeline**



3. **Towards foundation model for astrophysical source detection**

THE GAMMA-RAY SKY

- Different techniques for gamma-ray detection:

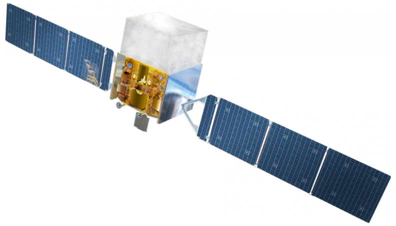


THE GAMMA-RAY SKY

- Different techniques for gamma-ray detection:

Fermi Large Area Telescope (LAT)

- Space-borne telescope
- Best sensitivity at ~200 GeV
- Angular resolution up to 0.1 deg
- 17 years of all-sky data

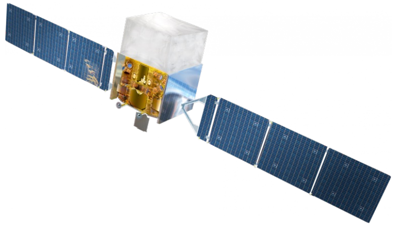


THE GAMMA-RAY SKY

- Different techniques for gamma-ray detection:

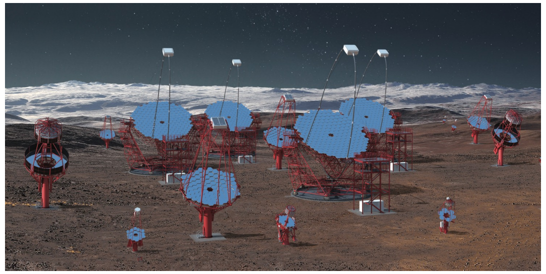
Fermi Large Area Telescope (LAT)

- Space-borne telescope
- Best sensitivity at ~200 GeV
- Angular resolution up to 0.1 deg
- 17 years of all-sky data



Cherenkov Telescope Array Observatory (CTAO)

- Ground-based telescope array
- Best sensitivity at ~1 TeV
- Angular resolution ~0.05 deg
- Deep dedicated surveys
- *O*(10) better sensitivity w.r.t. current telescopes



THE GAMMA-RAY SKY

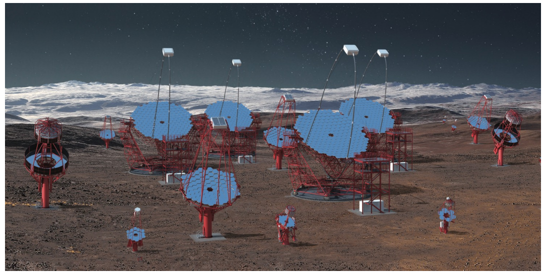
- Different techniques for gamma-ray detection:

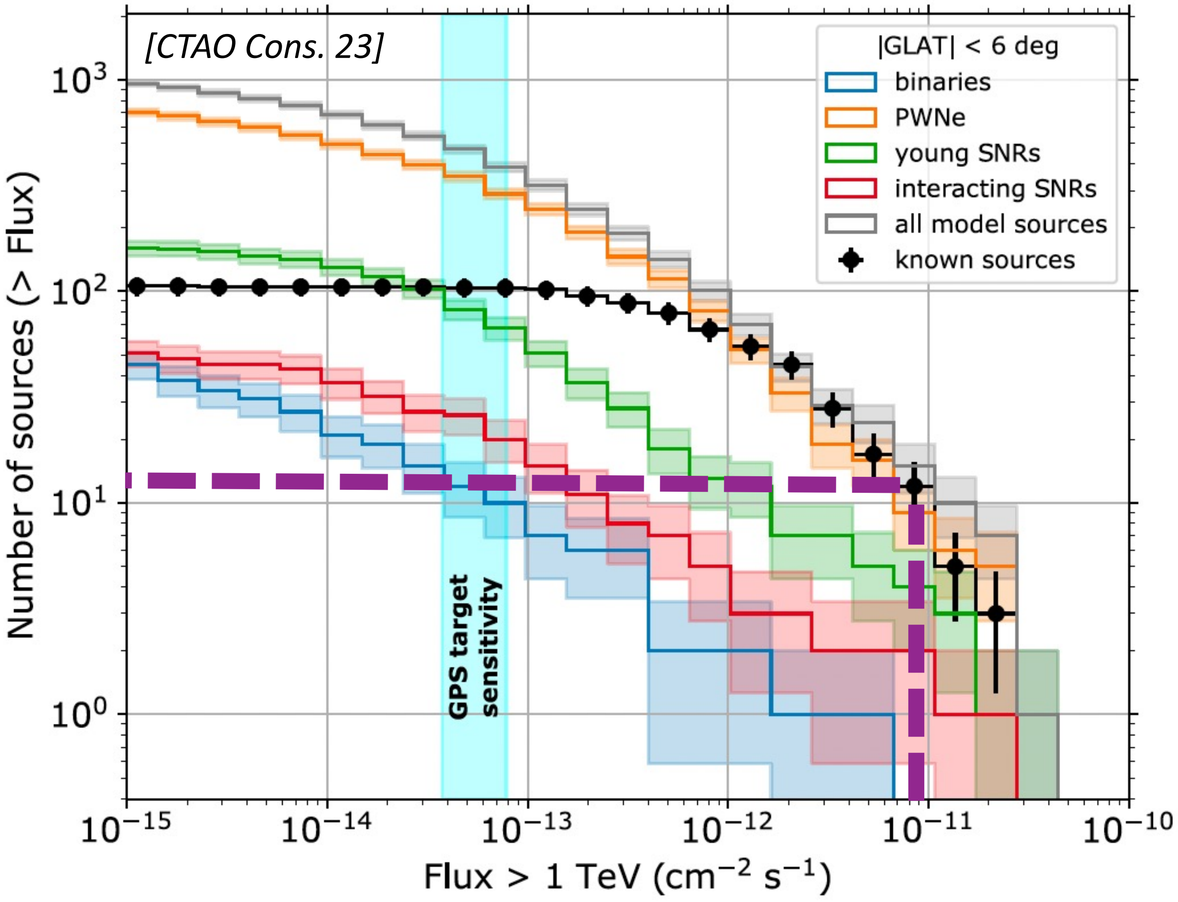
Fermi Large Area Telescope (LAT)

- Space-borne telescope
- Best sensitivity at ~200 GeV
- Angular resolution up to 0.1 deg
- 17 years of all-sky data

Cherenkov Telescope Array Observatory (CTAO)

- Ground-based telescope array
- Best sensitivity at ~1 TeV
- Angular resolution ~0.05 deg
- Deep dedicated surveys
- *O*(10) better sensitivity w.r.t. current telescopes





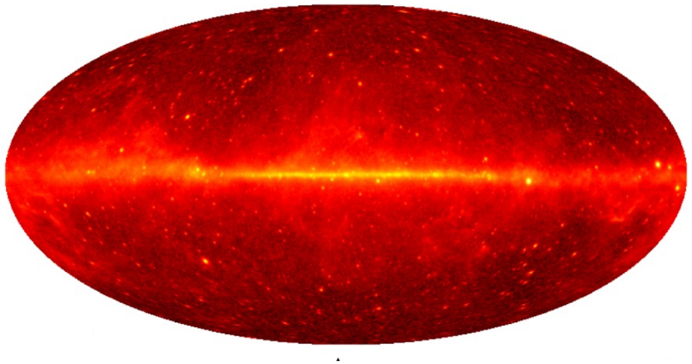
THE GAMMA-RAY SKY

- Different techniques for gamma-ray detection:

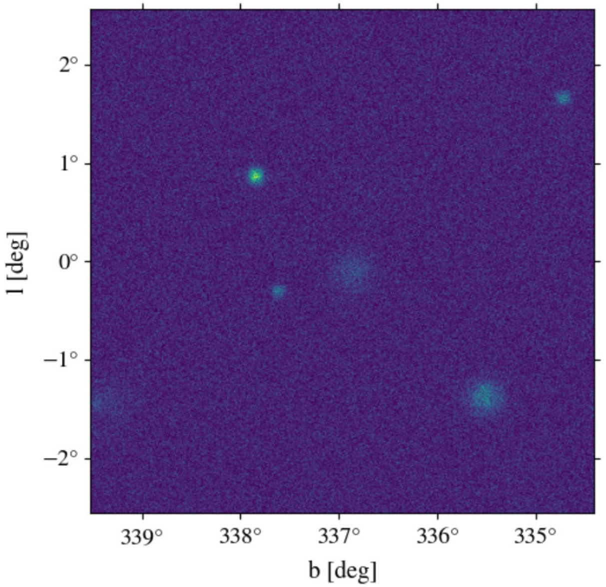
Fermi Large Area Telescope (LAT)

Cherenkov Telescope Array Observatory (CTAO)

Data



Data

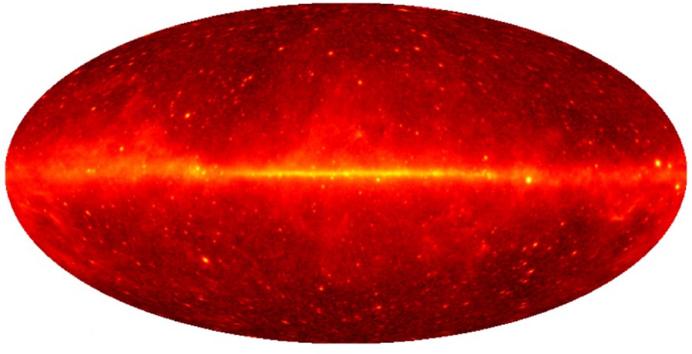


THE GAMMA-RAY SKY

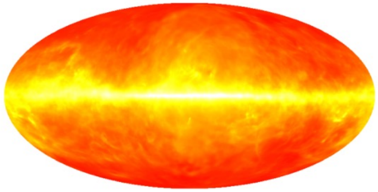
- Different techniques for gamma-ray detection:

Fermi Large Area Telescope (LAT)

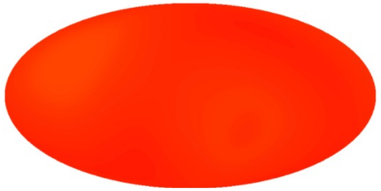
Data



Backgrounds



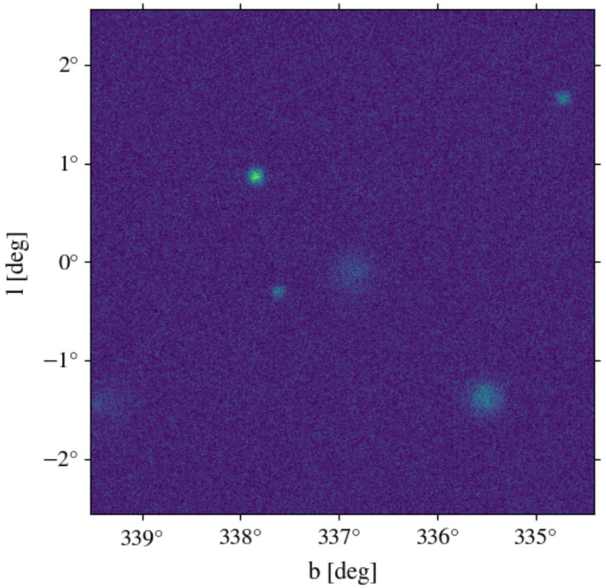
Interestellar emission (IE)



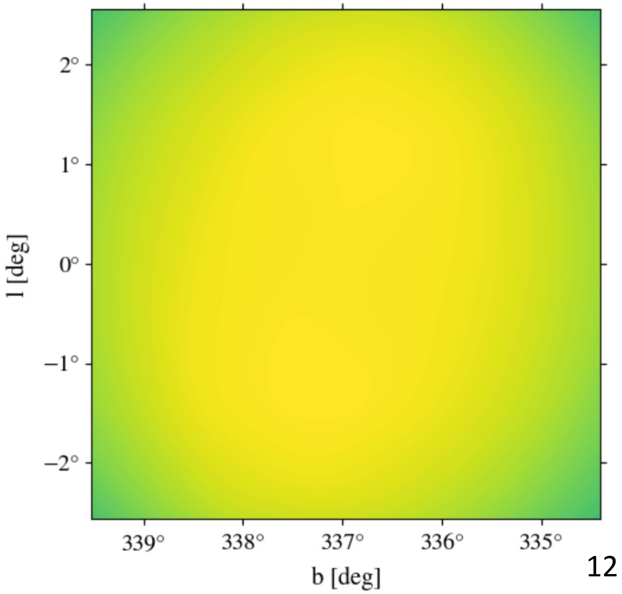
Isotropic background (IGRB)

Cherenkov Telescope Array Observatory (CTAO)

Data



Backgrounds



THE GAMMA-RAY SKY

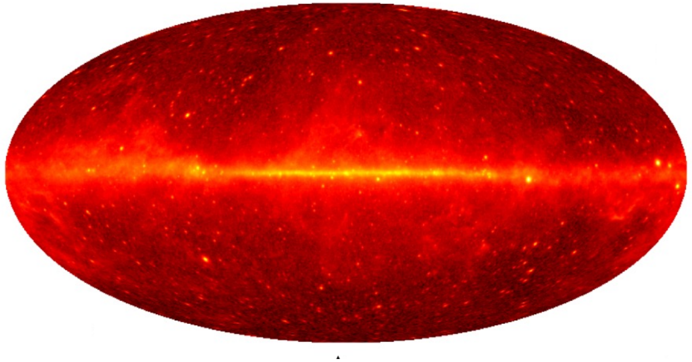
- Different techniques for gamma-ray detection:

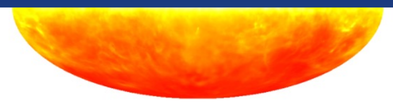
Fermi Large Area Telescope (LAT)

Cherenkov Telescope Array Observatory (CTAO)

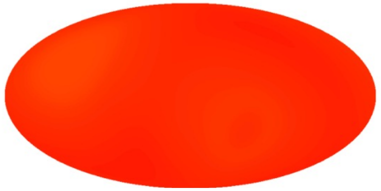
Removing/modelling backgrounds is critical for faint source detection

Data

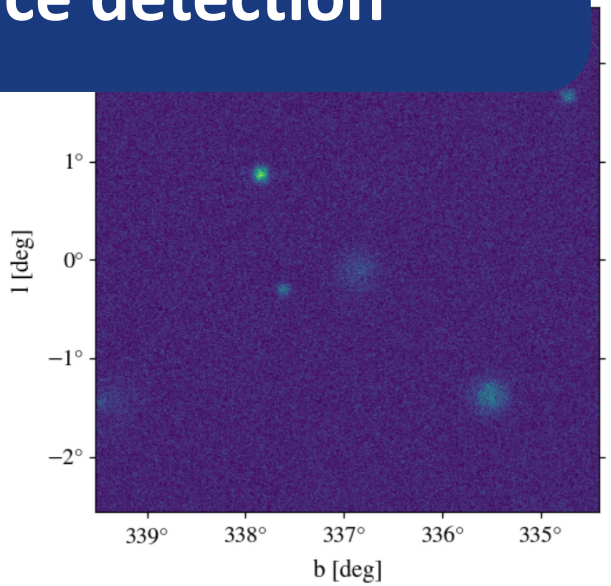




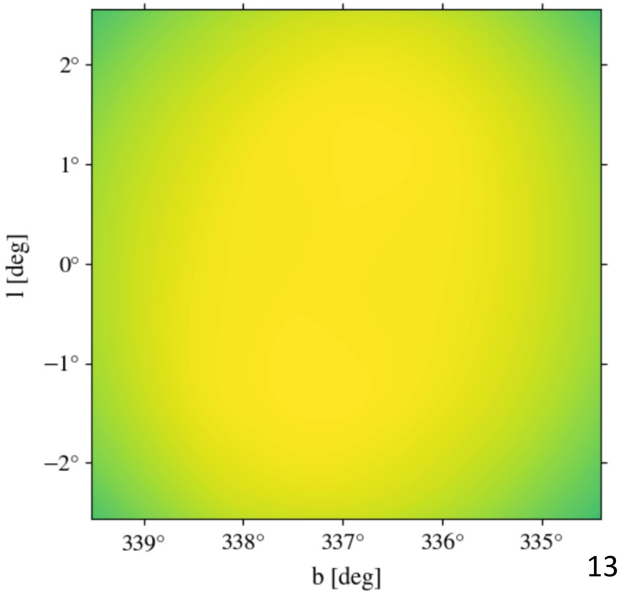
Interestellar emission (IE)



Isotropic backgound (IGRB)

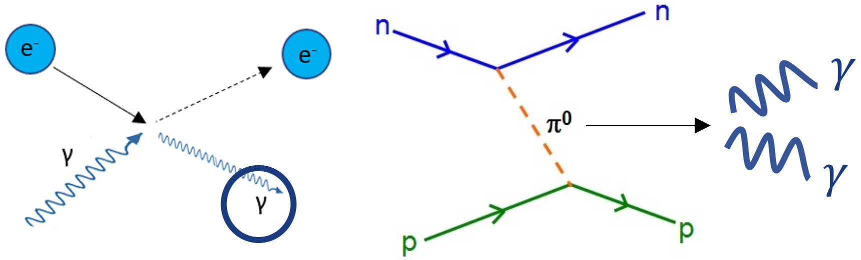


Backgrounds

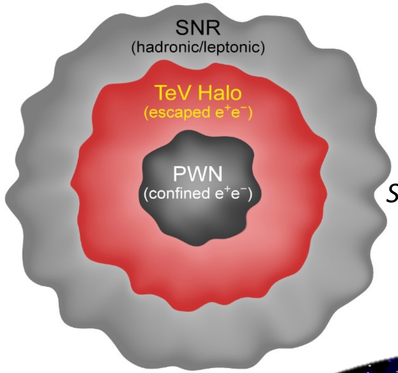


THE GAMMA-RAY SKY: FAINT SOURCES

- Study faint gamma-ray sources can lead to different discoveries:
  - **Cosmic ray production & populations**

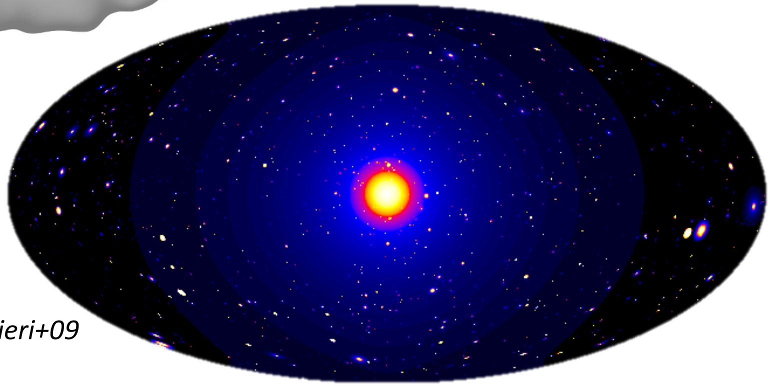


  - **New classes of sources**: TeV Halos *[Linden+17]*



Sudo+19

  - **Fundamental physics:** Indirect Dark Matter detection



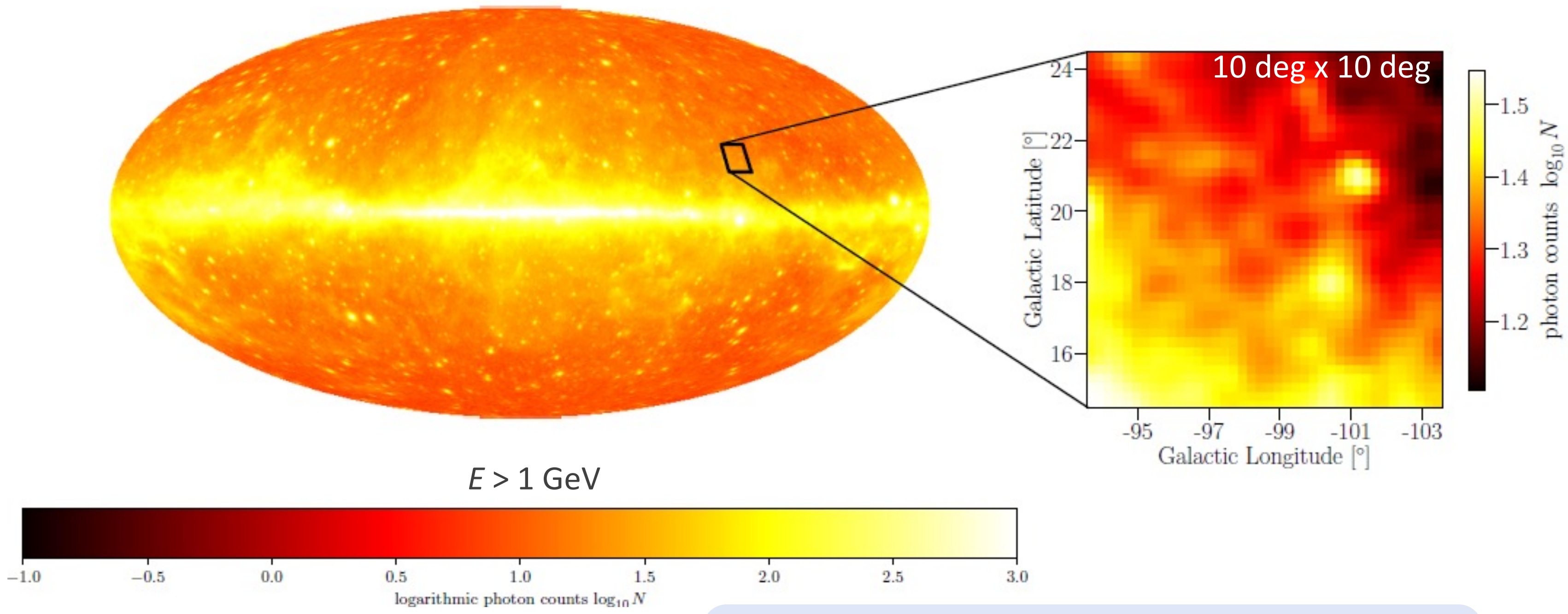
Pieri+09

FERMI-LAT: BUILDING A COMPLEMENTARY GAMMA-RAY CATALOG

- Simulation of 10 years of data with collaboration software [fermitools] and population models from previous catalogs [4FGL-DR2] for training

Analyzing gamma-rays of the Galactic Center with deep learning
S. Caron, G. A. Gómez-Vargas, L. Hendriks & R. Austri, JCAP05(2018)058, [arXiv: 1708.06706]
Identification of point sources in gamma rays using U-shaped convolutional neural networks and a data challenge
B. Panes, S. Caron, R. Austri, G. Zaharijas *et.al*, A&A (A62, 2021), [arXiv: 2103.11068]

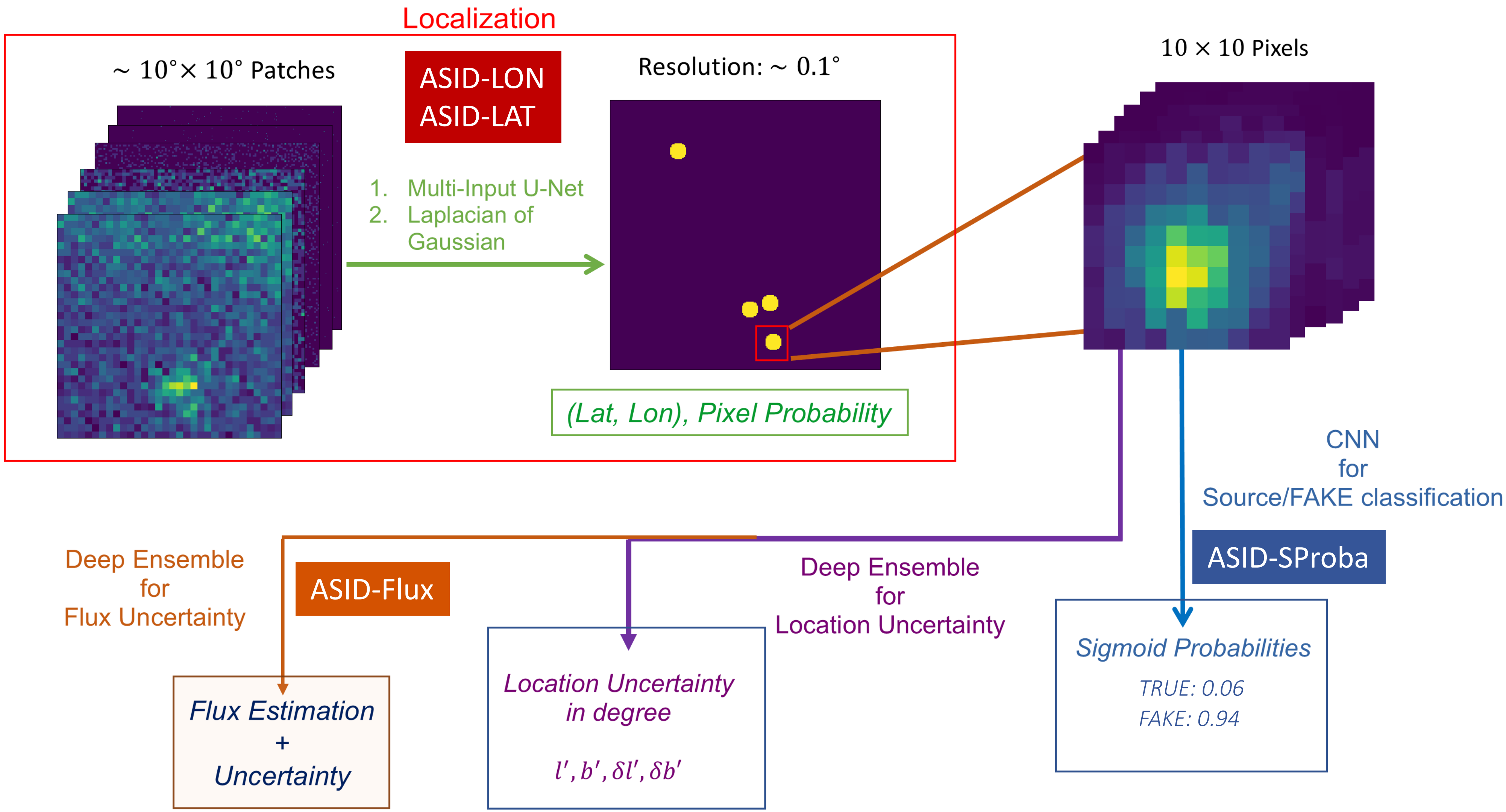
- 300 MeV < E < 1 TeV, 6 energy bins
- Spatial resolution increases with energy: from 0.8 deg at 0.3 GeV, to 0.1 deg above 7 GeV
- Test robustness against different IE models



IE small scale structures ⇄ Misidentification of faint sources → Can DNN help obtaining a background agnostic catalog?

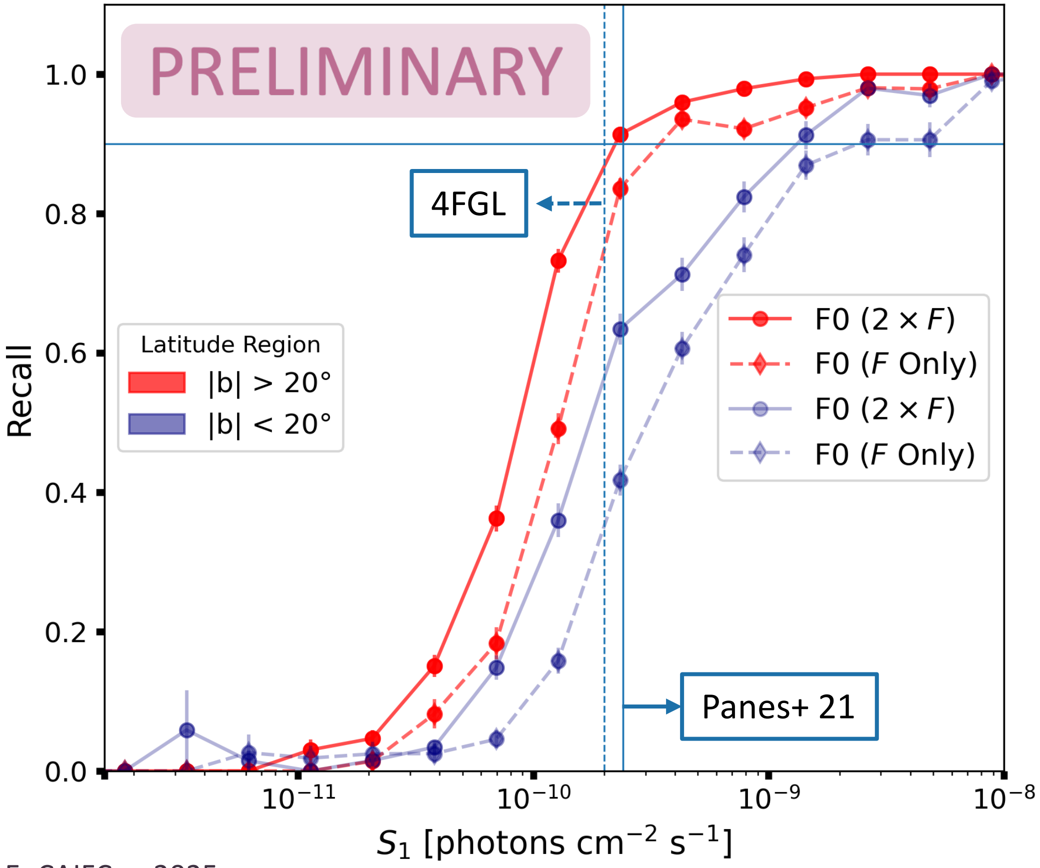
FERMI-LAT: BUILDING A COMPLEMENTARY GAMMA-RAY CATALOG

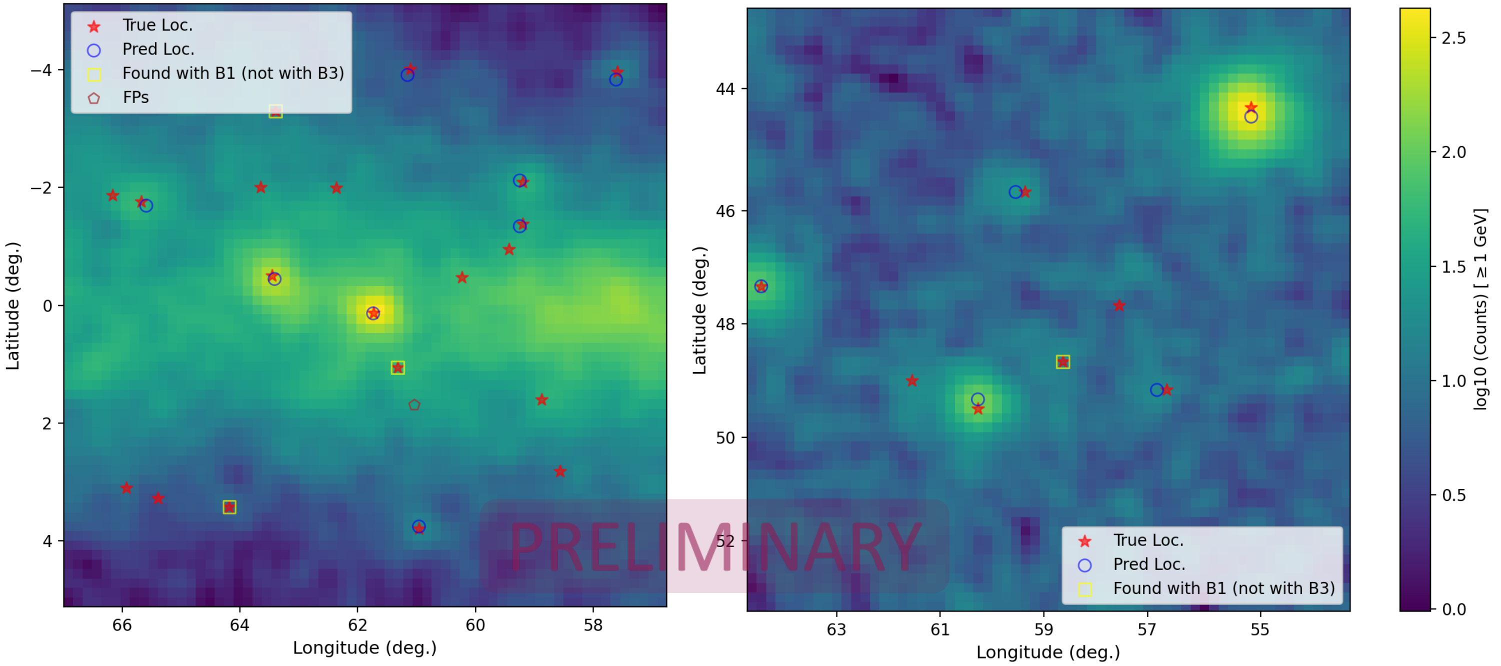
Complete Workflow of Source Location, Flux Estimation (+Uncertainty), Classification



FERMI-LAT: BUILDING A COMPLEMENTARY GAMMA-RAY CATALOG

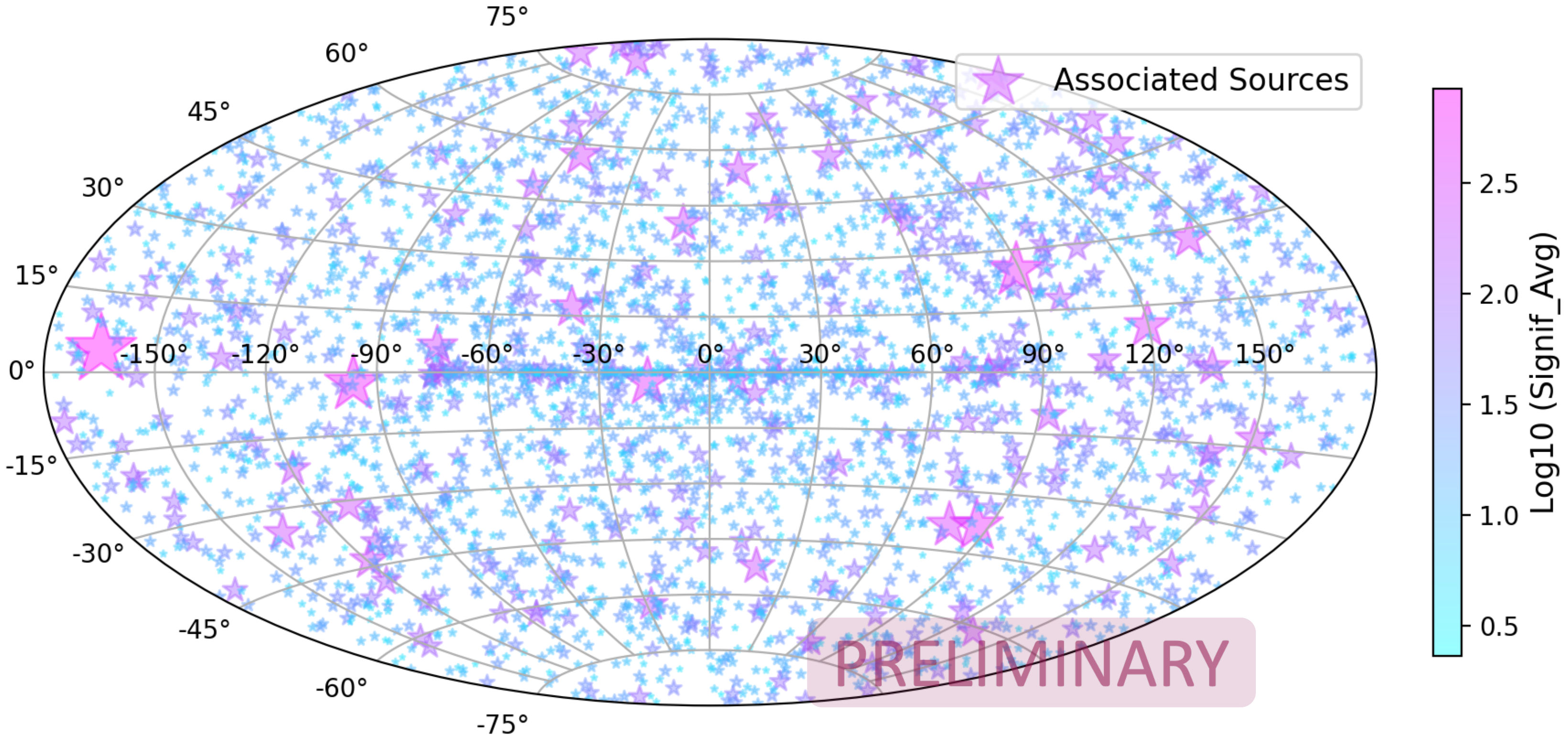
- Flux sensitivity of ASID is comparable to the 4FGL-DR2 detection threshold and to our previous work
- To test the robustness against different IEM models, use the pipeline trained with B1 IEM and test it with catalogs prepared using B1 and B2 IEMs separately





FERMI-LAT: BUILDING A COMPLEMENTARY GAMMA-RAY CATALOG

- Comparison with 4FGL-DR2 catalog:
  - For sources with $\boldsymbol{\sigma > 20}$, more than **90% association** independent on latitude
  - For sources with **20 < $\boldsymbol{\sigma}$ < 10** and **|*b*|> 20 deg, 90% association**
  - For sources with **20 < $\boldsymbol{\sigma}$ < 10** and **|*b*|< 20 deg, 77% association**



Final catalog-like product

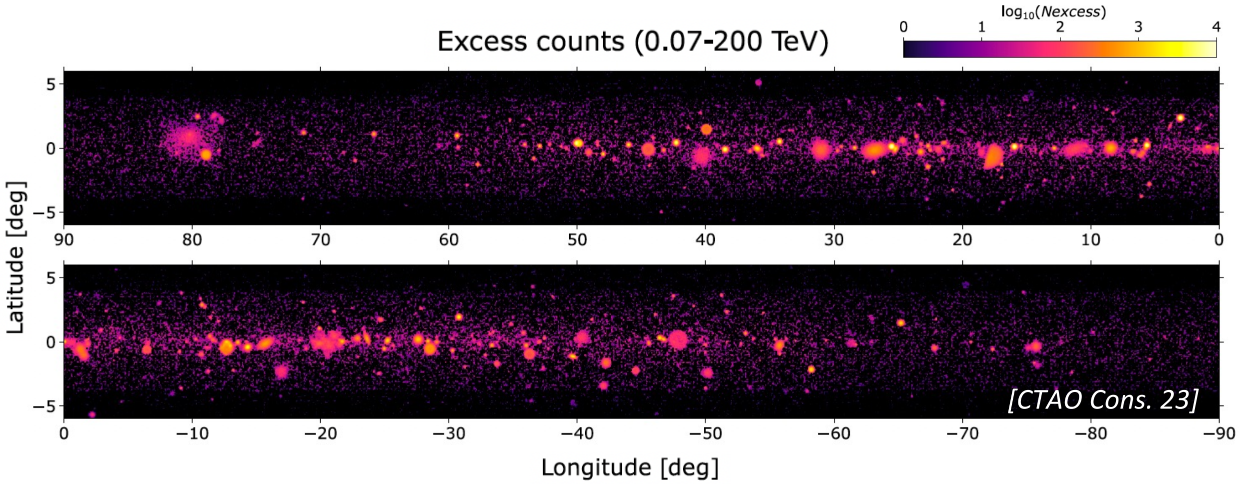
ASID-LON ° (deg.)	ASID-LAT ° (deg.)	ASID-SC Binary Class (0/1)	ASID-SProba Sigmoid Probability	ASID-Flux ph. cm^{-2}s^{-1}	DR2-Name
287.603	−0.627	0	$3.18e-12$	$6.037e-8$	4FGL J1045.1-5940
304.097	−45.137	0	$2.69e-4$	$2.827e-10$	4FGL J0040.7-7157
349.823	9.238	0	$3.94e-4$	$2.627e-10$	4FGL J1643.3-3148
⋮	⋮	⋮	⋮	⋮	⋮

CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY

- Improvement of CTAO sensitivity will imply detection of:

Fainter sources ⇄ Larger sources

- Crowded regions will be extremely complex to analyze



CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY

- Improvement of CTAO sensitivity will imply detection of:

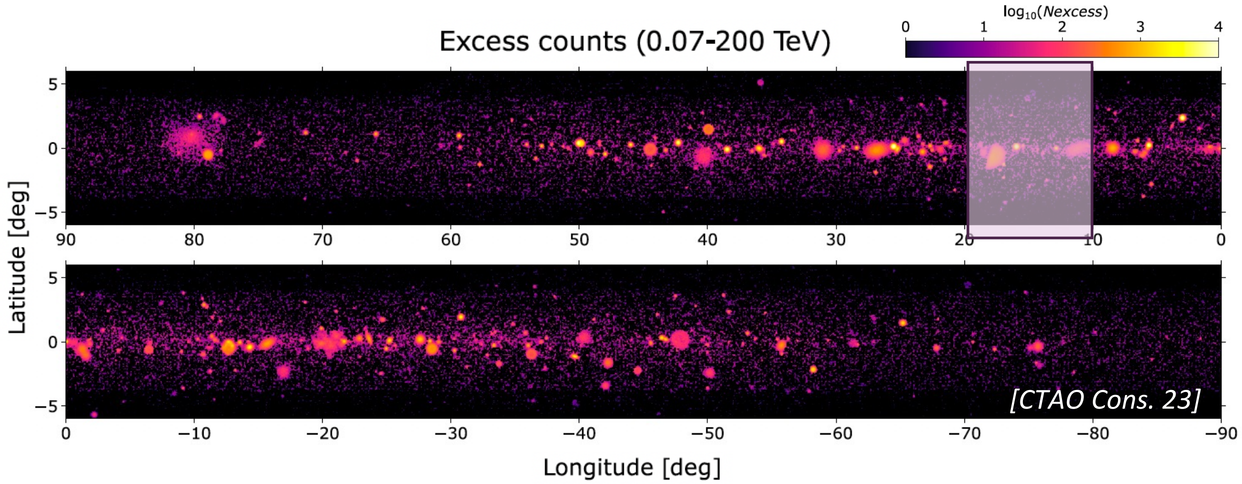
Fainter sources ⇄ Larger sources

- Crowded regions will be extremely complex to analyze

Simplify the problem

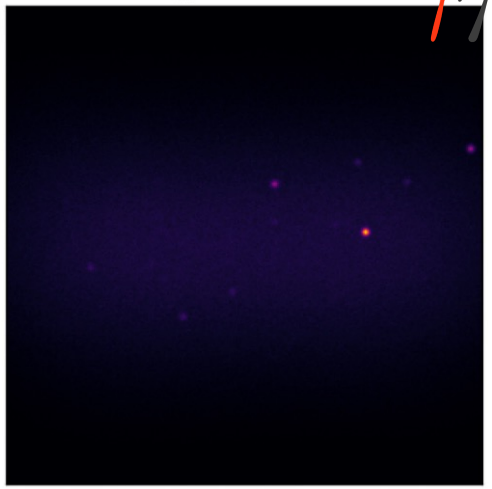
- Simulate a toy gamma-ray sky with only point-like sources
- Keep the original spatial and spectral distribution of sources
- One patch: 10 x 10 deg (512 x 512 pix), 3 energy bins

Proof of concept:
Learn how to improve detection and localization of sources in CTAO toy-simulation



Simplify the problem

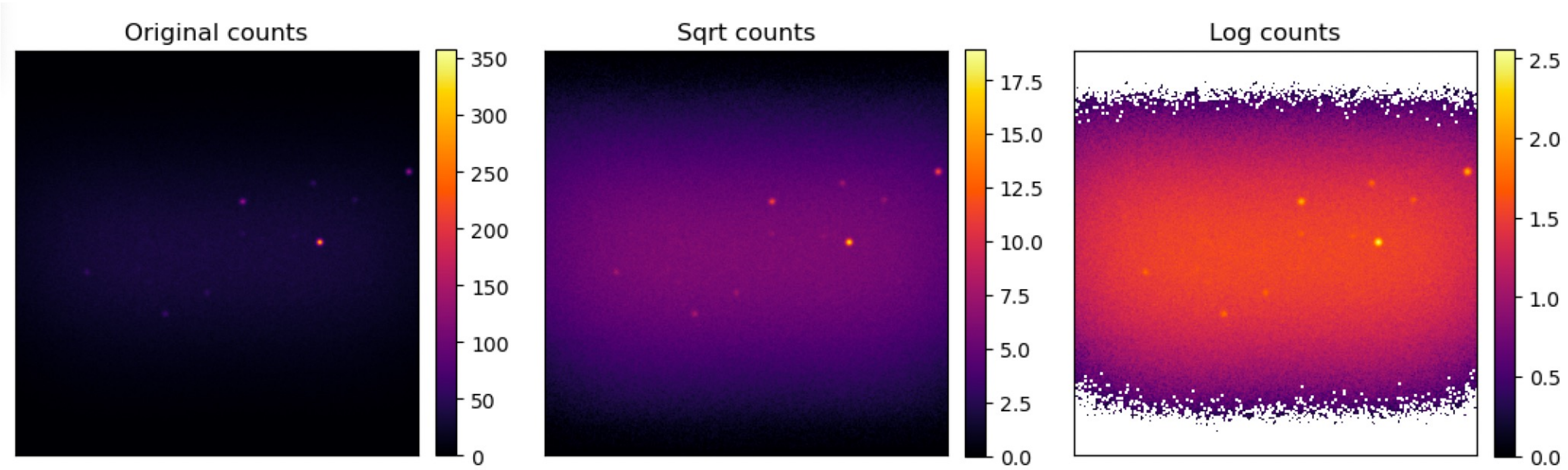
A **Python** package for **gamma-ray** astronomy



CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY

1. Image scaling using ASID

- Naturally enhances contrast between sources and background

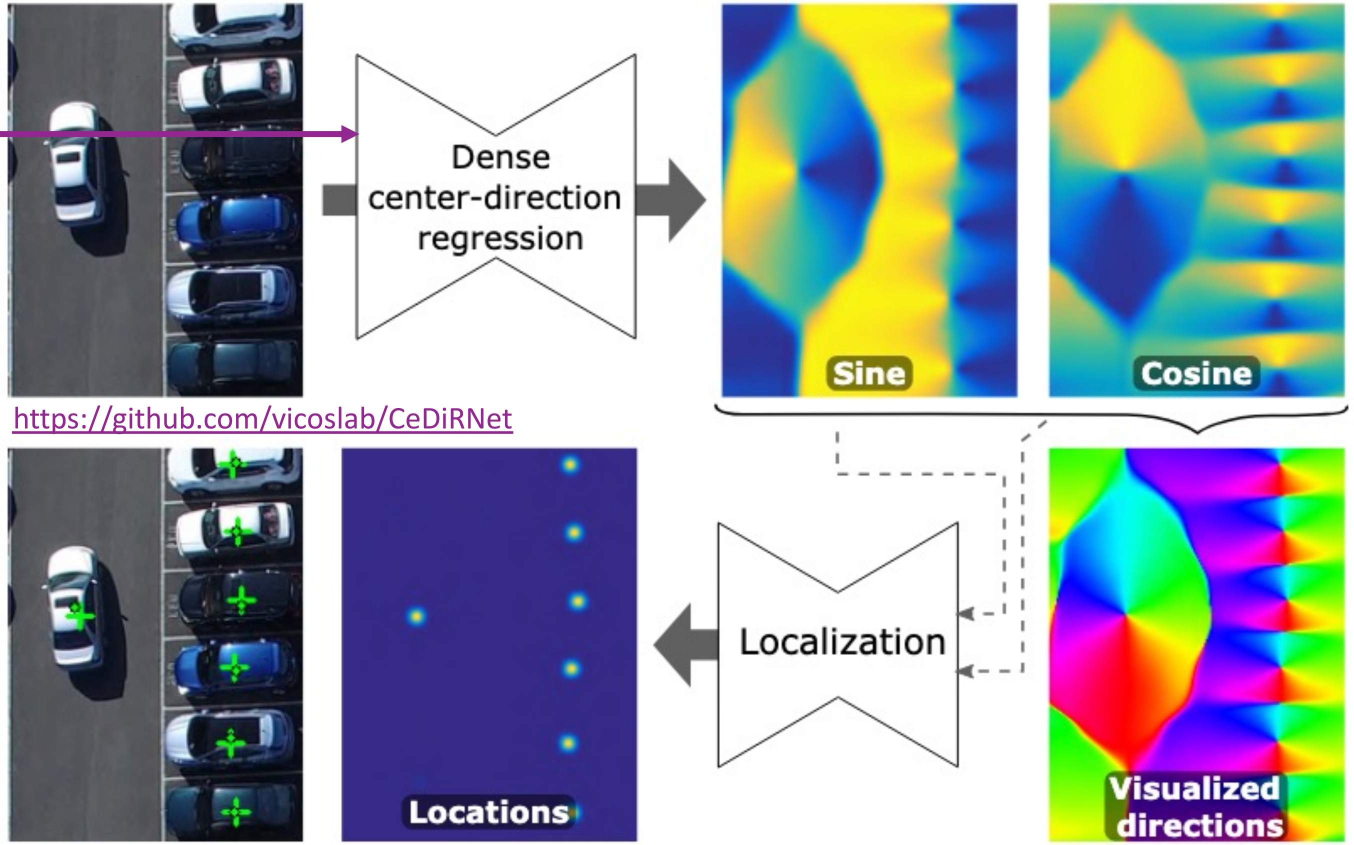


CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY

2. Dense center-direction regression approach

Dense Center-Direction Regression for Object Counting and Localization with Point Supervision,
D. Tabernik, J. Muhovič, D. Skočaj, J. Pat. Cog. 2024, 110540, [arXiv:2408.14457]

Densely regresses directions pointing to the nearest object center



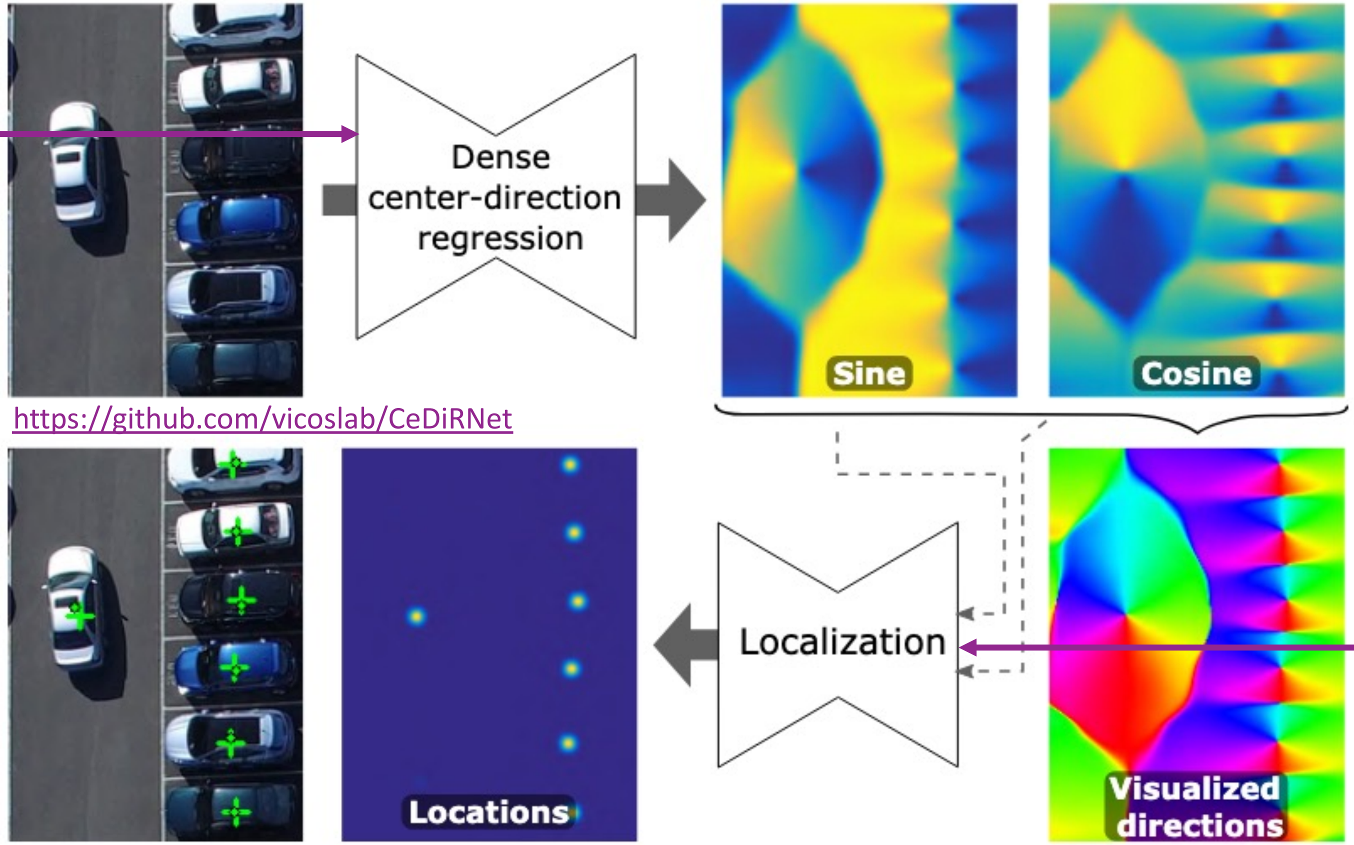
https://github.com/vicoslab/CeDiRNet

CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY

2. Dense center-direction regression approach

Dense Center-Direction Regression for Object Counting and Localization with Point Supervision,
D. Tabernik, J. Muhovič, D. Skočaj, J. Pat. Cog. 2024, 110540, [arXiv:2408.14457]

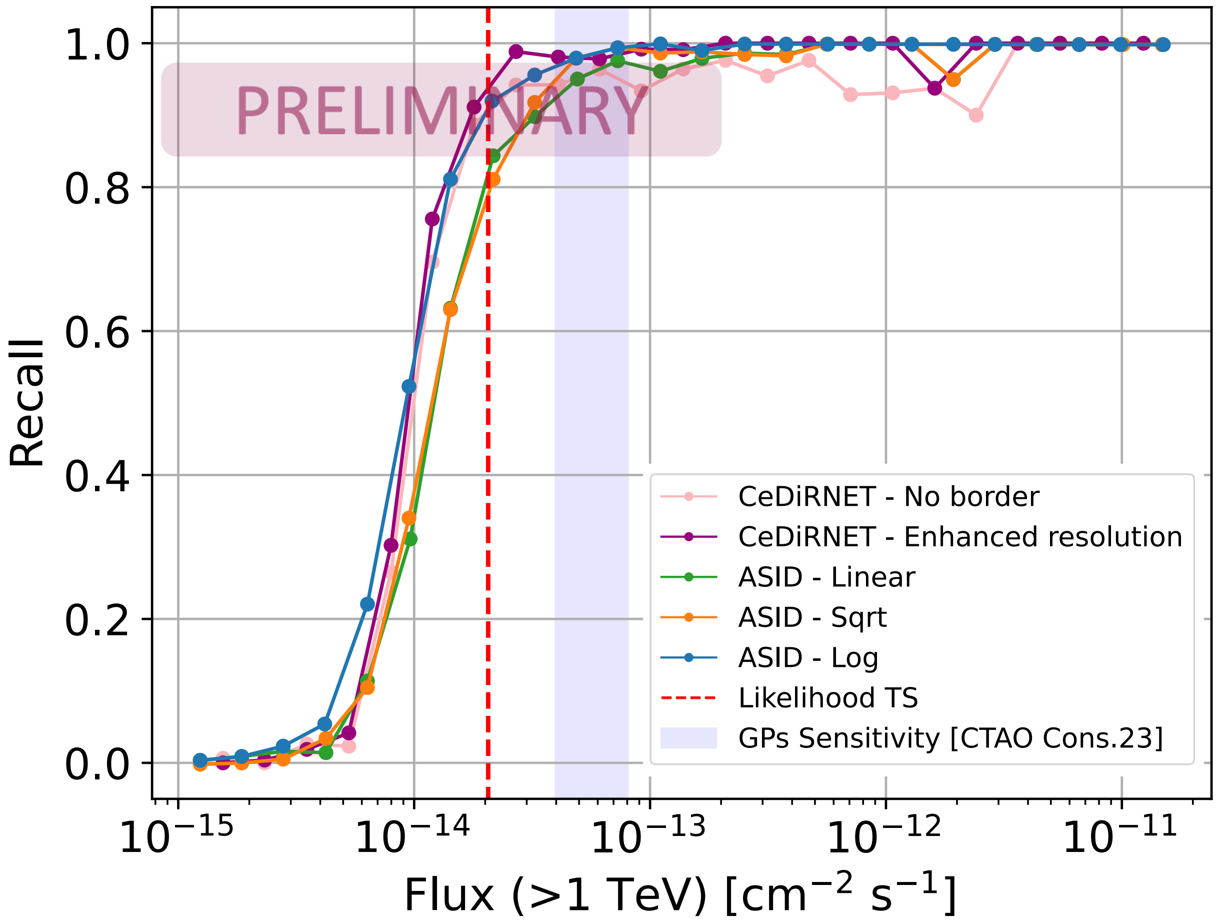
Densely regresses directions pointing to the nearest object center



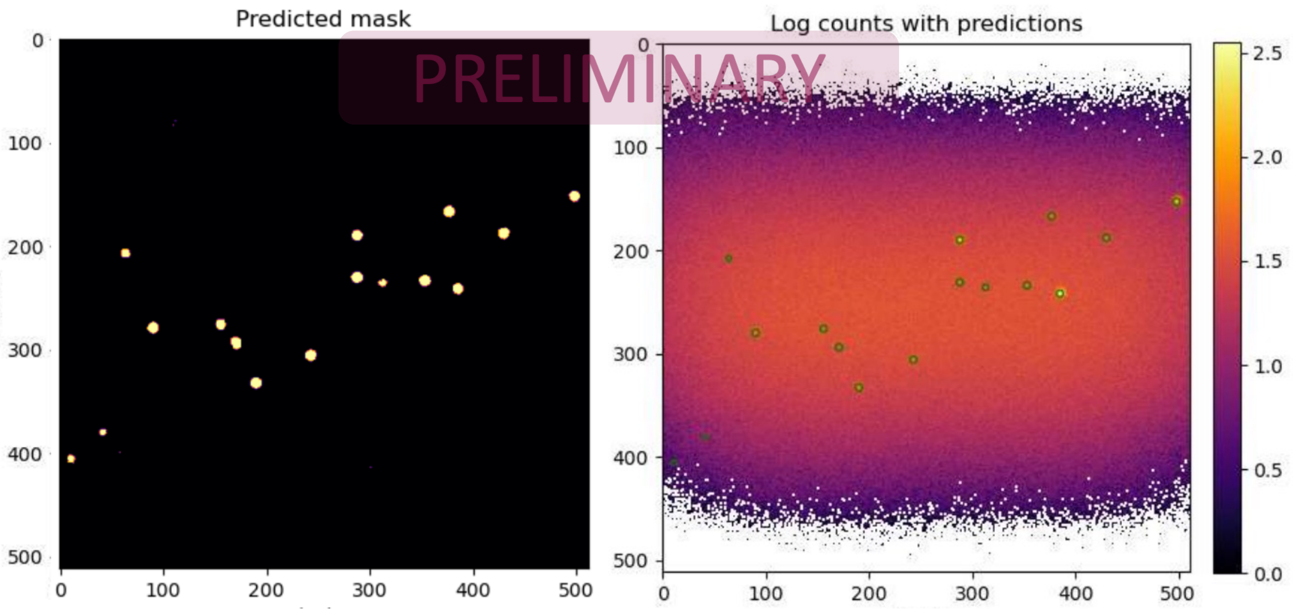
https://github.com/vicoslab/CeDiRNet

The localization network learns fully domain-agnostic

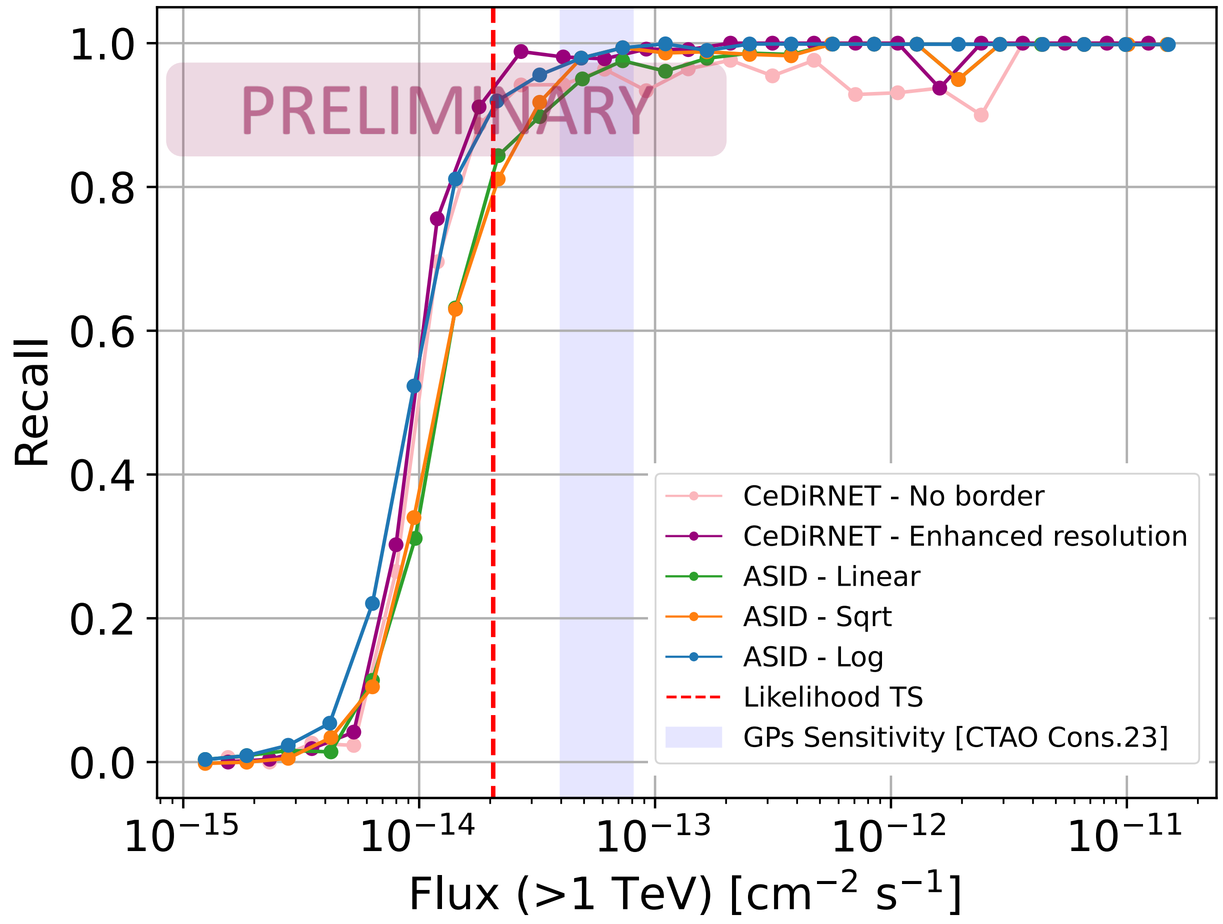
CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY

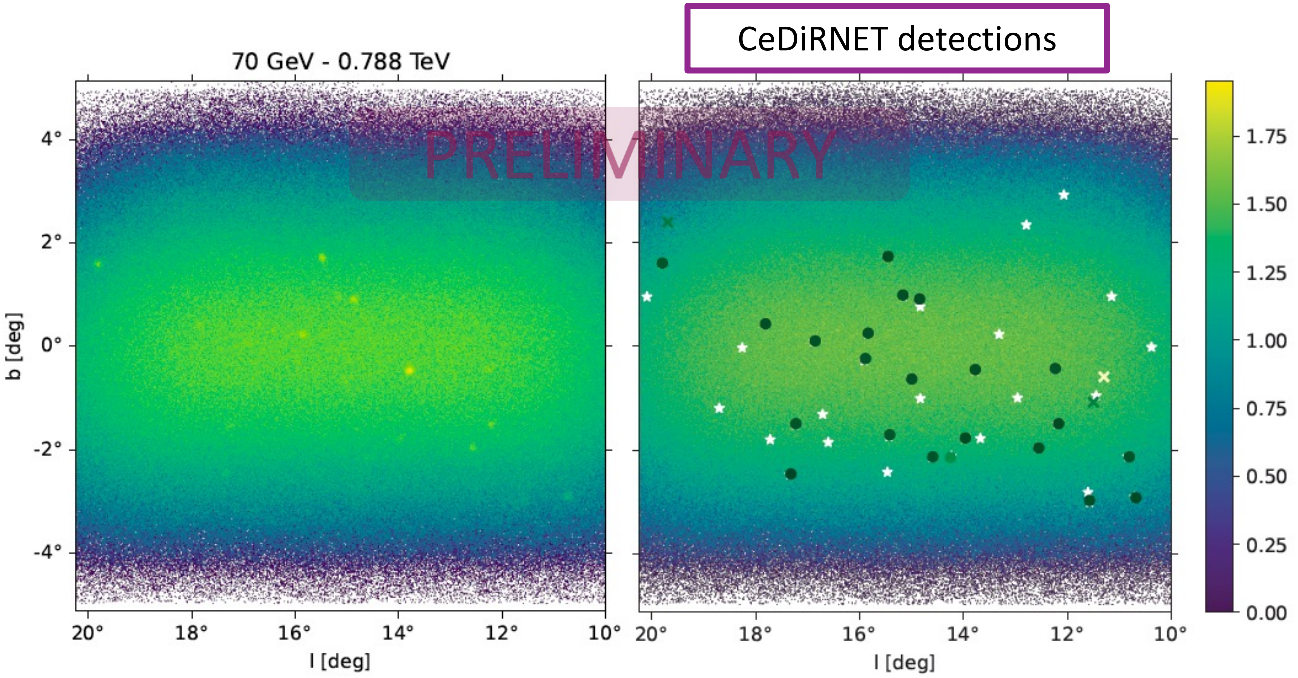


ASID detections

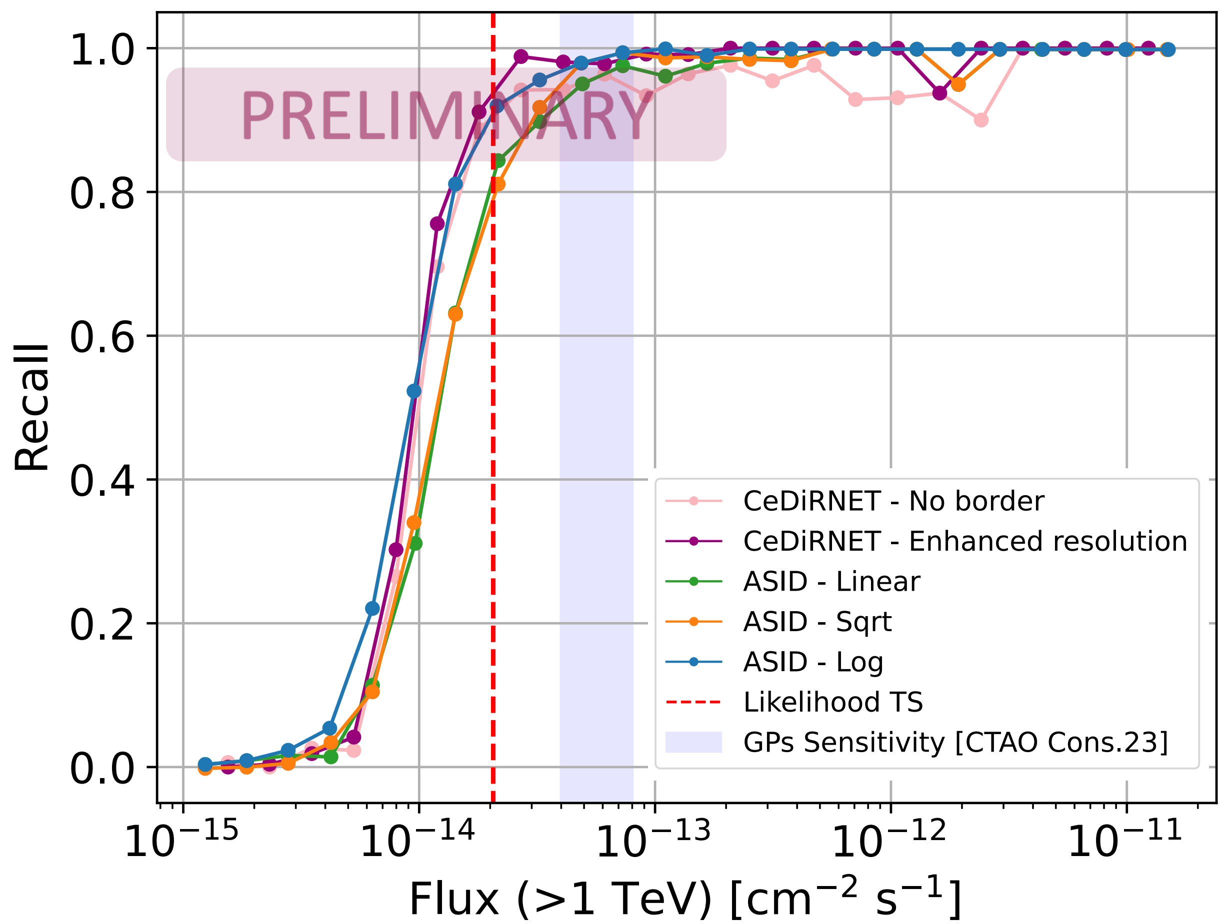


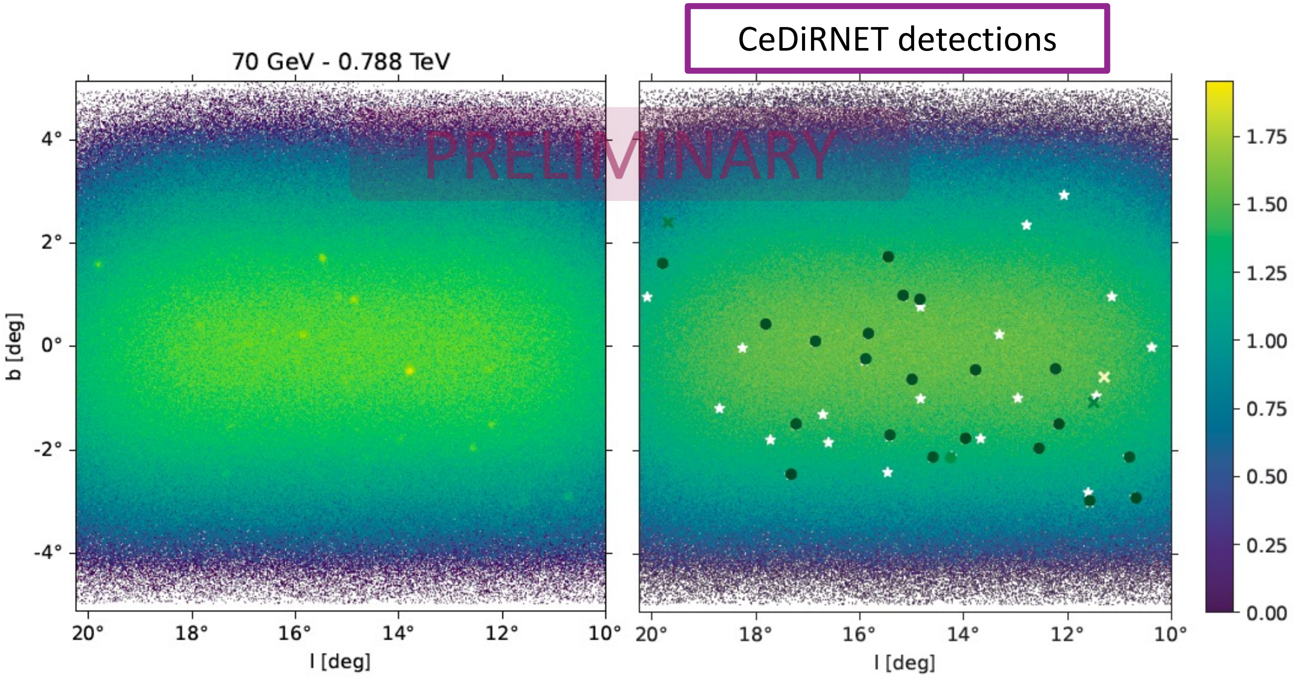
CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY





CTAO: THE UNEXPLORED HIGH-ENERGY GAMMA-RAY SKY





- On going work:
  - Denoising pipeline (U-Net, diffusion model…)
  - Realistic **CTAO** data (across the GP, state-of-the-art models for populations of sources)
  - Detection of **extended & overlapping** sources

MEERLICH: IDENTIFICATION OF SOURCES IN OPTICAL DATA

- Trained and tested with MeerLICHT data
  - 65 cm optical telescope with FoV = 2.7 deg^2
  - Images of fields with different source densities:
    1. Omega Cen. globular cluster,
    2. Fornax galaxy cluster
    3. “Empty” field
  - Each field is divided into 1681 patches of 256 × 256 pixels (total of 5043 patches)
- Automatic rejection of CR contaminants, satellite trails…

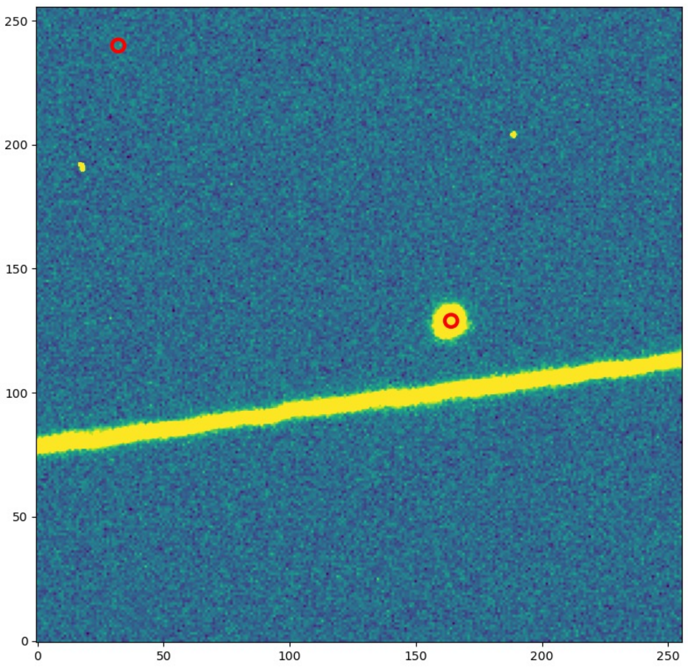
Can we extend the pipeline to other wavelengths?

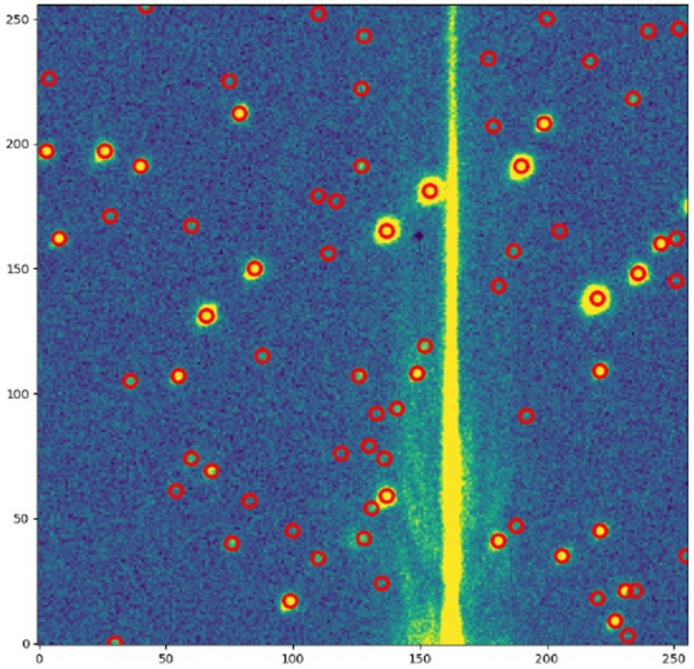
MEERLICH: IDENTIFICATION OF SOURCES IN OPTICAL DATA

- Trained and tested with MeerLICHT data
  - 65 cm optical telescope with FoV = 2.7 deg^2
  - Images of fields with different source densities:
    1. Omega Cen. globular cluster,
    2. Fornax galaxy cluster
    3. “Empty” field
  - Each field is divided into 1681 patches of 256 × 256 pixels (total of 5043 patches)
- Automatic rejection of CR contaminants, satellite trails…

Can we extend the pipeline to other wavelengths?

ASID-Light: Fast Optical Source Localization via U-Net and Laplacian of Gaussian
F. Stoppa *et. al.*, A&A (A109, 2022), [arXiv: 2202.00489]





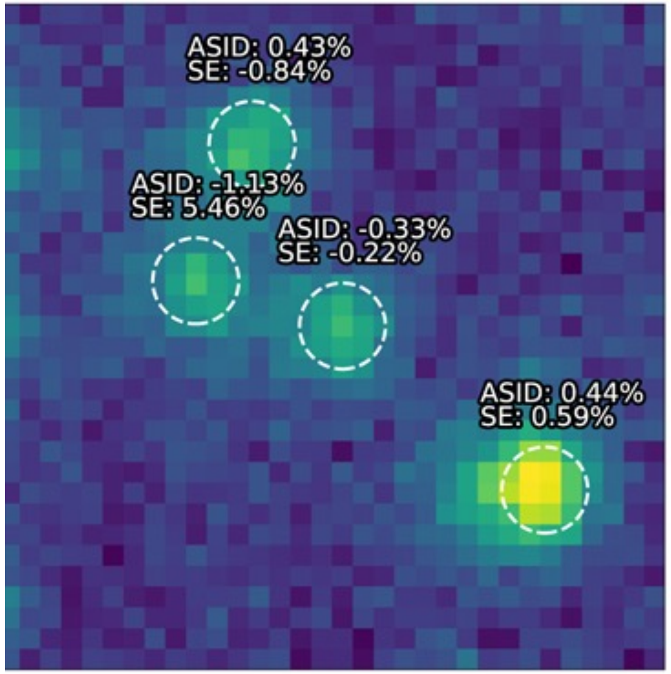
MEERLICH: IDENTIFICATION OF SOURCES IN OPTICAL DATA

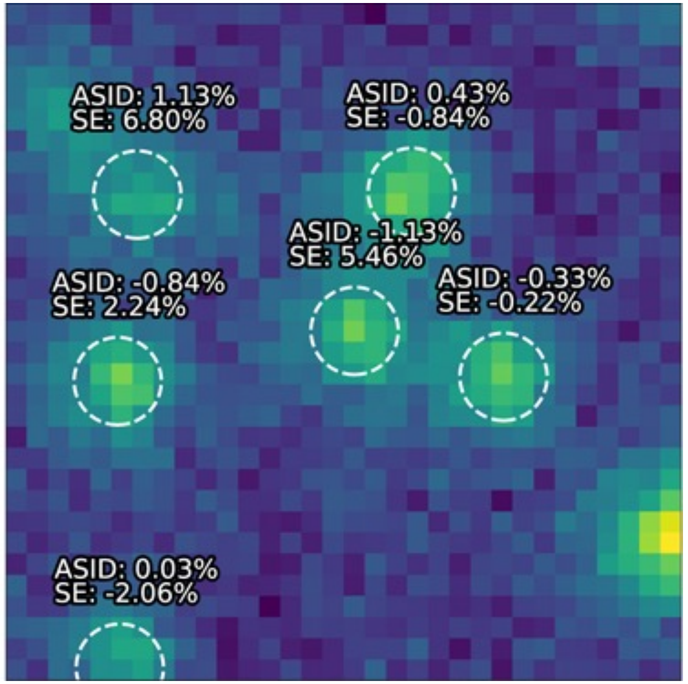
- Once localized, estimate flux with uncertainties (single band image cutout)

- Performs better in crowded field compared to source extractor; well-calibrated uncertainty

ASID-FE: Flux Estimation & Uncertainty Characterization
F. Stoppa *et.al.*, A&A (A108, 2023), [arXiv: 2305.14495]

Predicted flux percentage error



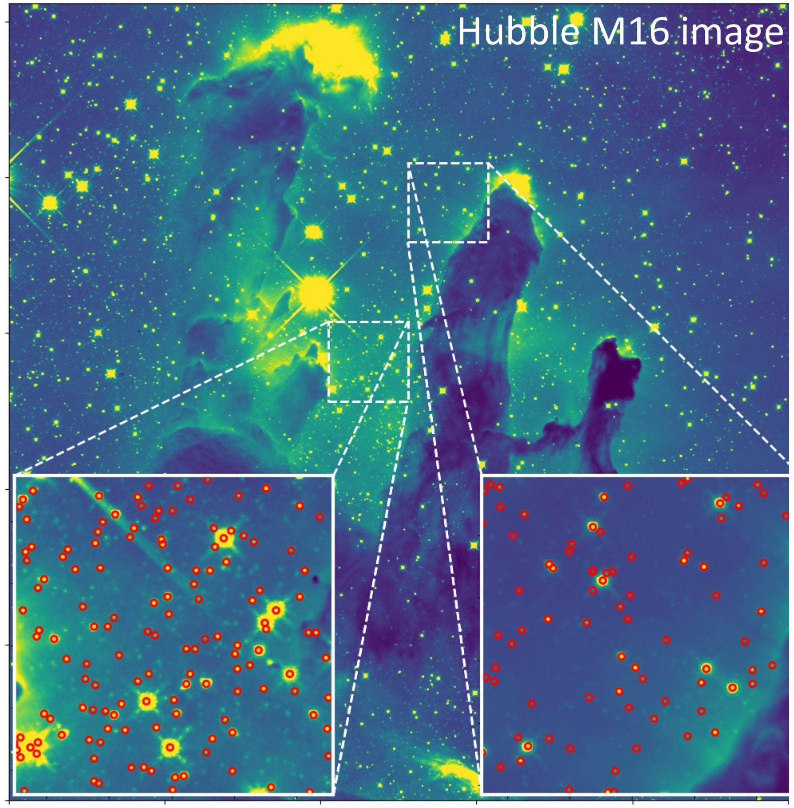


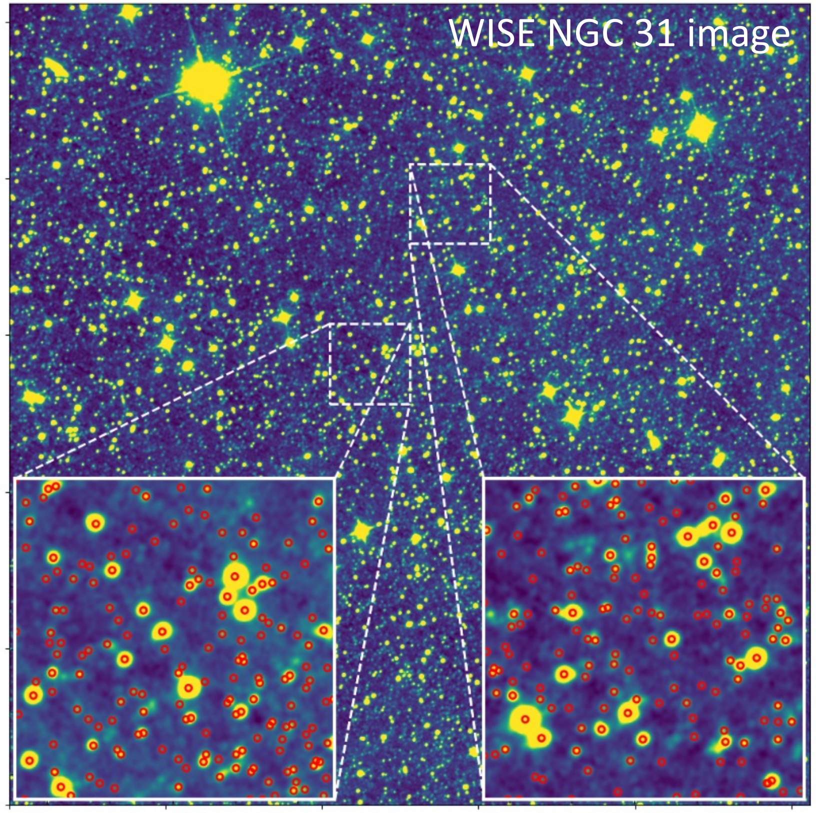
MEERLICH: IDENTIFICATION OF SOURCES IN OPTICAL DATA

ASID-Light [arXiv: 2202.00489]
ASID-FE [arXiv: 2305.14495]

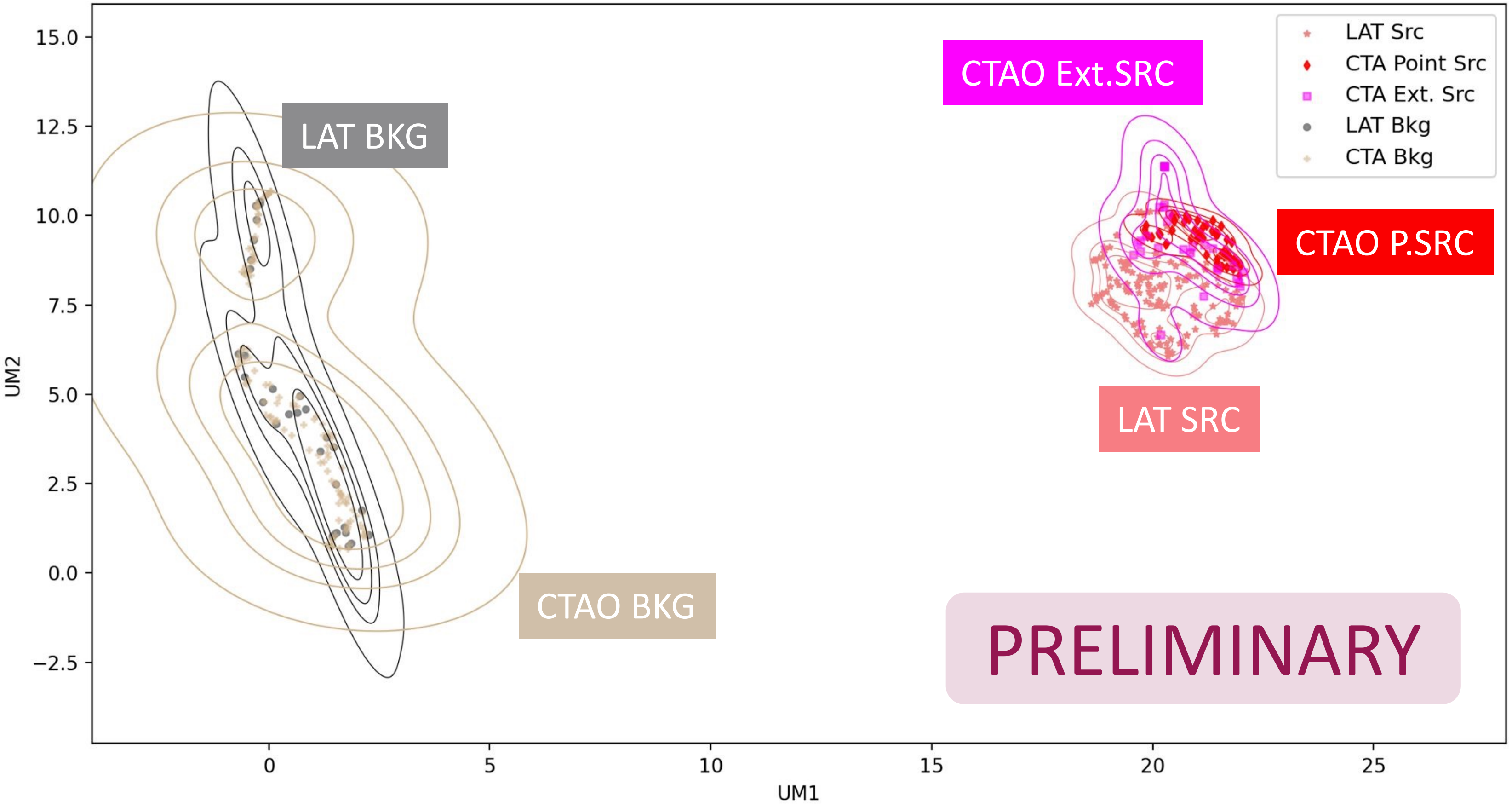
- Trained and tested with MeerLICHT data
- Try **transfer learning** with Hubble data
  - Hubble PSF: 0.11 arcsec
  - MeerLICHT telescope PSF: 2-3 arcssec
- Try **transfer learning** with WISE infrared data

First hints: is it possible to build a foundational model for source detection across wavelengths?



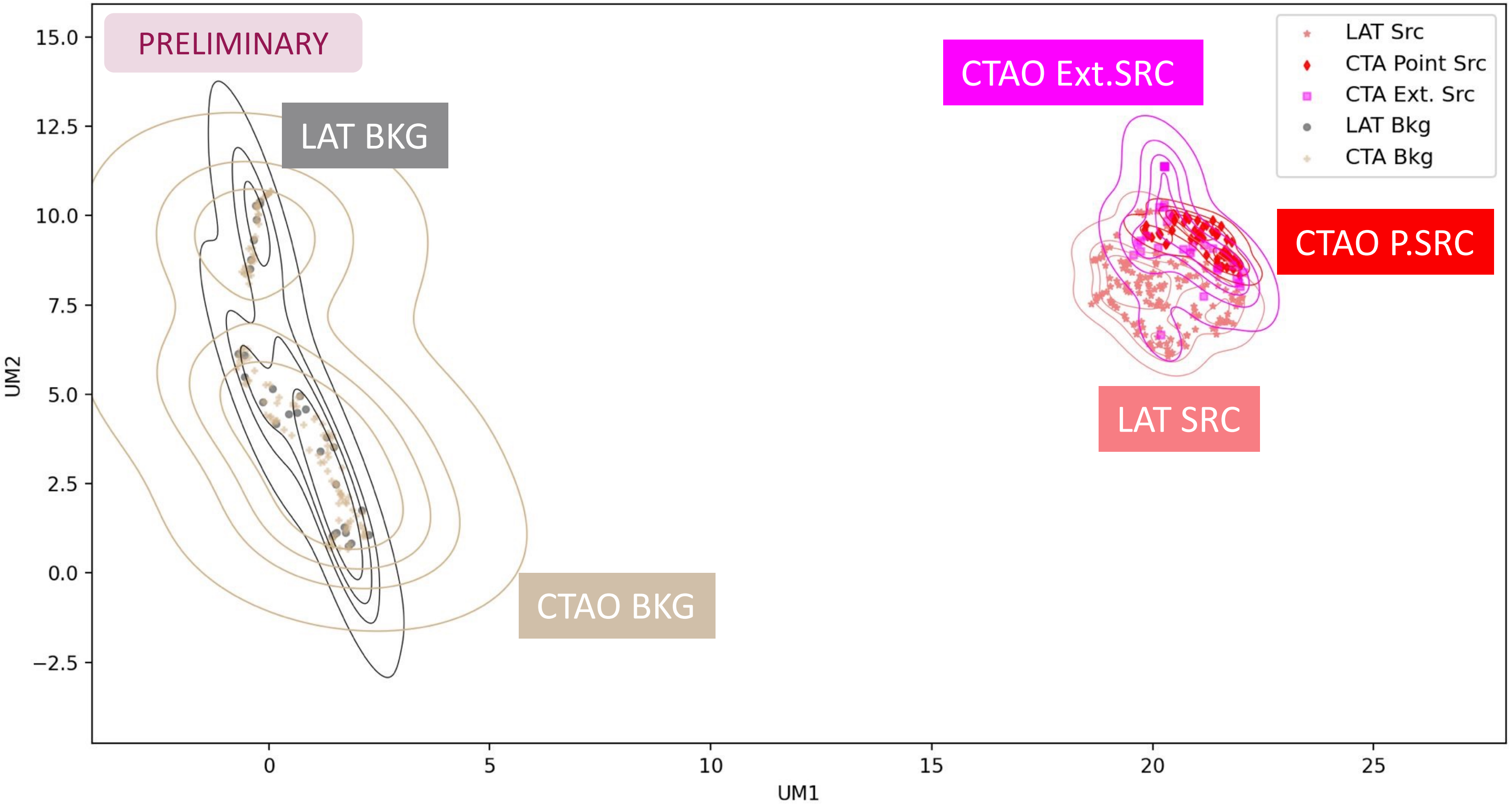


TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION



- Clear distinction between:
  - Gamma-ray sources
  - Gamma-ray backgrounds

TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION



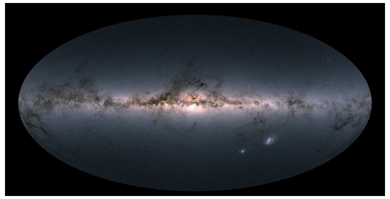
- Clear distinction between:
    - Gamma-ray sources
    - Gamma-ray backgrounds

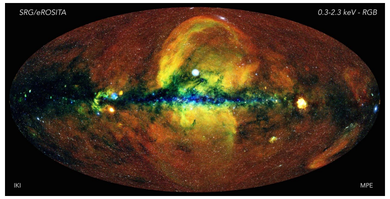
Next steps:

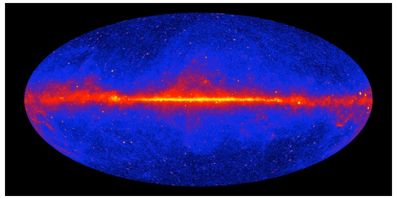
- Include datapoints from other wavelengths in latent space
- Tokenization

TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION

Input Data







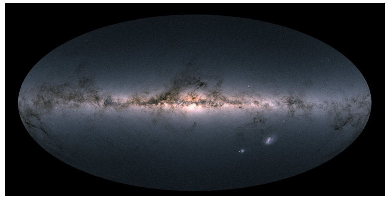
→ Detection + Localization

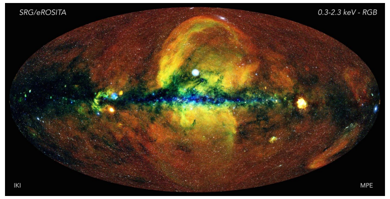
ASID-LON
ASID-LAT

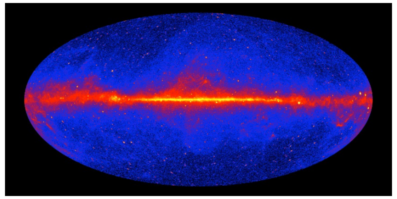
ASID-LIGHT

TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION

Input Data







Detection + Localization

ASID-LON
ASID-LAT

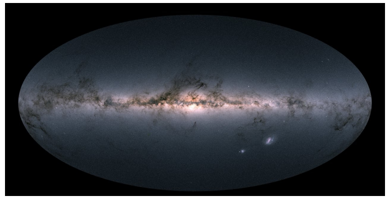
ASID-LIGHT

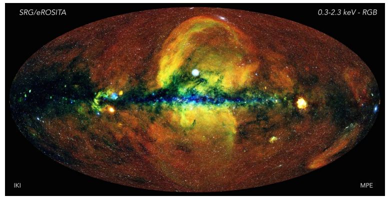
Characterization of Individual Sources: Catalog

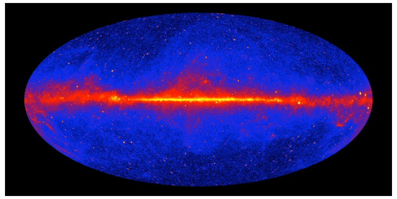
- Flux Prediction ASID-Flux ASID-FE
- Class Labels/Probabilities ASID-SProba
- Source Extension
- Location uncertainty
- ….

TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION

Input Data







Can we remove background?
Could this help next tasks?

Detection + Localization

ASID-LON
ASID-LAT

ASID-LIGHT

Can we generate high-quality data?

Characterization of Individual Sources: Catalog

- Flux Prediction ASID-Flux ASID-FE
- Class Labels/Probabilities ASID-SProba
- Source Extension
- Location uncertainty
- ….

AI 4 SCIENCE
LJUBLJANA, SLOVENIA

International Conference AI for SCIENCE 2025 joined by SMASHing Conference

https://ai4science.si/

22-26 September 2025
Ljubljana, Slovenia

Important dates for thematic tracks (excluding Discovery Science conference):

- **24 June, 2025**
  Paper/abstract submission deadline
- **1 July, 2025**
  Notification of acceptance

28th Discovery Science Conference +

- AI & Digital Humanities
- AI & Environmental Science
- AI & Life Sciences
- AI & Physics

Co-funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101081355.

I FEEL SLOVENIA

Thanks for your attention!

SMASH
machine learning for science and humanities postdoctoral program

Co-funded by the European Union

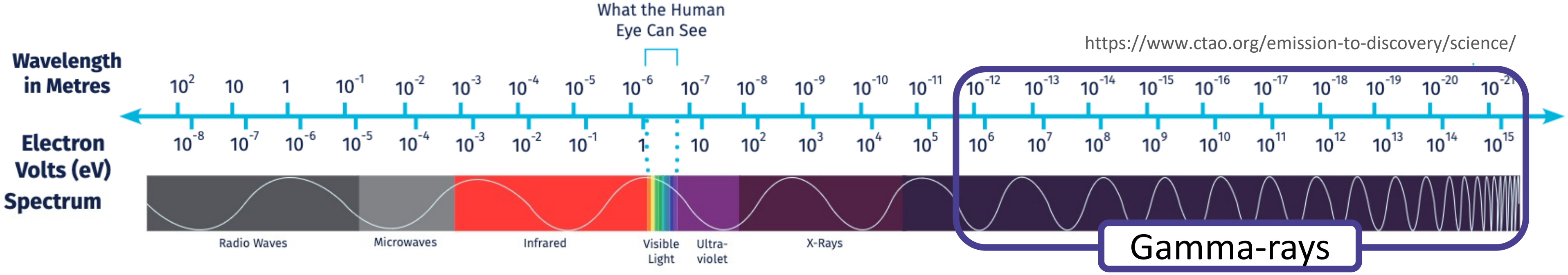
This project has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101081355.

I FEEL SLOVENIA

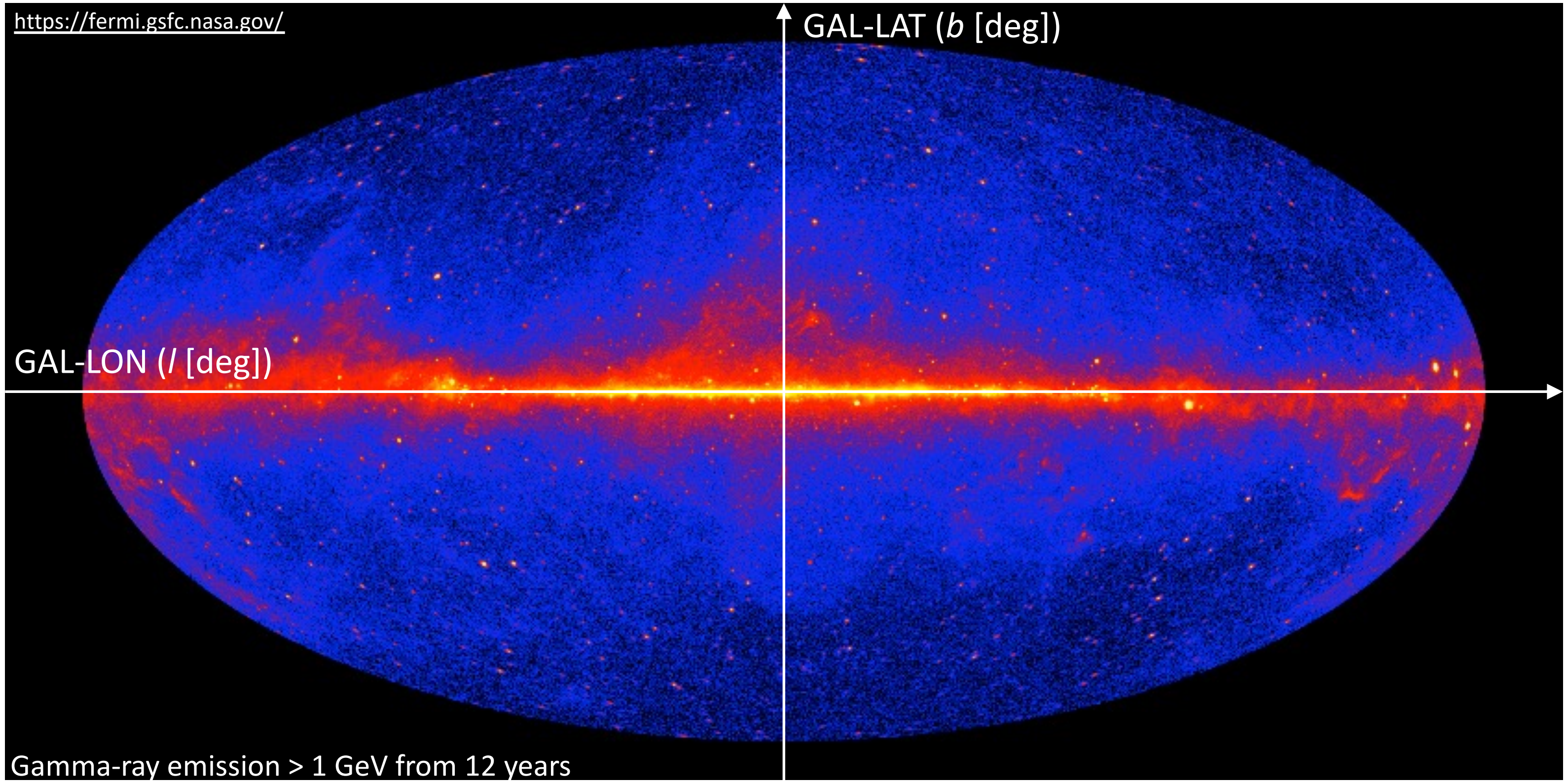
Back up slides

SMASH machine learning for science and humanities postdoctoral program

THE GAMMA-RAY SKY



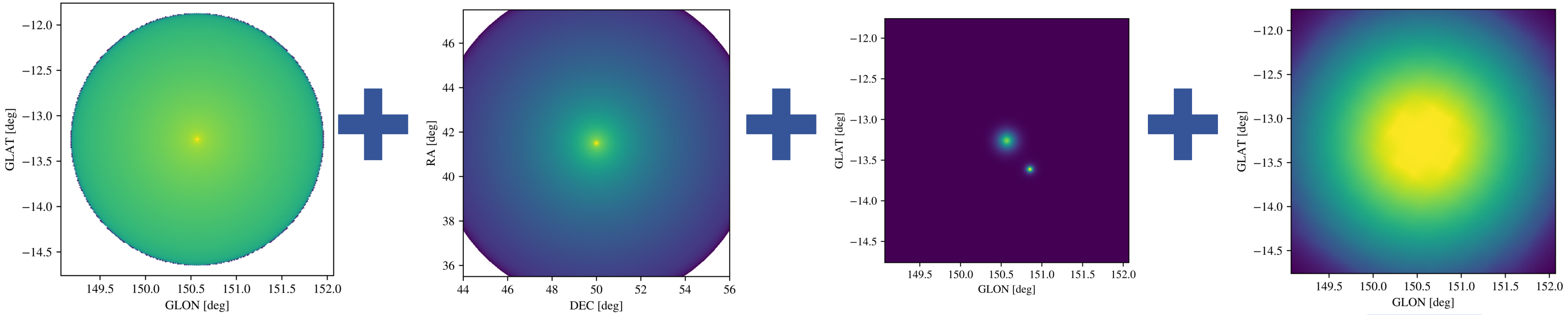
Gamma-rays are electromagnetic radiation > 100 keV



- Gamma-rays travel in straight lines, allowing to determine their origin
- Originated from non-thermal processes
- Gamma-ray sources:
  - Galactic
  - Extragalactic
  - Exotic (?)

GAMMA-RAY STANDARD ANALYSIS

- Includes all expected gamma-ray sources: Target + Astrophysical Backgrounds (BKG) + BKG from Instrument Response Function (IRFs)



Target emission

BKG astrophysical emission

BKG astrophysical emission

BKG

[from IRFs]

Most realistic physical scenario

- Use likelihood ratio test to fit the models to the simulated data:

$$\ln\mathcal{L}(\vec{\theta}|D) = \sum_i \tilde{M}_i(\vec{\theta}) - d_i \ln(\tilde{M}_i(\vec{\theta}))$$

Poissonian likelihood for each parameter

$$TS = 2\,\log\left[\frac{\mathcal{L}\left(A_\chi, \hat{\nu}\right)}{\mathcal{L}_{\text{null}}\left(A_\chi = 0, \hat{\nu}\right)}\right]$$

- $TS < 25$ ➡ No signal

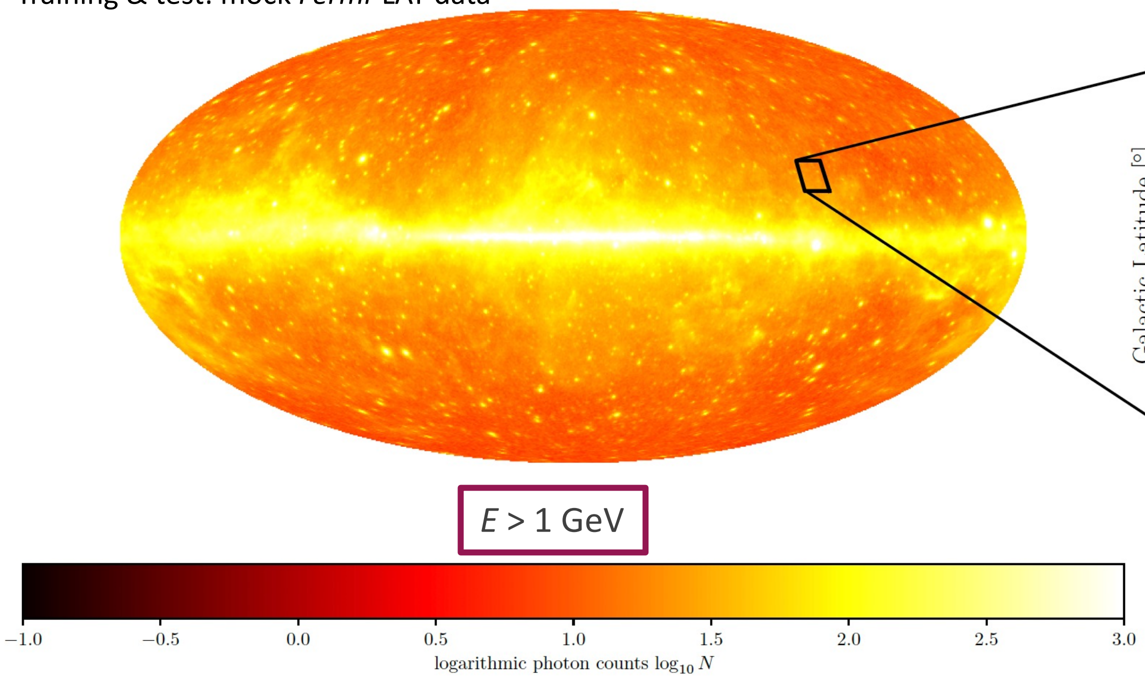
ML TO DETECT FAINT GAMMA-RAY SOURCES

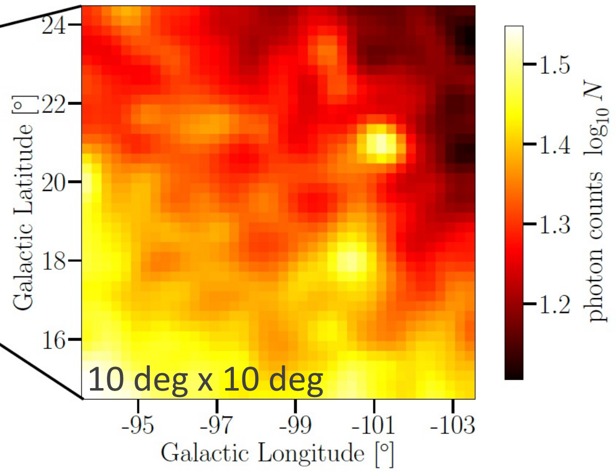
- Standard gamma-ray analysis:

Models for: signal + **background** (with uncertainties) + Likelihood fitting → Lack of knowledge of backgrounds can introduce strong biases

AutoSourceID (ASID) *[Panes+21]*
https://github.com/bapanes/AutoSourceID

Training & test: mock *Fermi*-LAT data





- CNN pipeline based on U-Net algorithms
- Goal: **detect** and classify **point-like** sources

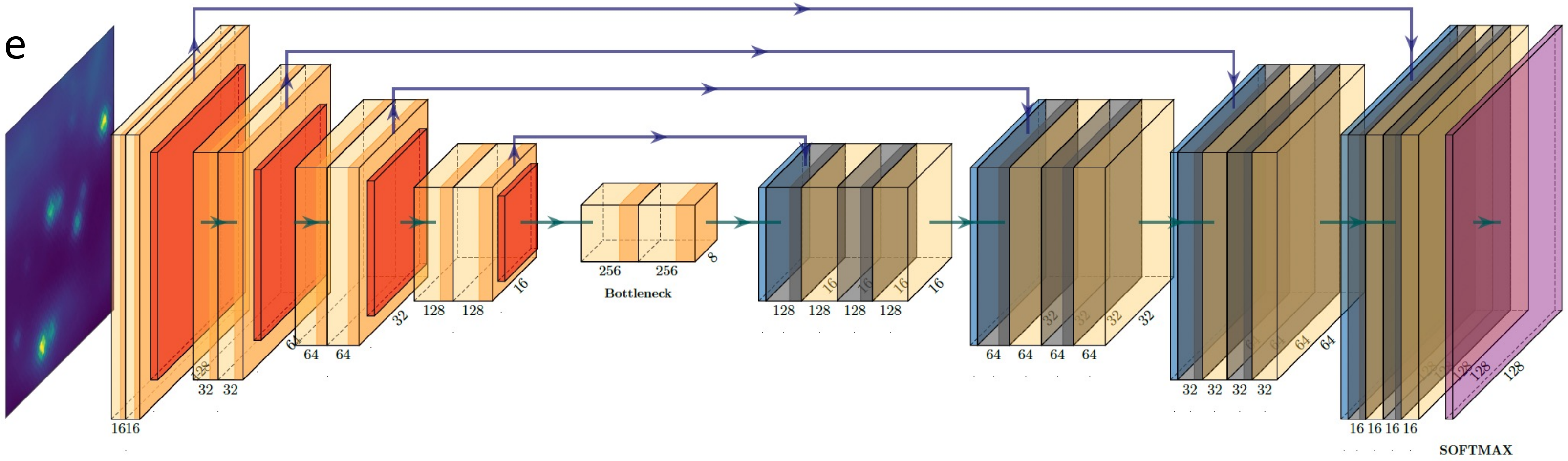
1. Detection: U-NET + clustering algorithm (*k-means*, Centroid-NET)
2. Classification: deep NN to classify different sources (from energy features)

ASID METHODOLY FOR DETECTION OF POINT-LIKE SOURCES

AutoSourceID (ASID) *[Panes+21]*
https://github.com/bapanes/AutoSourceID

ML tool to directly analyse gamma-ray image datasets

U-NET scheme



- Convolutional Neural Network (CNN) pipeline based on U-Net algorithms
- Goal: **detect** (localize) **point-like** sources

Semantic segmentation

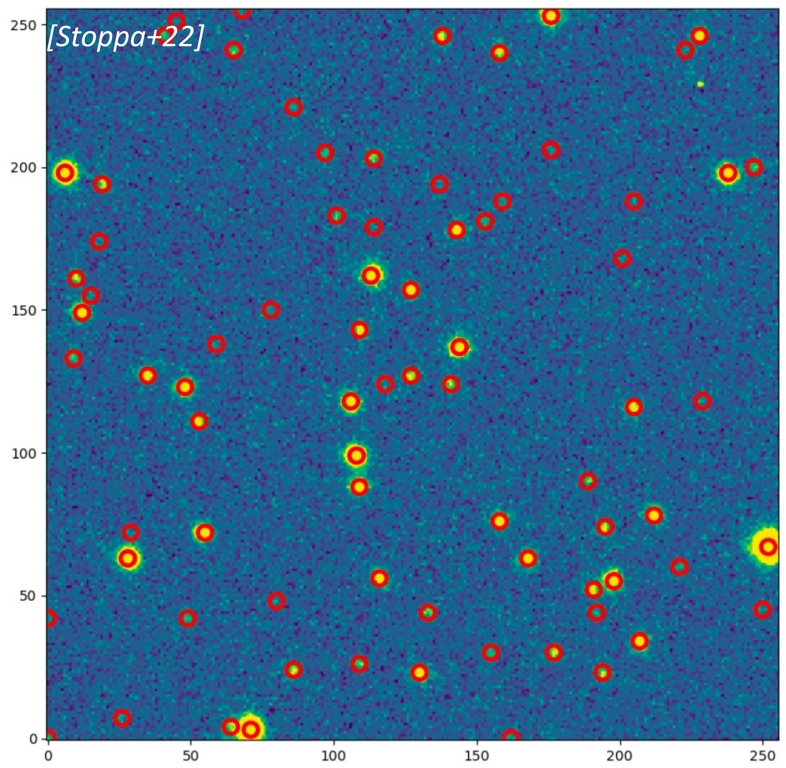
- U-Net produces segmented regions around point sources
- For each input patch there is per-pixel classification (background vs. foreground)
- Label scores: ~1 (for pixels in the region around a point source) and ~0 (otherwise)
- To translate this to positions, apply a clustering algorithm

ML TO DETECT FAINT GAMMA-RAY SOURCES

Localization

Laplacian of Gaussian filter

$$\mathrm{LoG}(x, y; \sigma^2) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right] e^{-\frac{x^2+y^2}{2\sigma^2}}$$

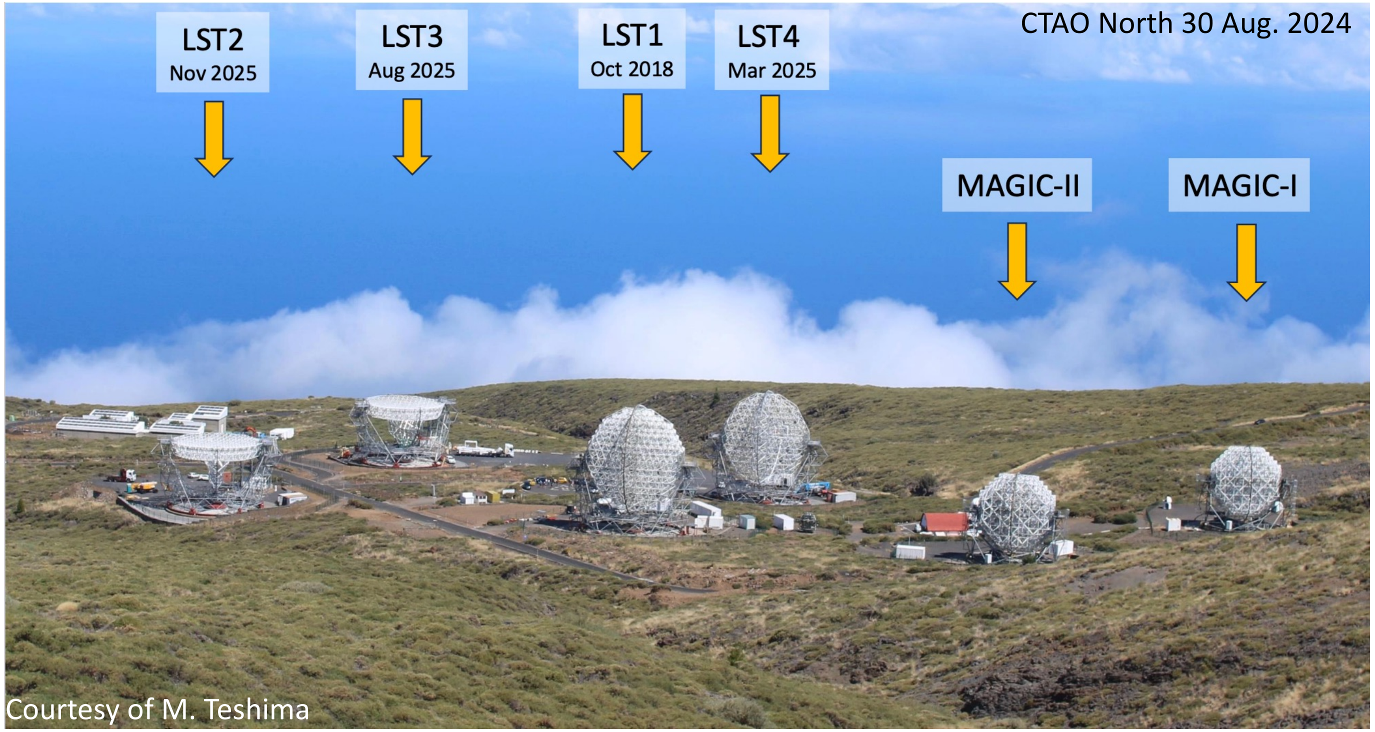


AutoSourceID-Light (ASID-L) *[Stoppa+22]*

https://github.com/FiorenSt/AutoSourceID-Light

ML TO DETECT FAINT GAMMA-RAY SOURCES: CTAO

- Future of Imaging Atmospheric Cherenkov Telescopes for VHE gamma-ray astronomy
- 2 arrays: Northern Array (La Palma, Spain) and Southern Array (Paranal, Chile)
- First LST already in operations!



https://www.cta-observatory.org/

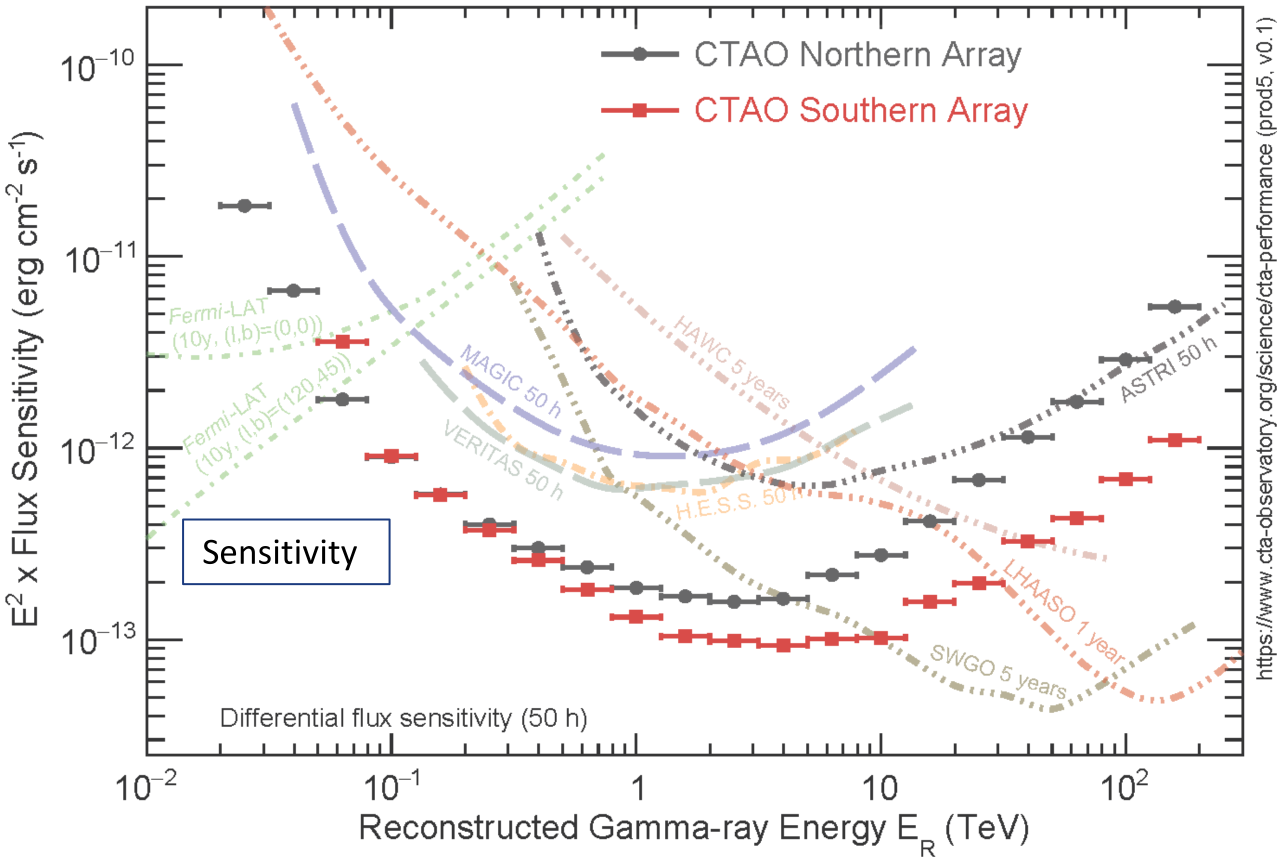
SST	MST	LST	Energy range
5 - 300 TeV	150 GeV - 5 TeV	20 - 150 GeV	20 GeV - 300 TeV
$D_{\emptyset}$ = 4.3m	$D_{\emptyset}$ = 11.5m	$D_{\emptyset}$ = 23m	

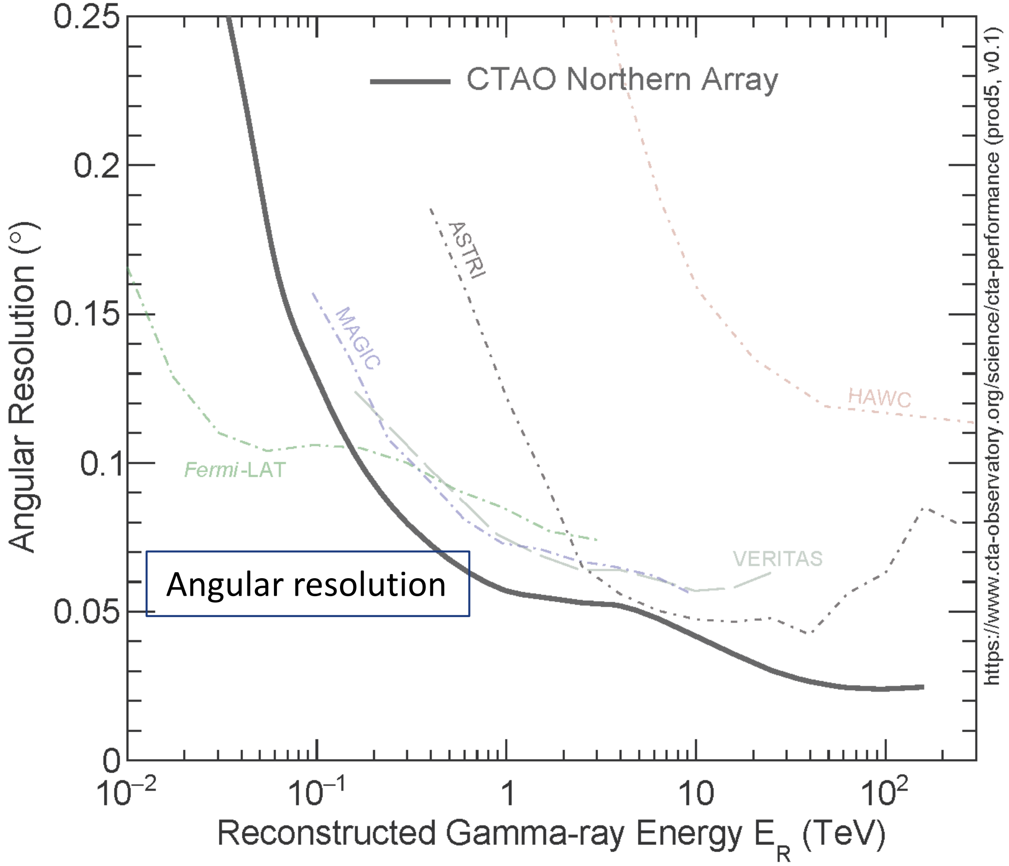
ML TO DETECT FAINT GAMMA-RAY SOURCES: CTAO

- We need to prepare for the data & analysis: we will for sure detect **fainter sources, ML (U-Net) can help**

Preliminary Performance Capabilities of the Alpha Configuration

https://www.ctao.org/for-scientists/performance/





ABOUT THE CTAO DATA

- We generate the data (telescope is under construction phase)
- We only have one sky
- Data is generated according to state-of-the-art physical models on the population of the different kind of sources that we know
- We only have simulated data (we use most updated characterization of the detector to make it as realistic as possible)
- To increase number of data, we take advantage of the uncertainty in populations, making realizations of the sky given the models
- This means we have as training data as many as we need (reasonable in space and time)

ABOUT THE DATA: PRELIMINARY TOY SIMULATIONS

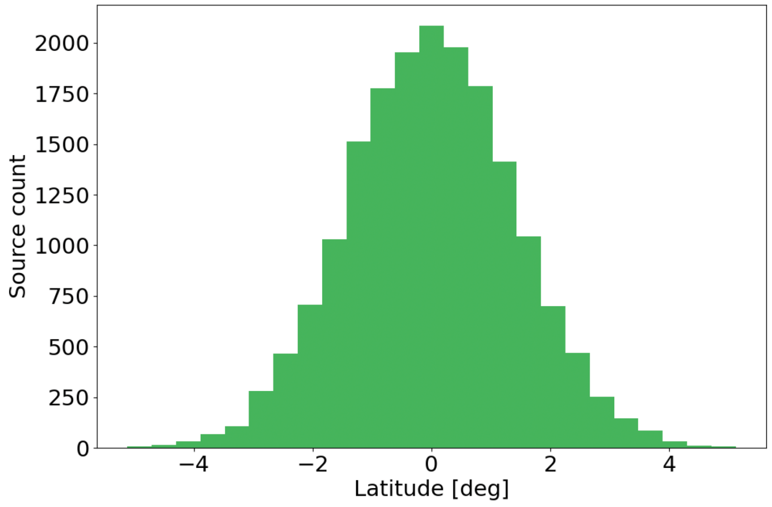
GEOM

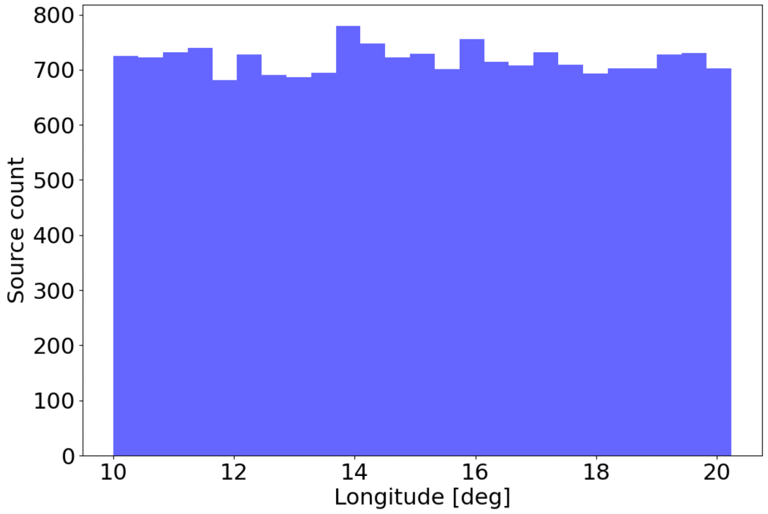
- Observation plan in [*2310.02828*] and the corresponding GPS pointings list from the GPS repository (non-equilateral double-row pattern and duration 30 min per pointing)
- IRFs prod5-v0.1 (South_z20_50h)
- The RoI centered on (15.12 deg, 0 deg) in gal. coords. and 10.24 x 10.24 deg^2
- The spatial binning 0.02deg
- Energy bins range from 0.07 TeV to 100 TeV logarithmically binned in 3 (70 GeV - 0.788 TeV, 0.788 TeV - 8.88 TeV, 8.88 TeV - 100 TeV)
- Always same patch

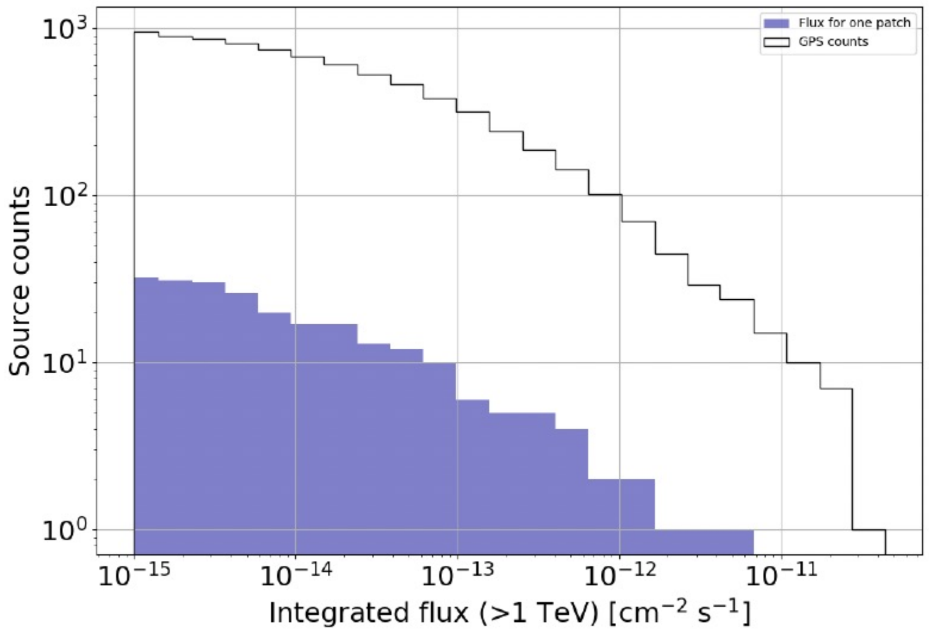
SIMS

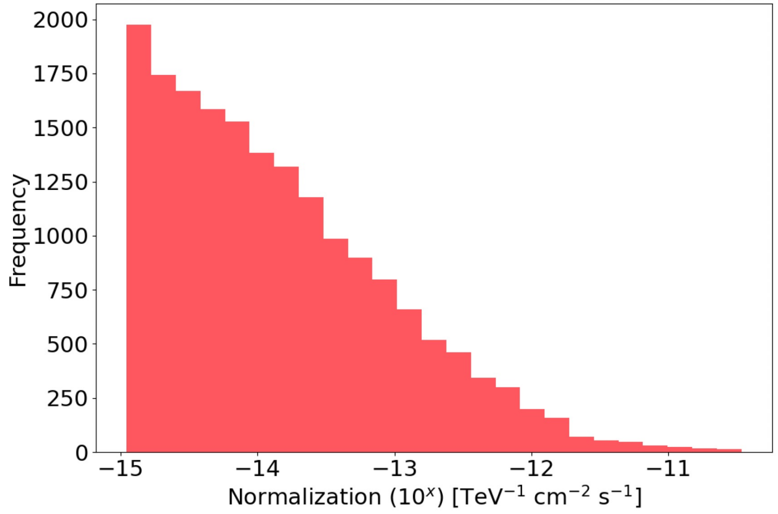
- number of sources randomly between [20, 40]
- $10 < l < 20.24$ deg
- $-5.12 < b < 5.12$ deg, with a variance of 1.4 so most of sources in $|b| < 1$ deg
- Each source follows power-law spectral distribution $dN/dE = K0(E/E0)-\gamma$
- To obtain the distributions of the parameters used GammaCat
- E0 fixed at 1 TeV, and γ has a normal distribution with a mean of 2.2 and variance of 0.05
- K0 was modeled as 10^x TeV−1 cm−2 s−1, where x was adjusted so that the resulting cumulative distribution of integrated source flux above 1 TeV aligned with the same distribution of all model sources from the GPS paper

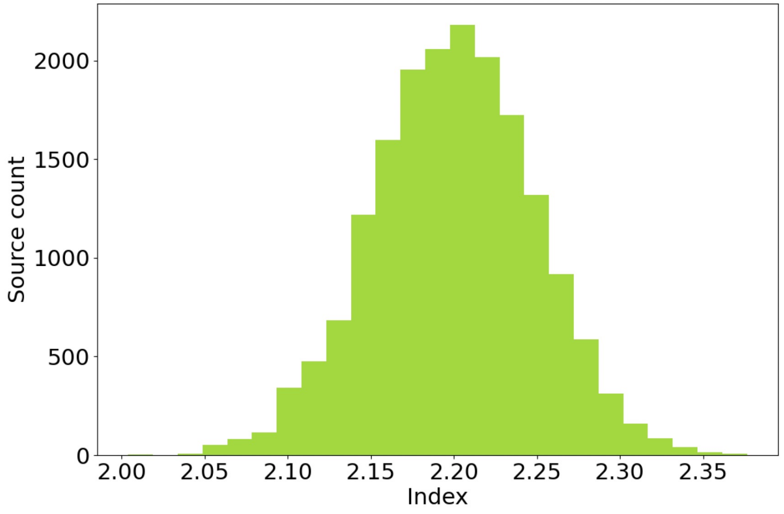
ABOUT THE DATA: PRELIMINARY TOY SIMULATIONS





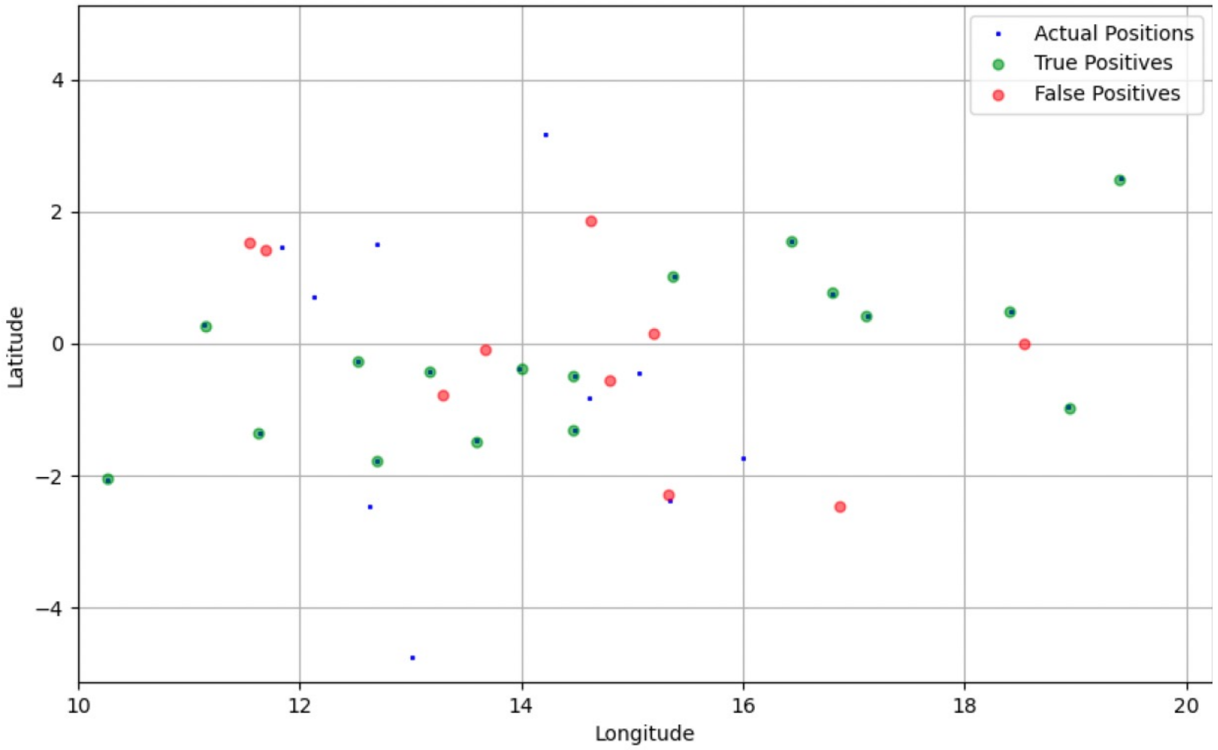


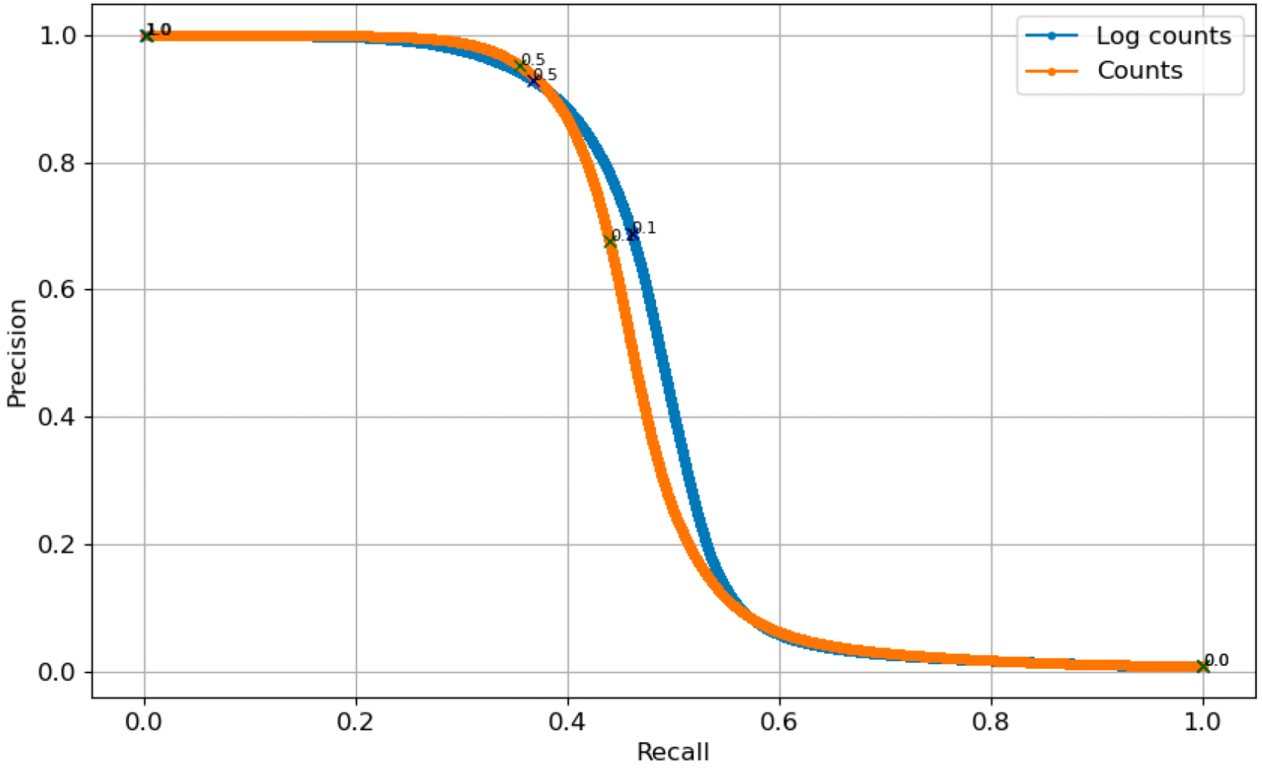




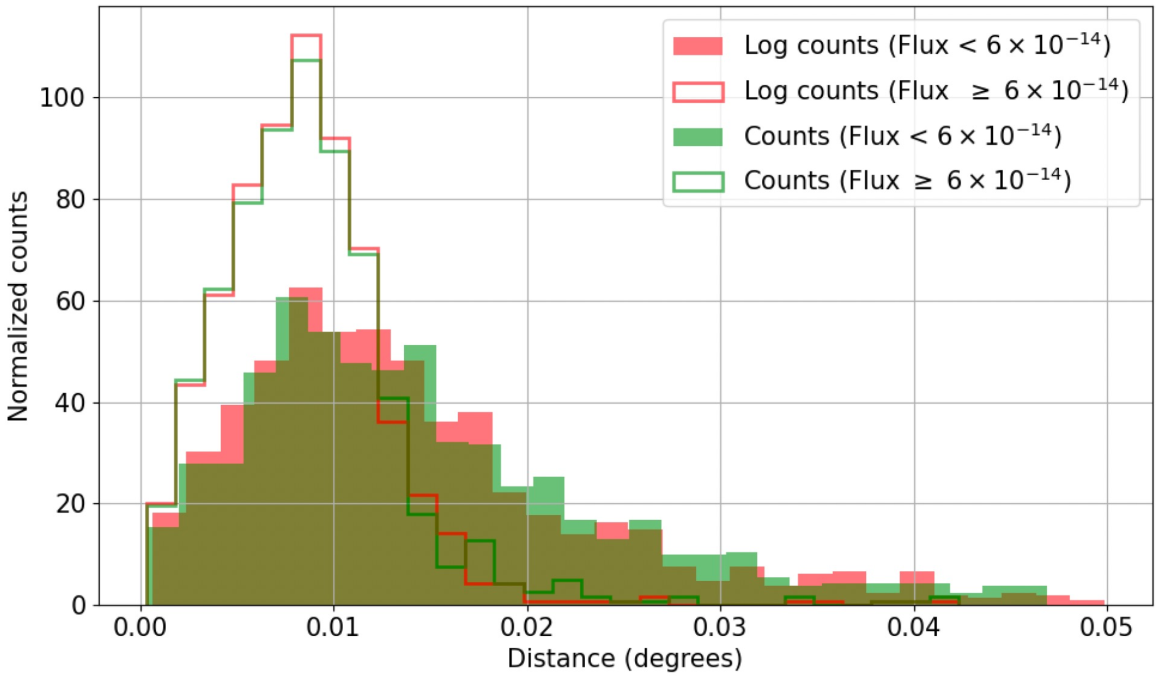
CTAO EXTRA-RESULTS: PRELIMINARY TOY SIMULATIONS

$$\text{Precision} = \frac{TP}{TP + FP} \qquad \text{Recall} = \frac{TP}{TP + FN}$$

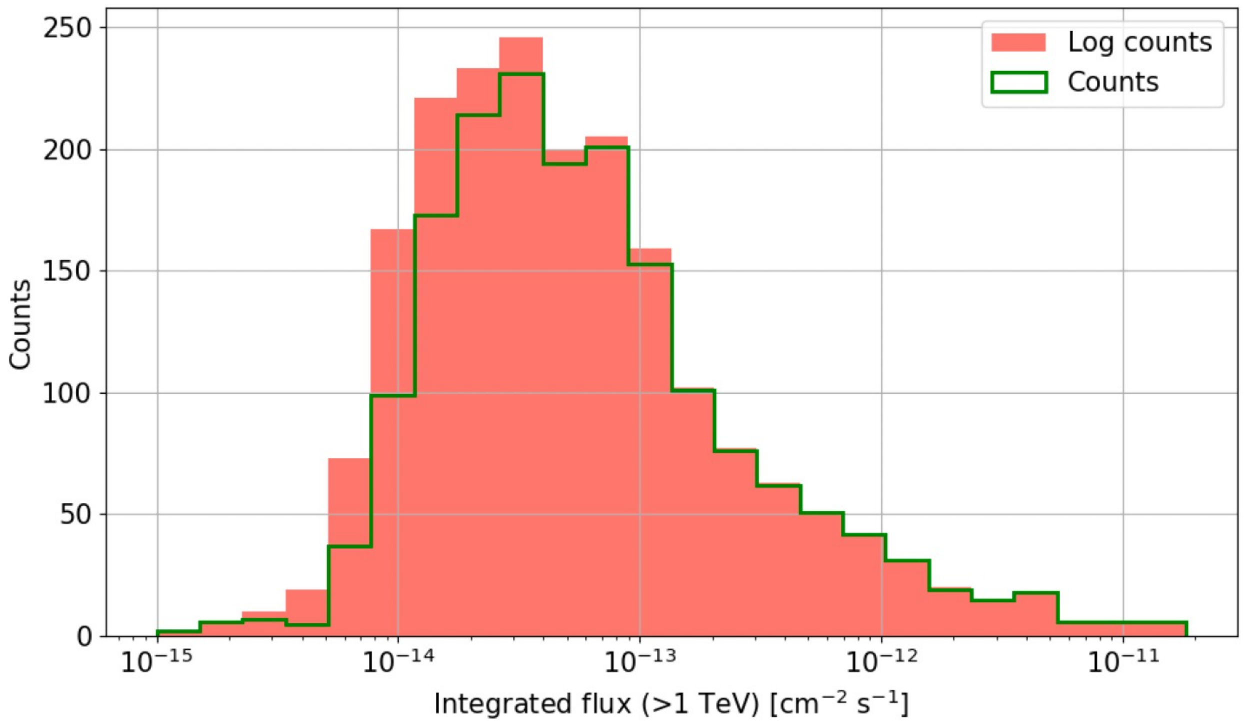




CTAO EXTRA-RESULTS: PRELIMINARY TOY SIMULATIONS



Distribution of separations of predicted sources from true sources below a separation threshold of 0.05◦. Only original counts and log-transformed counts are included and for each we plot for separations of sources with flux below and above the GPS sensitivity

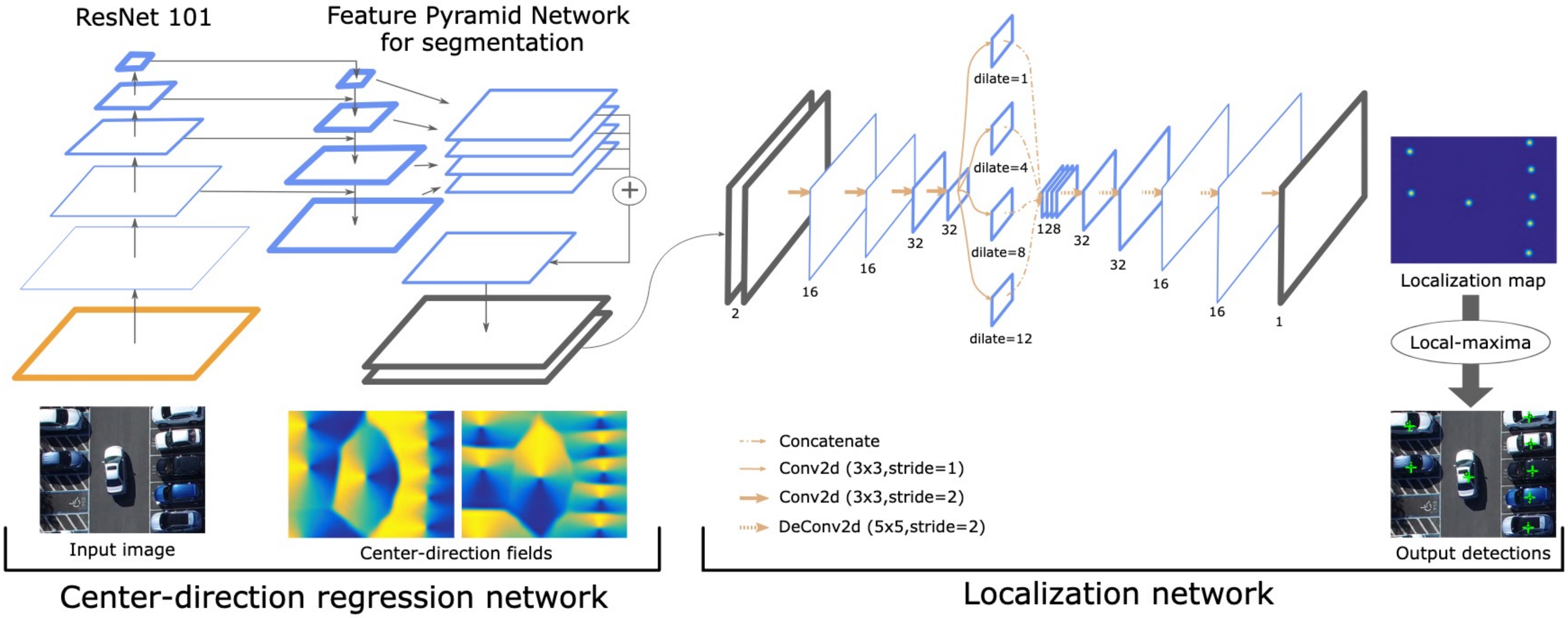


Integrated flux (> 1TeV) distribution of true recovered sources by the algorithm.

CEDIRNET METHODOLY

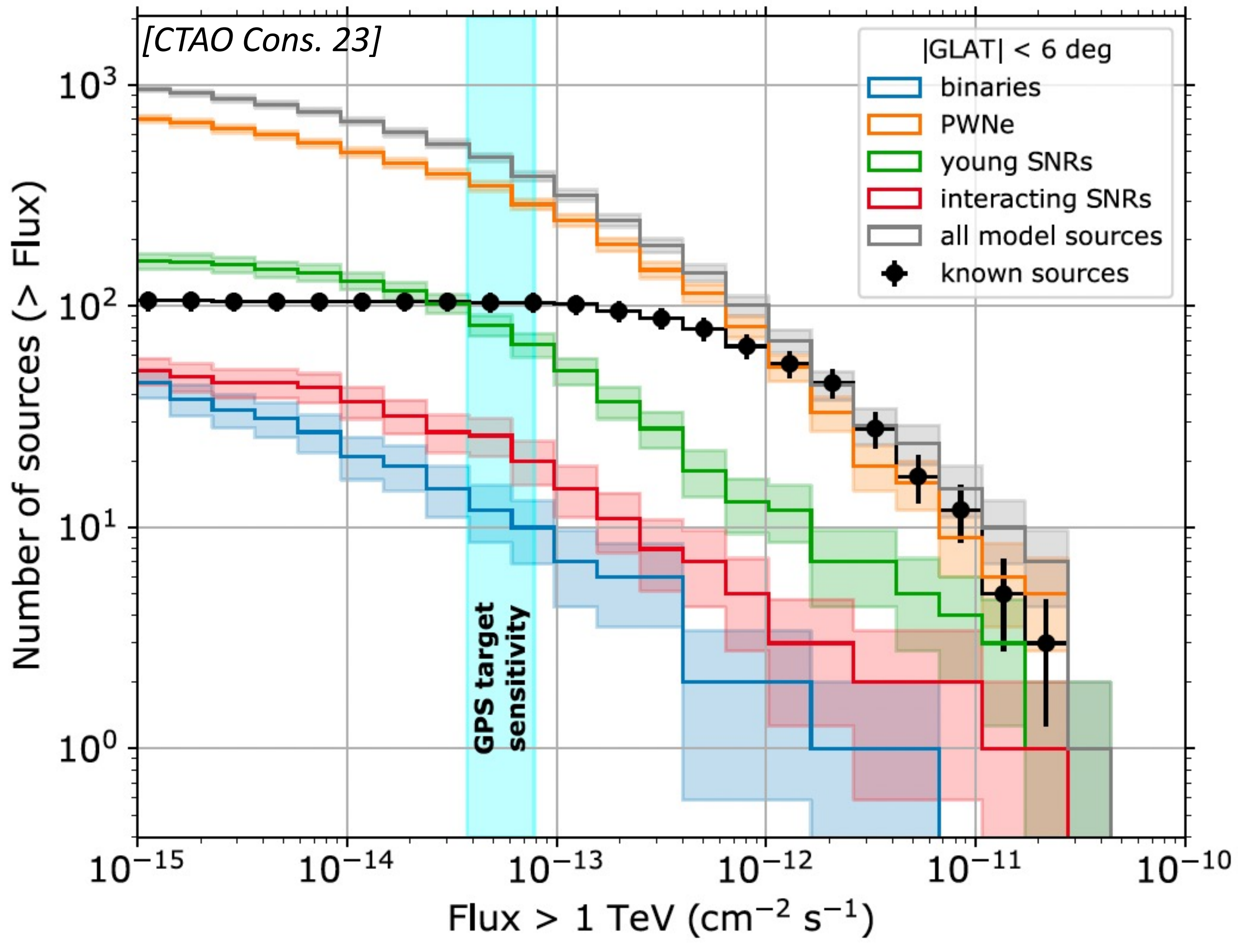
- Algorithm CeDirNet: https://github.com/vicoslab/CeDiRNet

  [Tabernik, Muhovič & Skočaj 24]

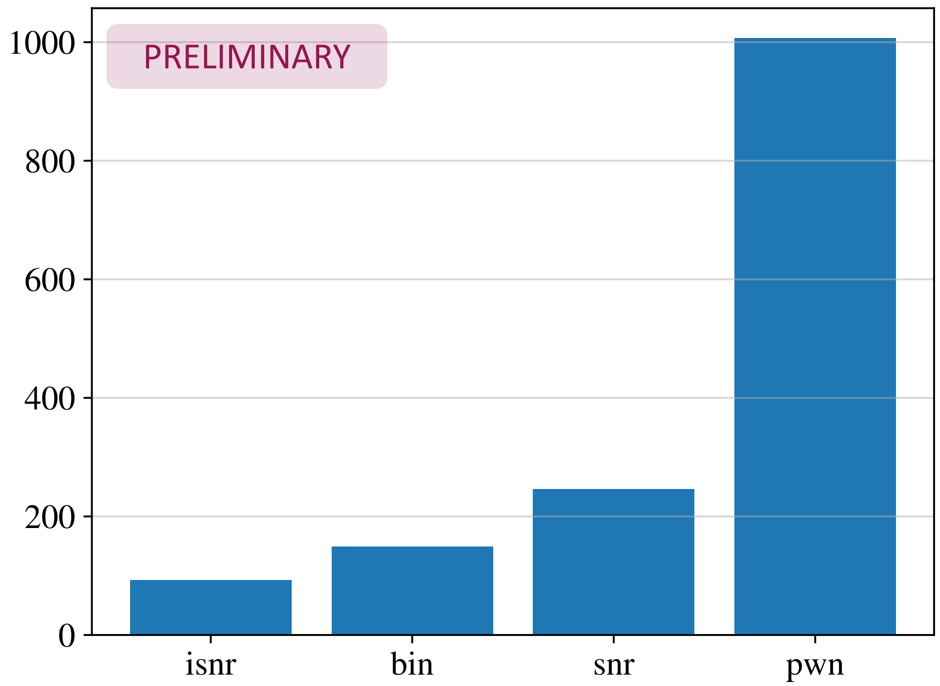


CTAO EXTRA-RESULTS: CTAO GALACTIC PLANE SIMULATION

- Original simulated population on the galactic plane



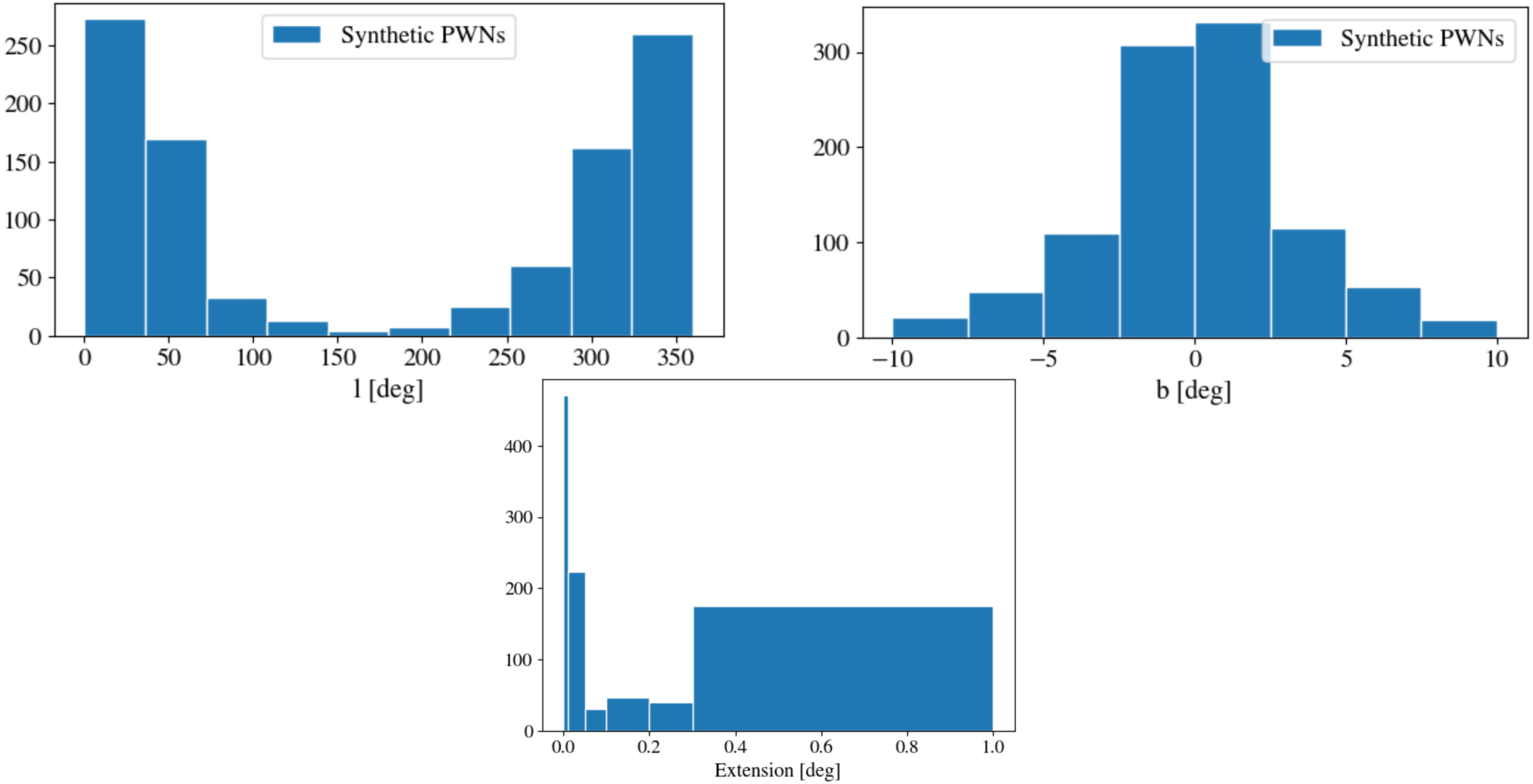
- We need several realizations (simulations) of the GP
- Extract the physical distributions of the sample



- The difference between types is not on the image but on the spectrum (flux vs. energy)

CTAO EXTRA-RESULTS: CTAO GALACTIC PLANE SIMULATION

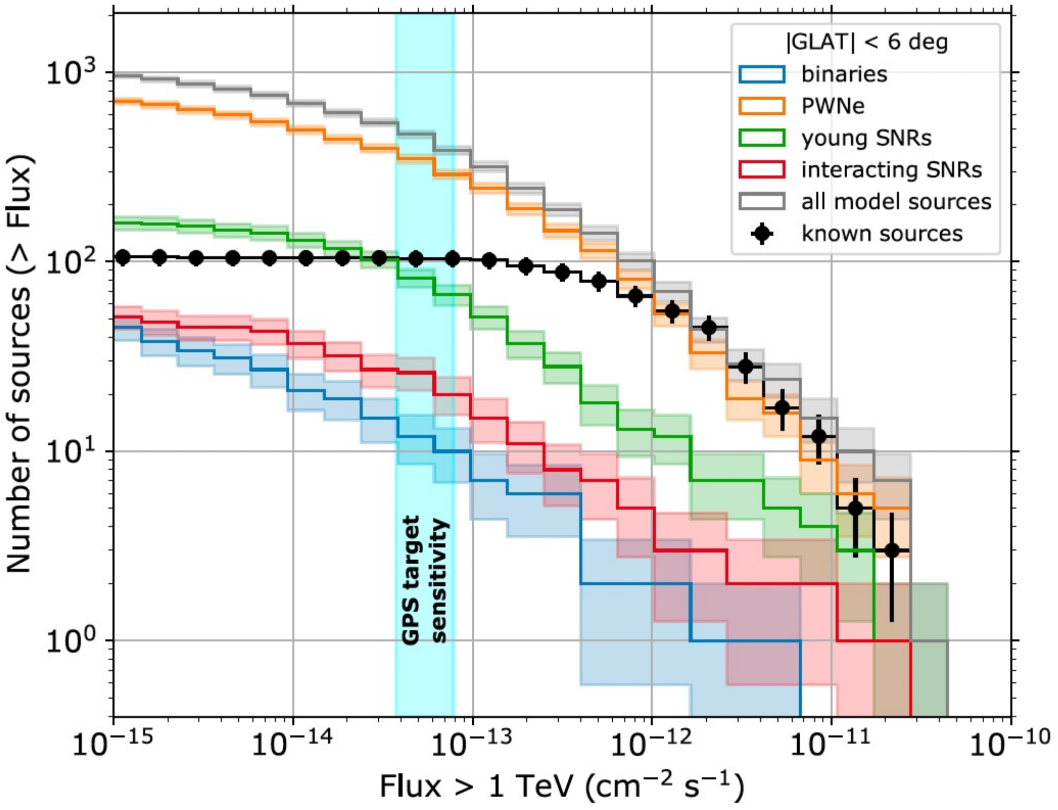
- Original simulated population on the galactic plane



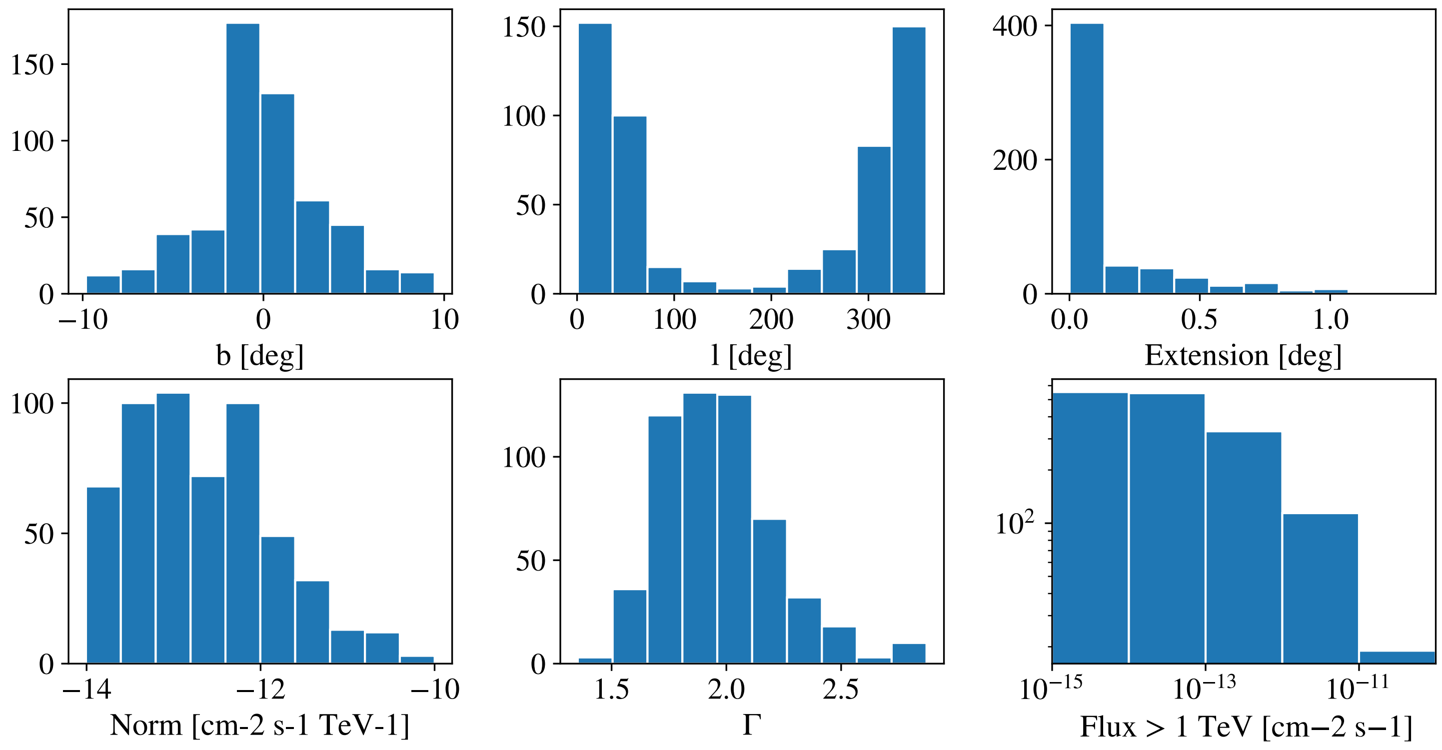
CTAO EXTRA-RESULTS: CTAO GALACTIC PLANE SIMULATION

- Original simulated population on the galactic plane

[CTAO Cons. 23]



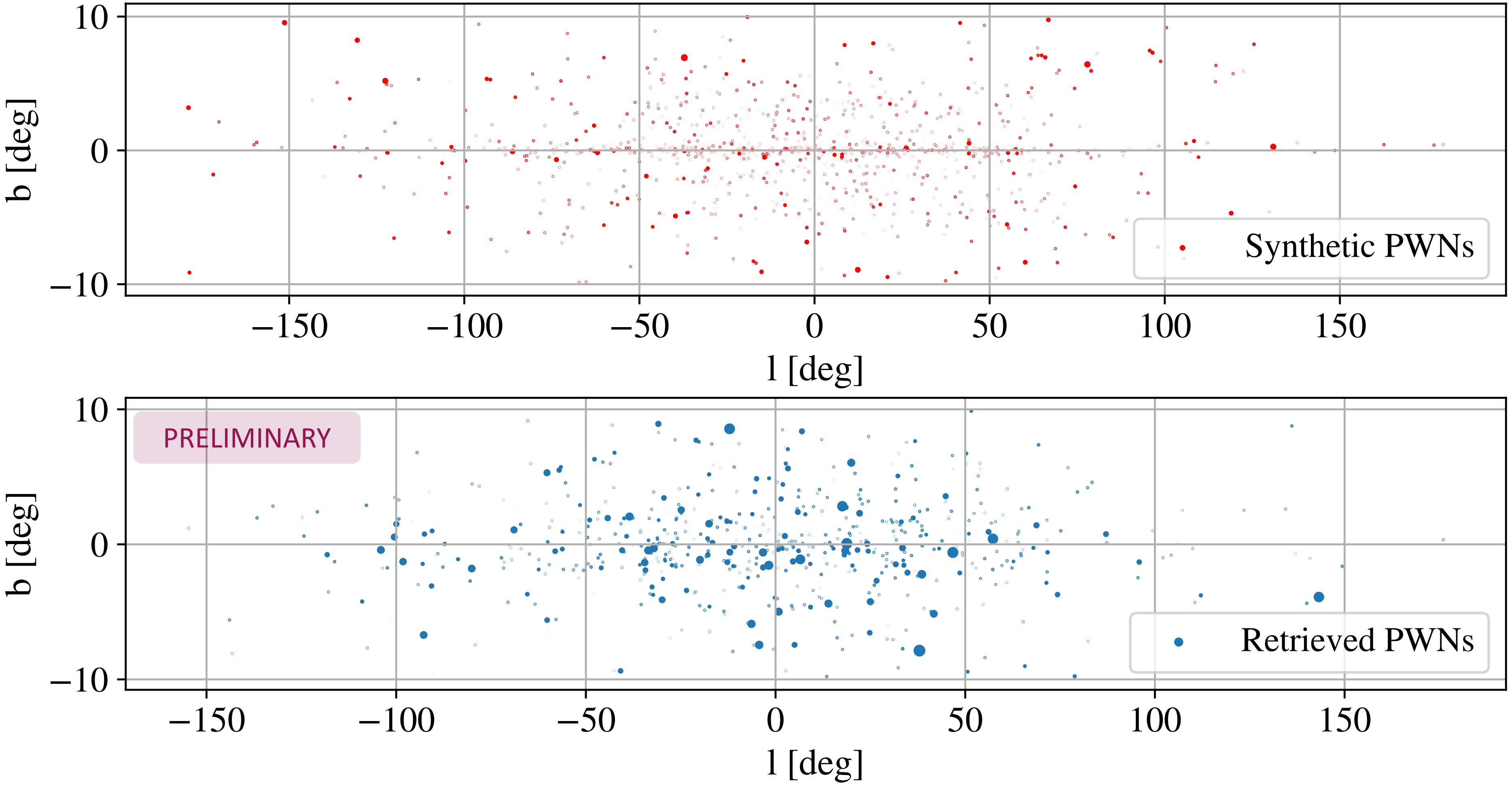
- We need several realizations (simulations) of the GP
- Extract the physical distributions of the sample $\phi(E) = \phi_0 \cdot \left(\frac{E}{E_0}\right)^{-\Gamma}$



- Sources fainter than 1/3 of target sensitivity are removed

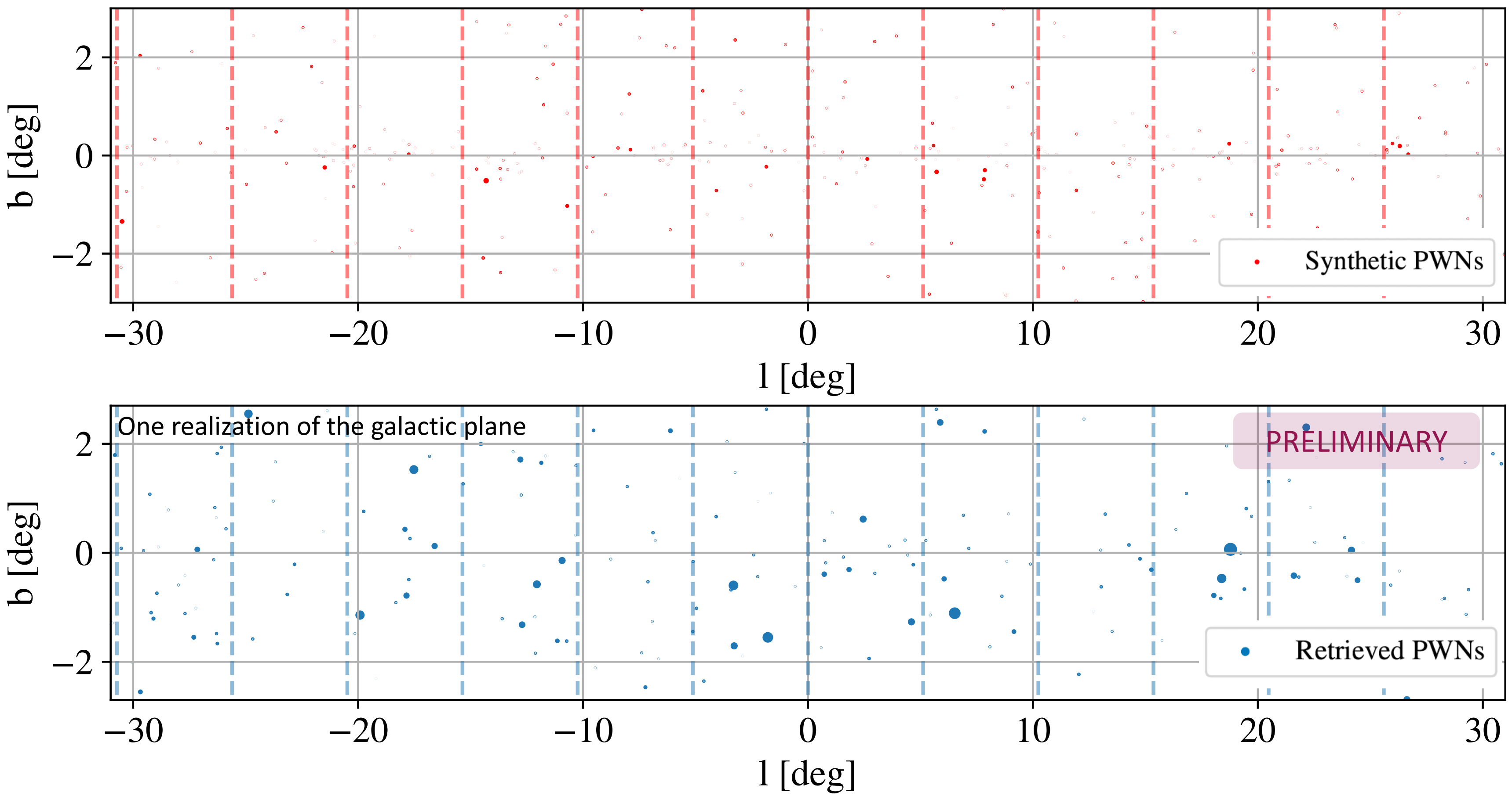
CTAO EXTRA-RESULTS: CTAO GALACTIC PLANE SIMULATION

- Comparison of original sample vs. one drawn realization from the physical distributions



CTAO EXTRA-RESULTS: CTAO GALACTIC PLANE SIMULATION

- Focus on the most crowded region
- Cover through patches: $-30 < l < 30$ deg $\quad -2.5 < b < 2.5$ deg



CTAO EXTRA-RESULTS: CTAO GALACTIC PLANE SIMULATION

$\gamma\pi$ A **Python** package for **gamma-ray** astronomy
[*Donath et al. 2023*]

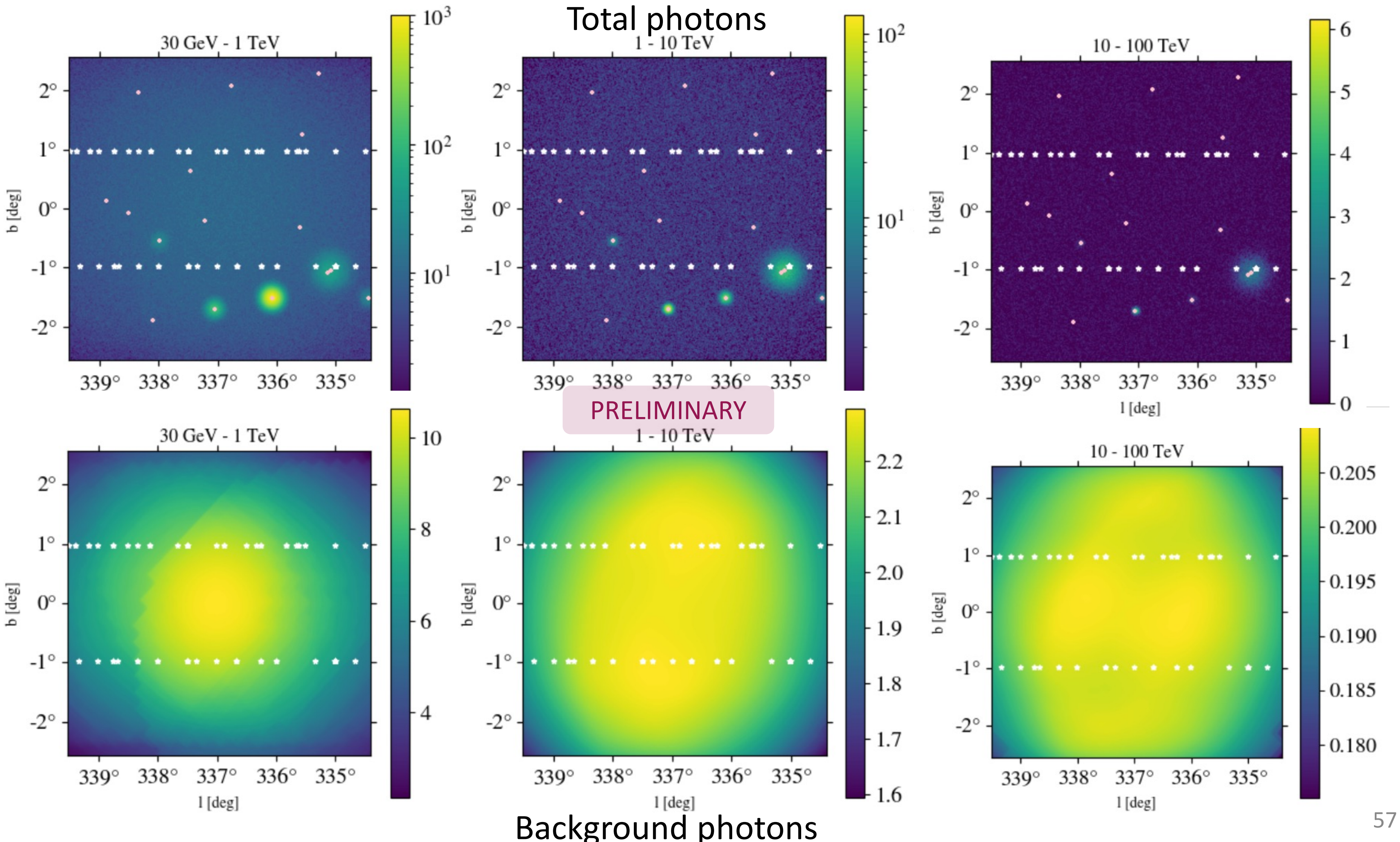
- Cover the galactic plane through patches

$$-30 < l < 30 \text{ deg}$$
$$-2.5 < b < 2.5 \text{ deg}$$

- 12 patches per each complete simulation of the galactic plane

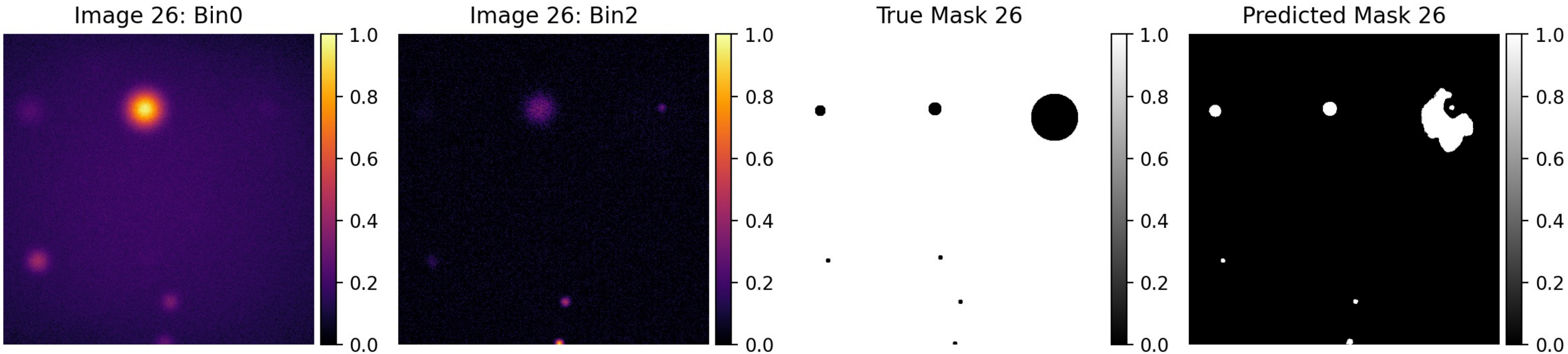
$$512 \text{ pix} \times 512 \text{ pix}$$
$$5.12 \text{ deg} \times 5.12 \text{ deg}$$

- 3 energy bins (following the instrument’s sensitivity)



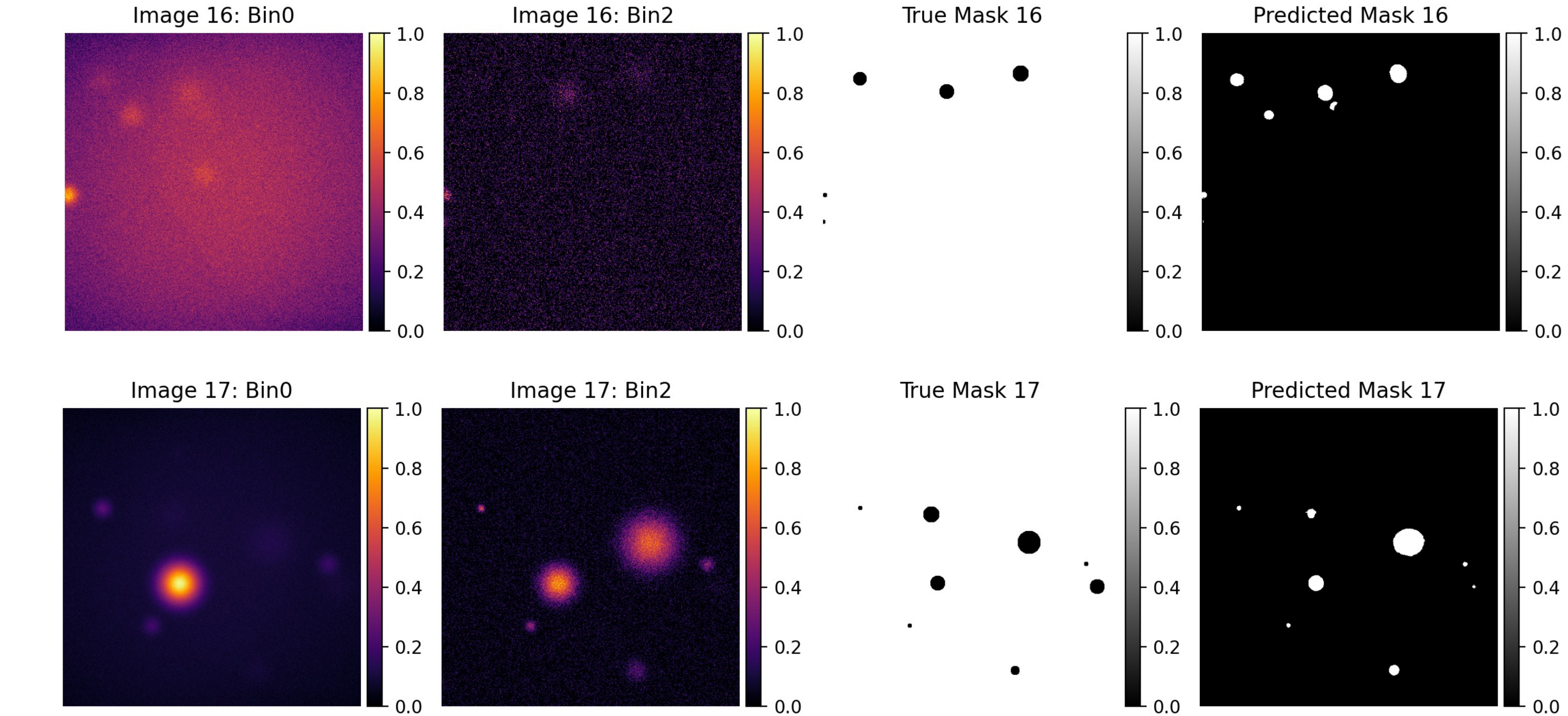
CTAO EXTRA-RESULTS: FIRST TRIALS USING ASID

- Running U-Net + LoG is problematic, since the mask we use for training, have an extension the original sigma used for generating the simulation, and some sources lay behind the larger ones, confusing the segmentation part



CTAO EXTRA-RESULTS: FIRST TRIALS USING ASID

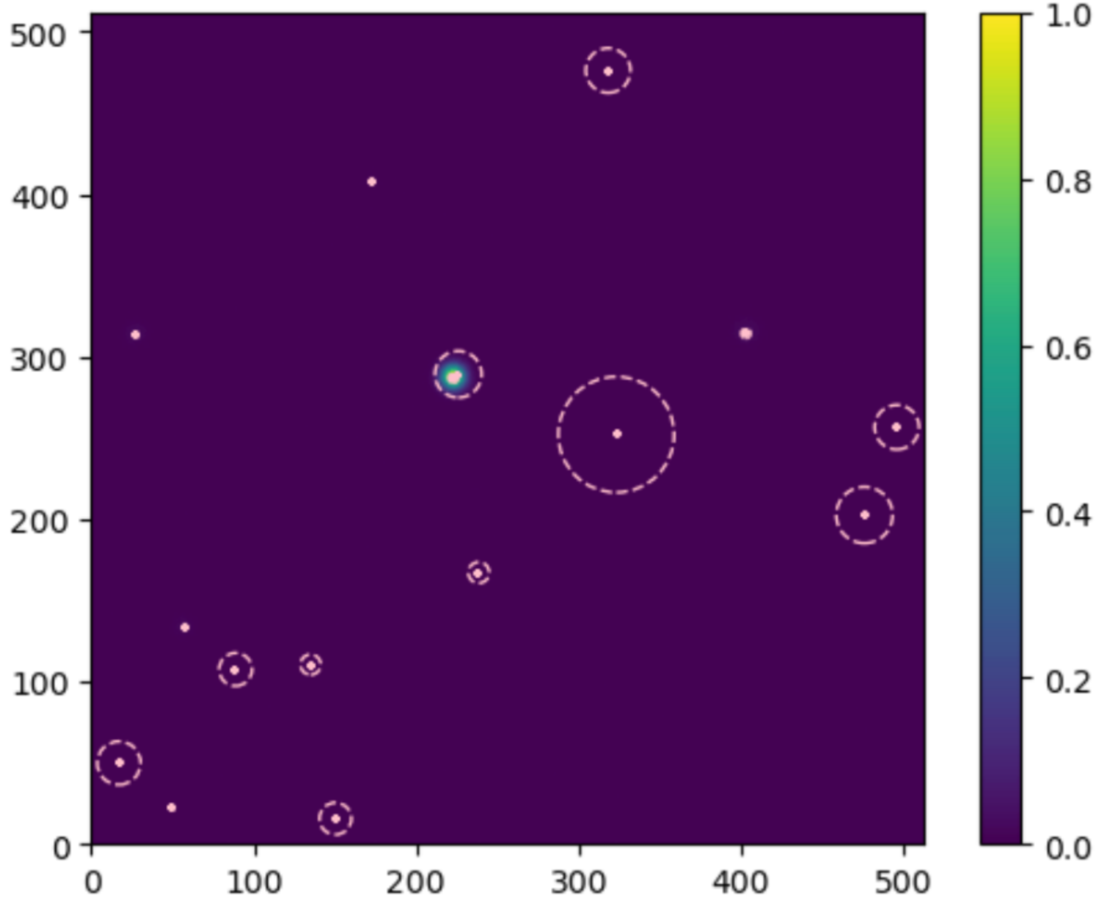
- Approach 1: Remove from the training the fainter sources, with a different flux threshold depending on each size, original size then should more or less correspond to size in last energy bin, for point-like source we use value of CTAO PSF at those energies (0.05 deg)



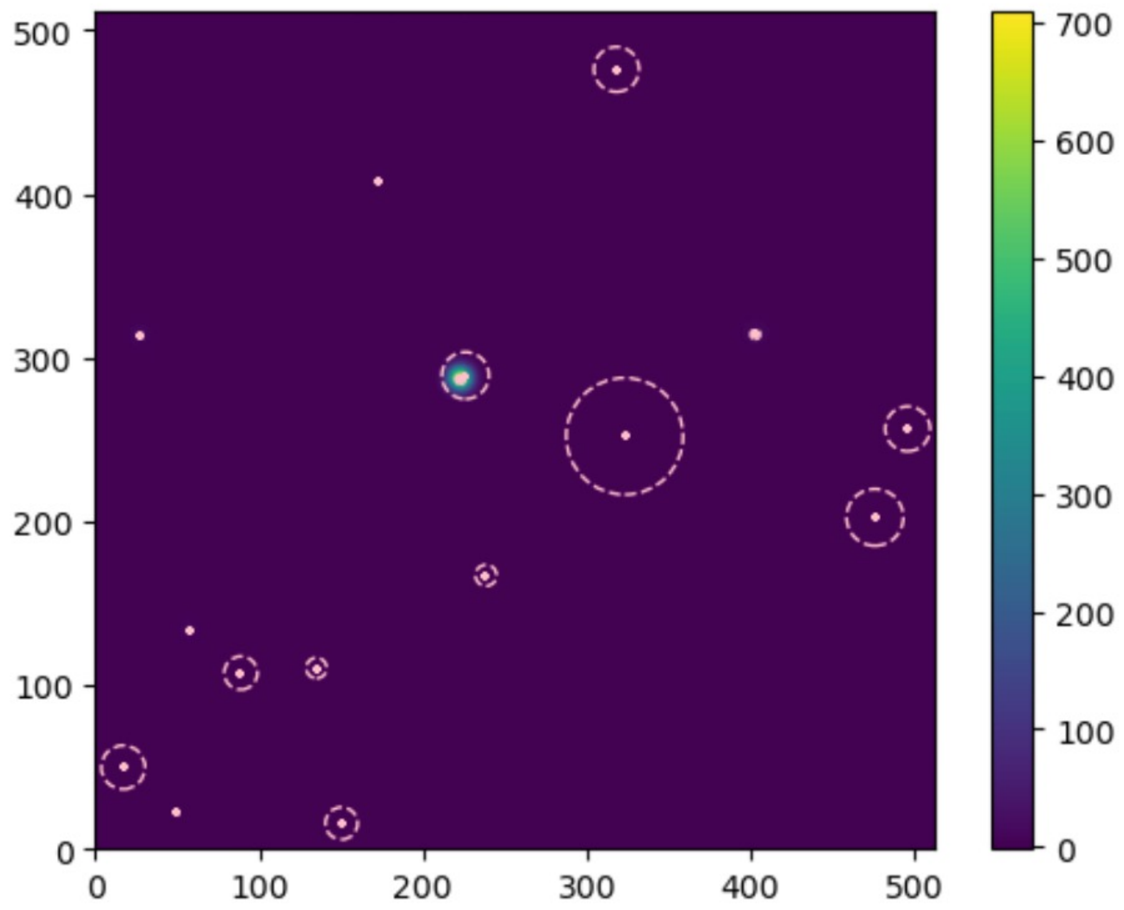
CTAO EXTRA-RESULTS: OVERLAPPING PROBLEMS

- Example for patches_v1 patch_0, bin 1

Normalized to max image

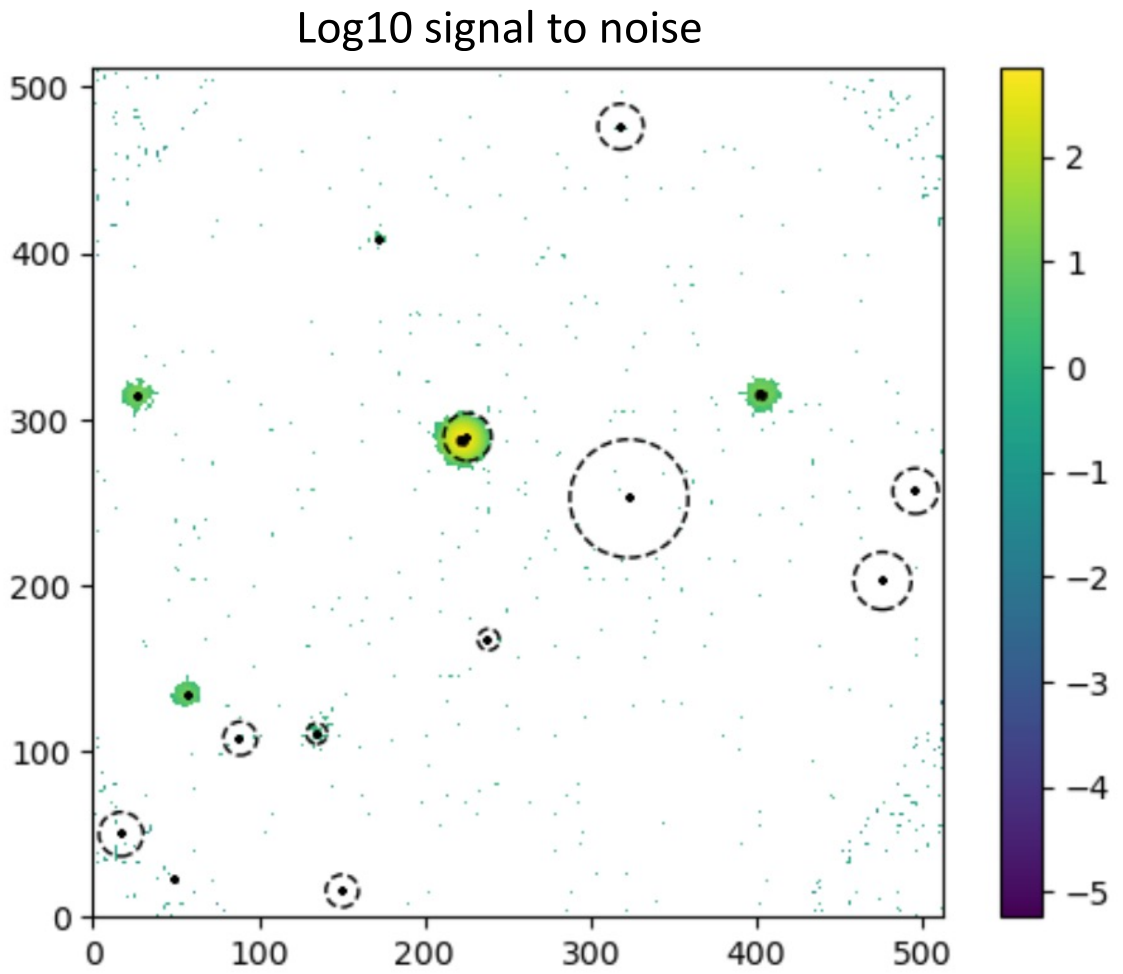


Signal to noise



CTAO EXTRA-RESULTS: OVERLAPPING PROBLEMS

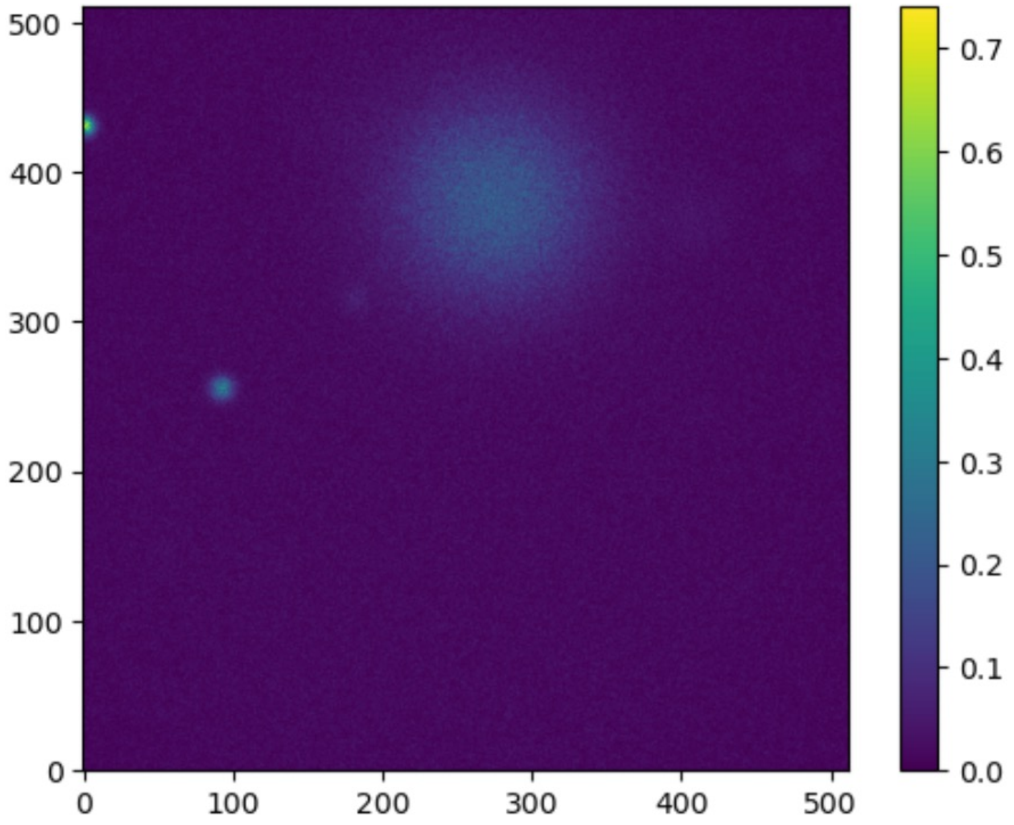
- Example for patches_v1 patch_0, bin 1



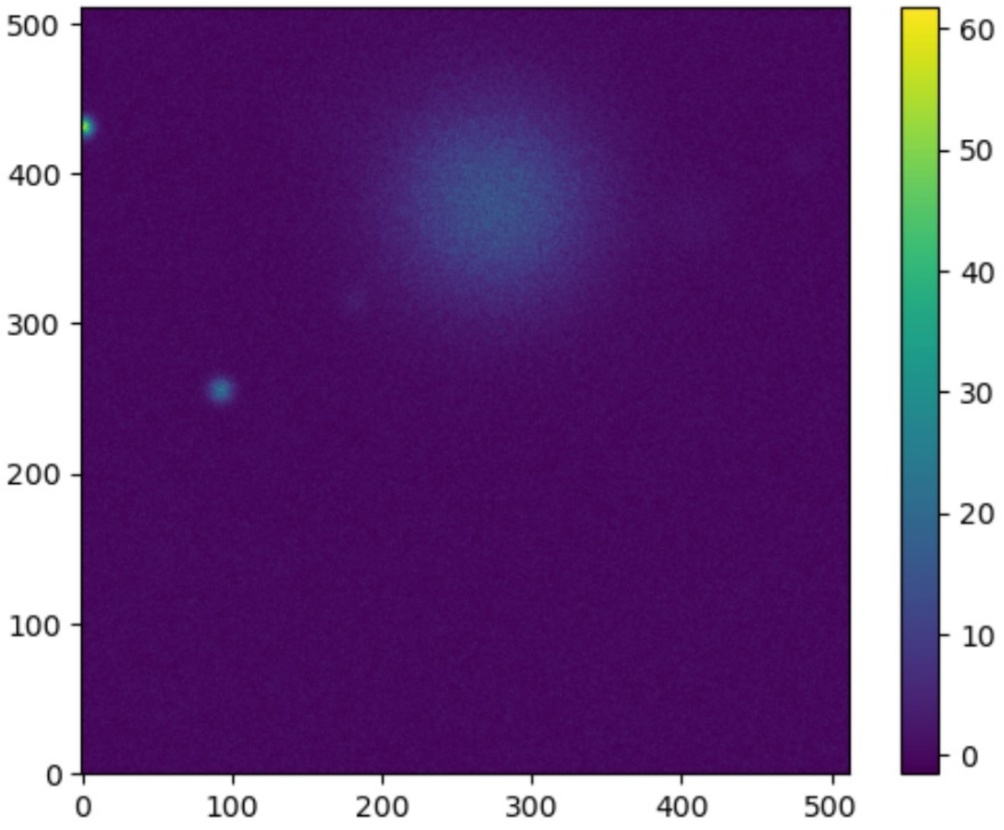
CTAO EXTRA-RESULTS: OVERLAPPING PROBLEMS

- Example for patches_v1 patch_1, bin 1

Normalized to max



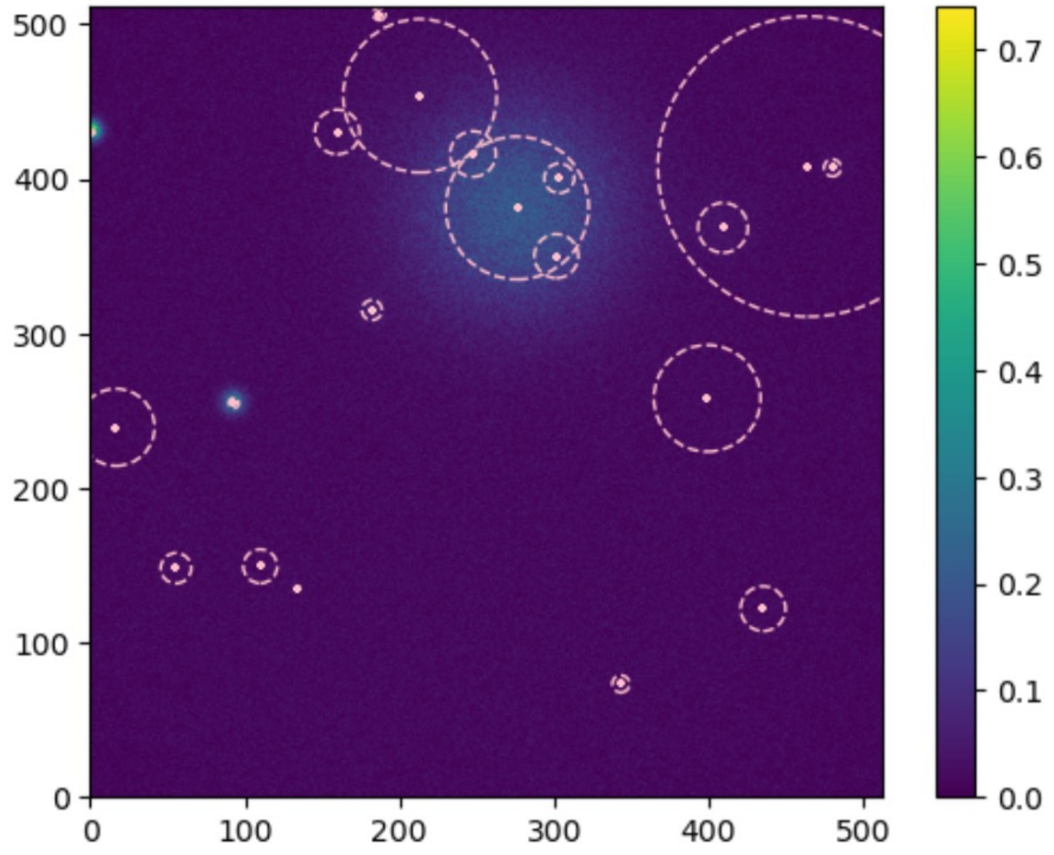
Signal to noise



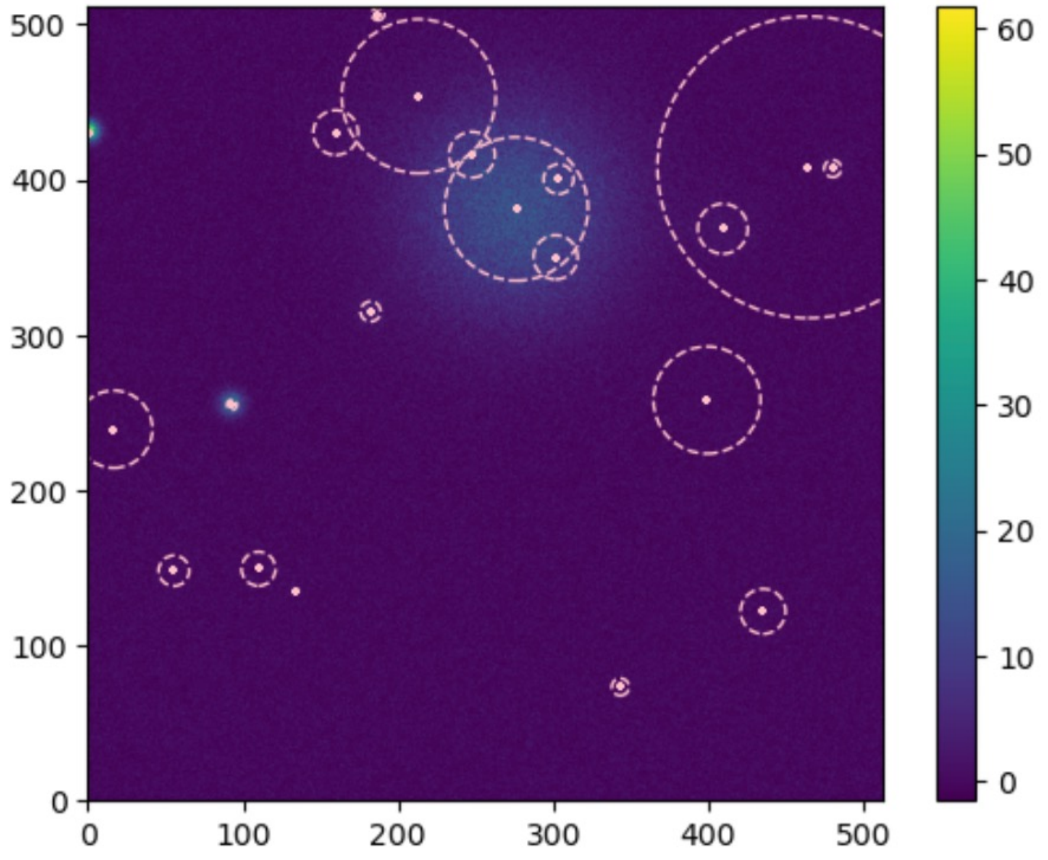
CTAO EXTRA-RESULTS: OVERLAPPING PROBLEMS

- Example for patches_v1 patch_1, bin 1

Normalized to max image



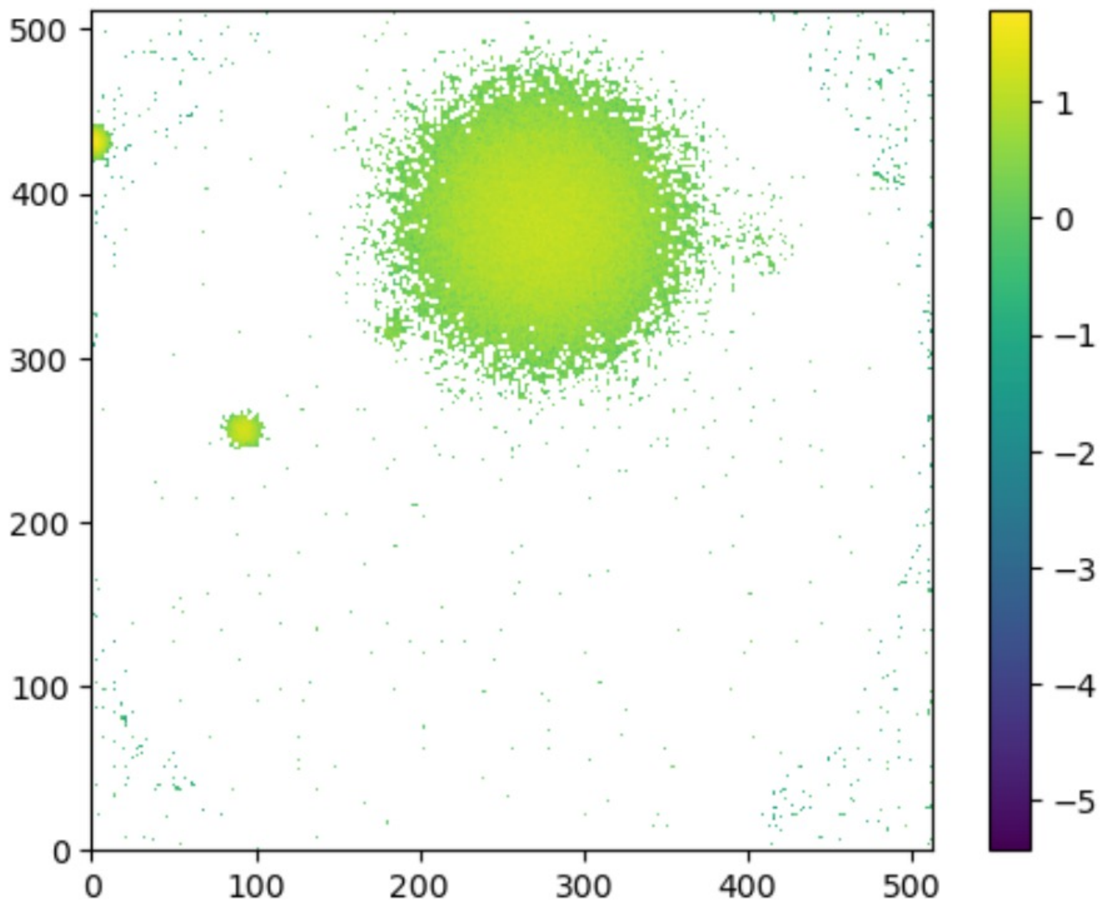
Signal to noise



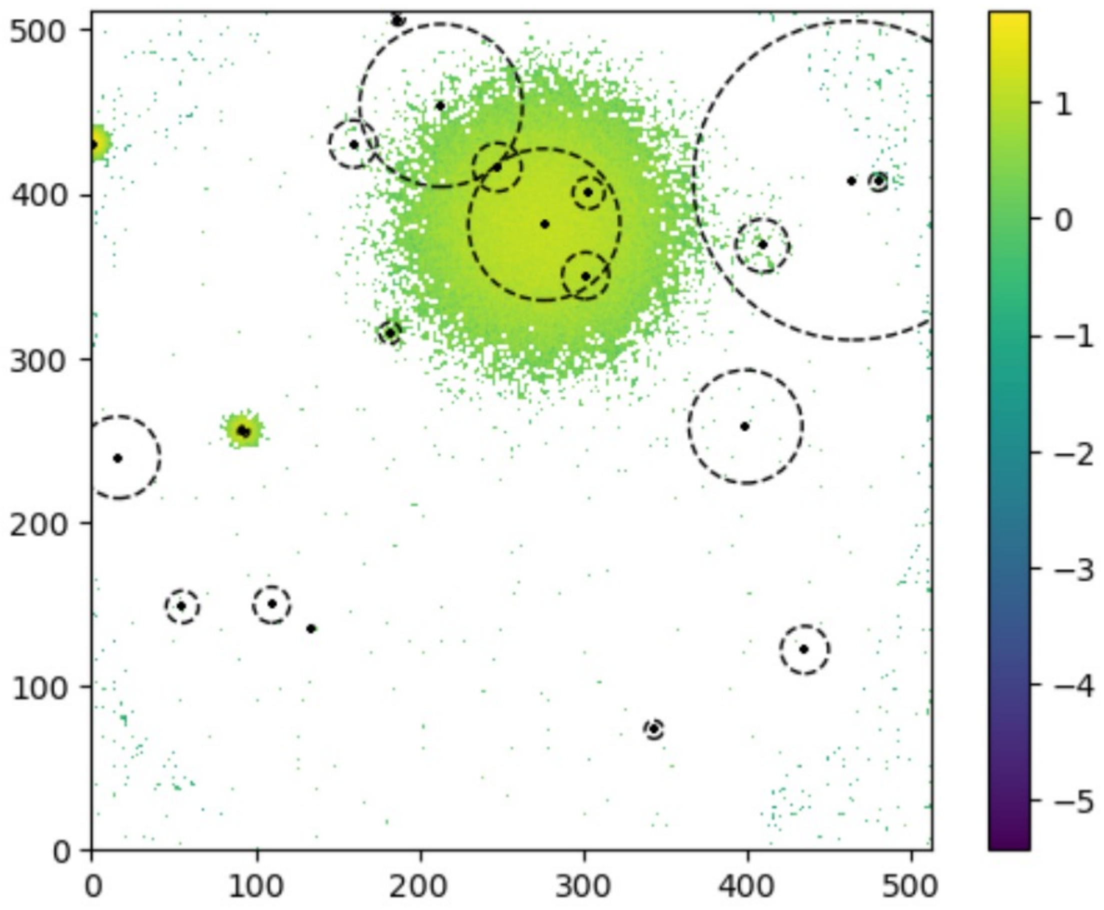
CTAO EXTRA-RESULTS: OVERLAPPING PROBLEMS

- Example for patches_v1 patch_1, bin 1

Log10 signal to noise

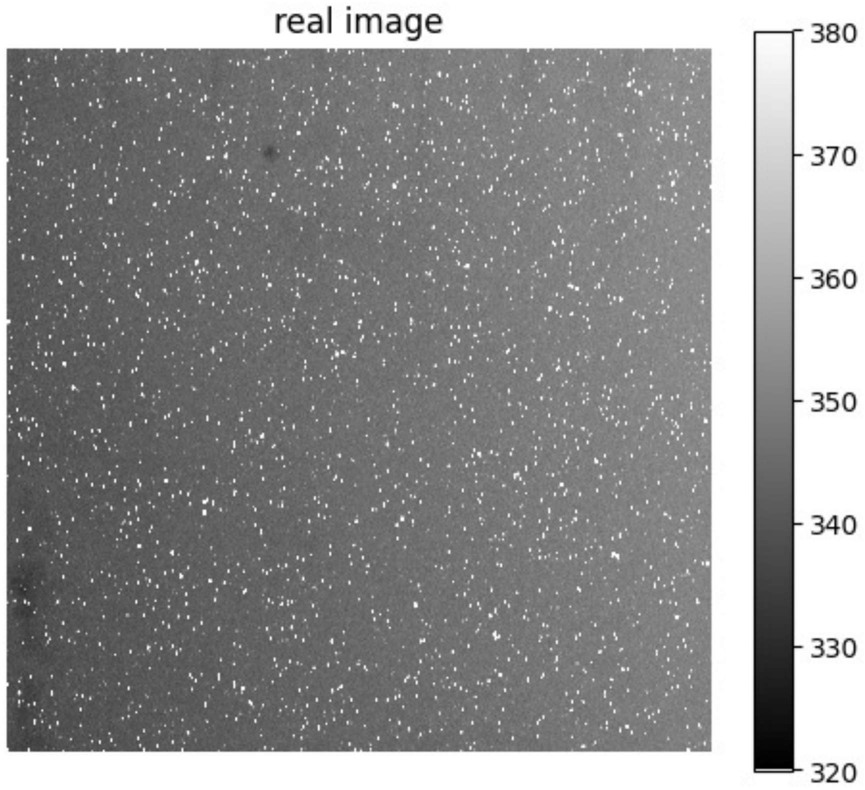


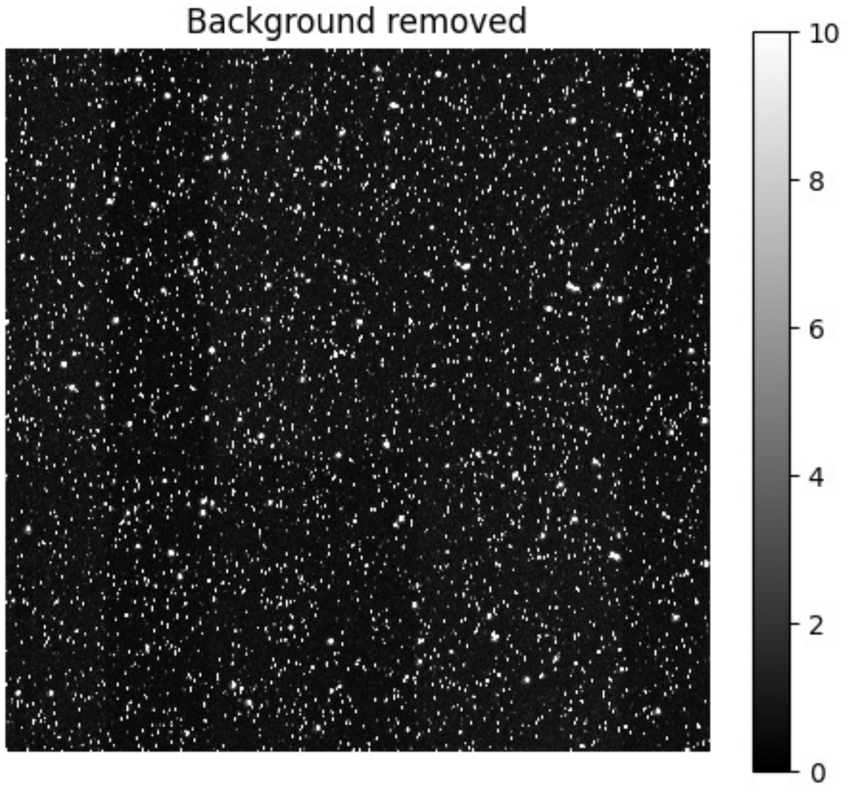
Log10 signal to noise to noise



TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION

Background removal: Optical Data
Model: Denoising Diffusion (Attention U-net as backbone)

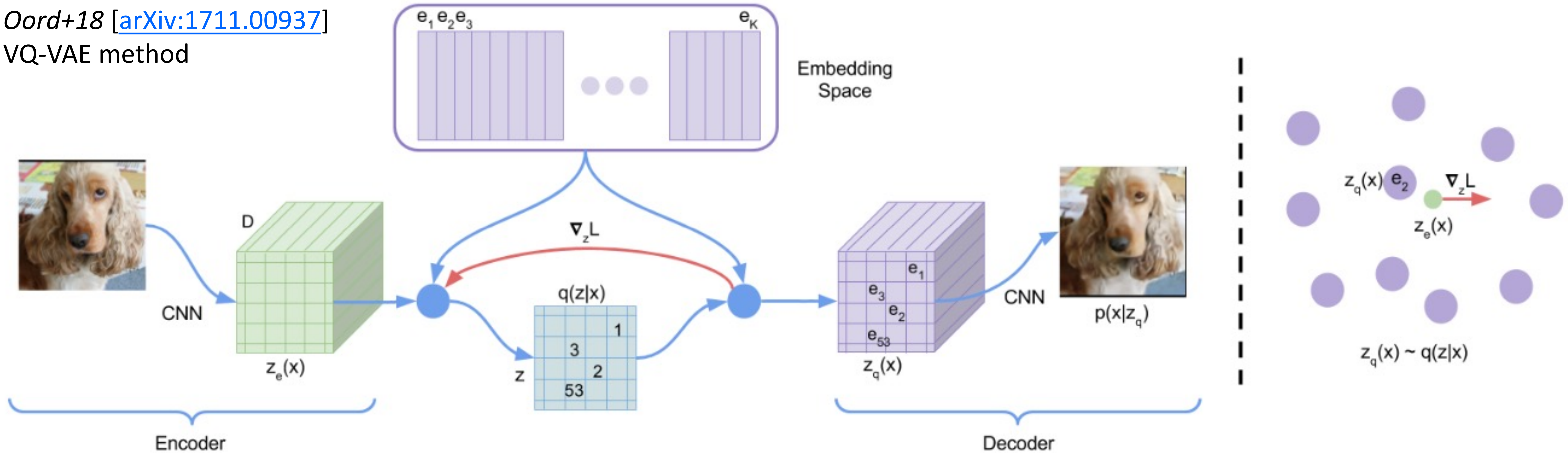




TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION

Through the embedding matrix we try to find most important words ➡ Find most important features in latent space

Oord+18 [arXiv:1711.00937]
VQ-VAE method



TOWARDS FOUNDATION MODEL FOR ASTROPHYSICAL SOURCE DETECTION

- Main objective: