EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE EuCAIFCon 2024

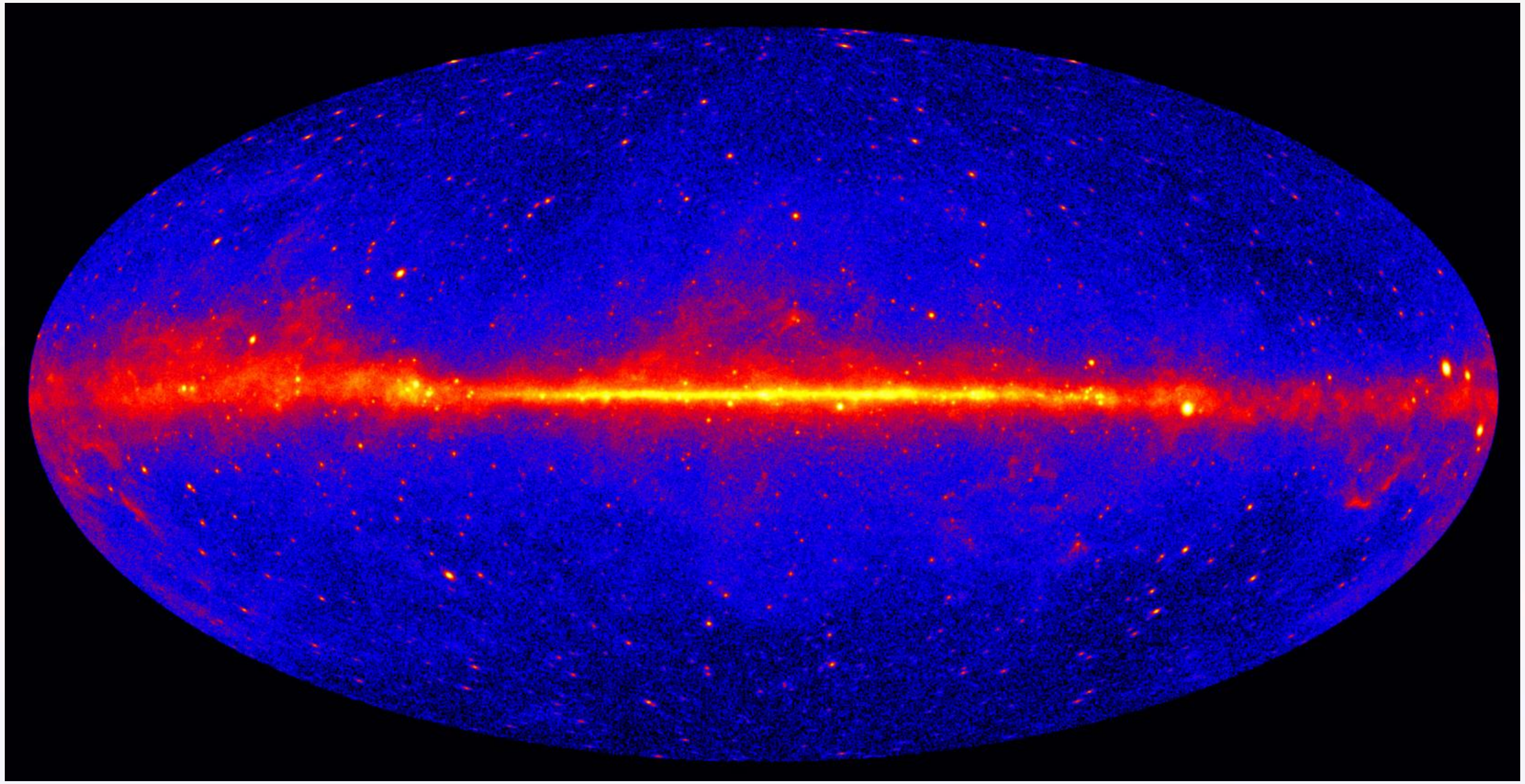
Characterizing High Energy Gamma-Ray Sources Using Deep Learning (& More…)

ID: 141

S. Bhattacharyya, F. Stoppa*, R. Austri, S. Caron, G. Principe, D. Malyshev, G. Zaharijas, R. Nicolas et.al.,

SMASH machine learning for science and humanities postdoctoral program

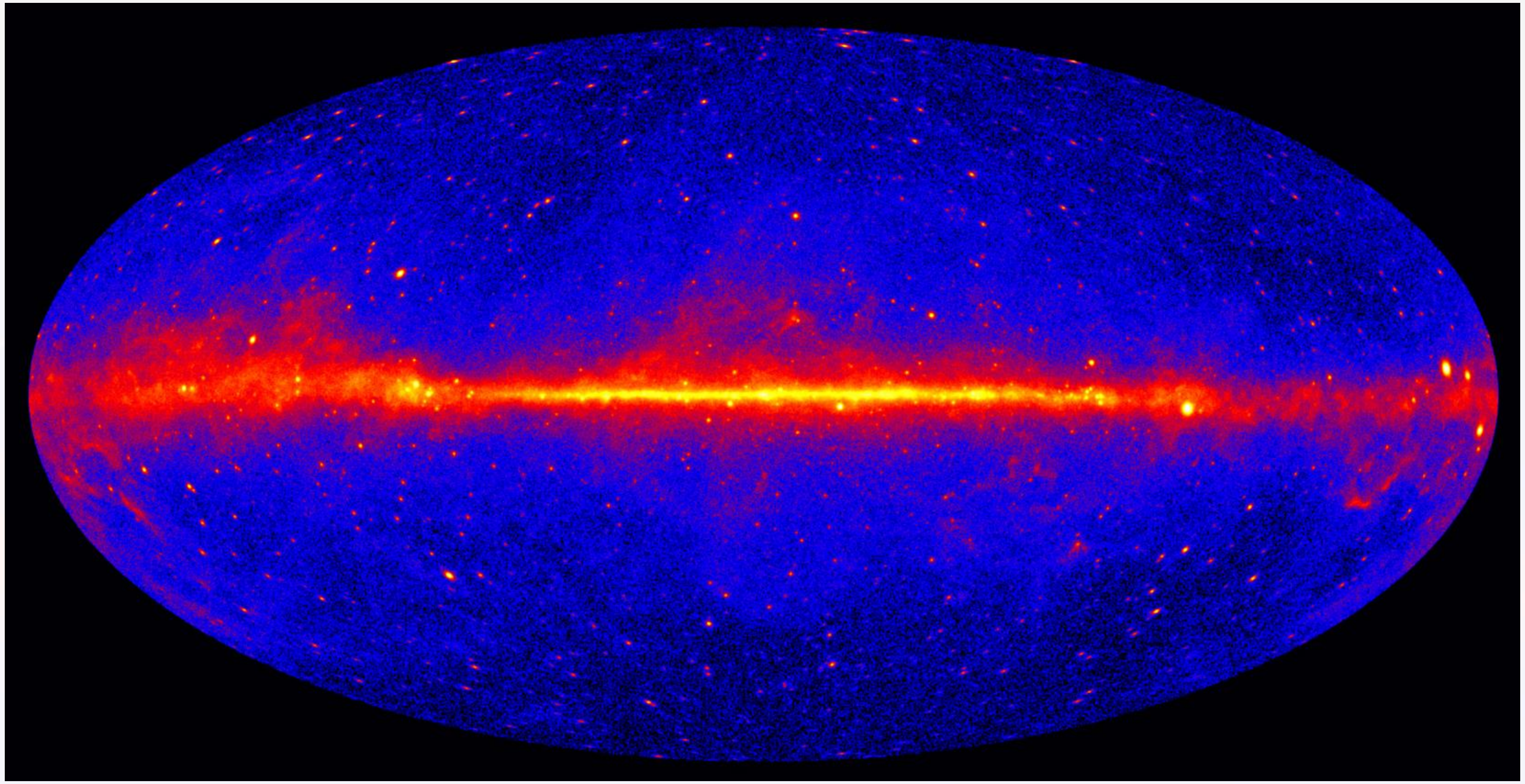
Objective



Source: 5 years of Fermi-LAT observation (E>1 GeV)

Given a gamma-ray sky-map, can a DNN-based pipeline detect the point sources, predict precise locations (including uncertainties), and eventually, characterize them?

Objective

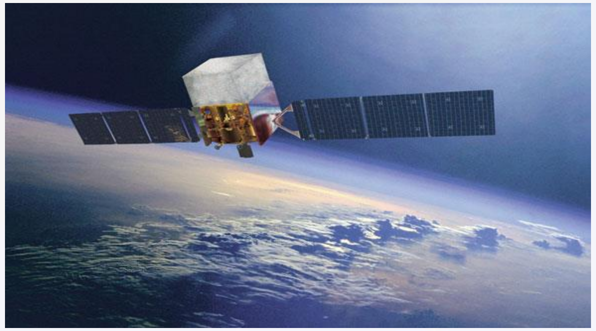


Source: 5 years of Fermi-LAT observation (E>1 GeV)

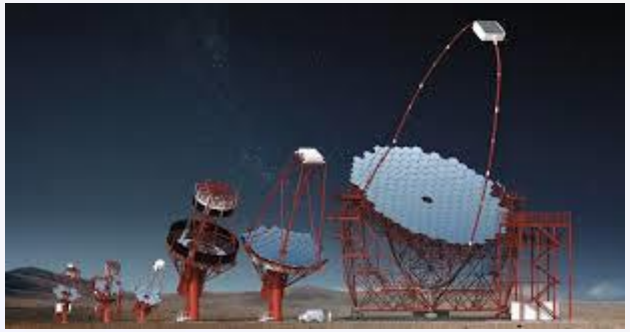
Given a gamma-ray sky-map, can a DNN-based pipeline detect the point sources, predict precise locations (including uncertainties), and eventually, characterize them?

Can these methodologies be applicable at a different region of the EM spectrum (e.g. Optical)?

Gamma-Ray Telescopes (Considered in this Study): Fermi-LAT & CTA



- Fermi Large Area Telescope (LAT)
- Space-based detector (collecting data from 2008 onwards).
- Sensitive to $\sim 300\ \mathrm{MeV} \leq E \leq\sim 100\ \mathrm{GeV}$ photons.



- Cherenkov Telescope Array (CTA)
- Ground-based detector.
    - Two sites: La-Palma, Chile.
- Sensitive to $\sim 30\ \mathrm{GeV} \leq E \leq\sim 100\ \mathrm{TeV}$ photons.

Getting Started with Fermi-LAT: Supervised ML and Data Generation

- To learn a mapping from input to output based on example input-output pairs. ‘Supervised Learning’
  - Only one ‘instance’ from real data; we prepare realistic simulated data.

Getting Started with Fermi-LAT: Supervised ML and Data Generation

- To learn a mapping from input to output based on example input-output pairs. 'Supervised Learning'
  - Only one 'instance' from real data; we prepare realistic simulated data.

- Create a set of sky-maps with astrophysical source properties based on the Current Data (10 years of Observation).
  - Include properties of Active Galactic Nuclei (AGNs), Pulsars (PSRs) and Supernovae (SNe).

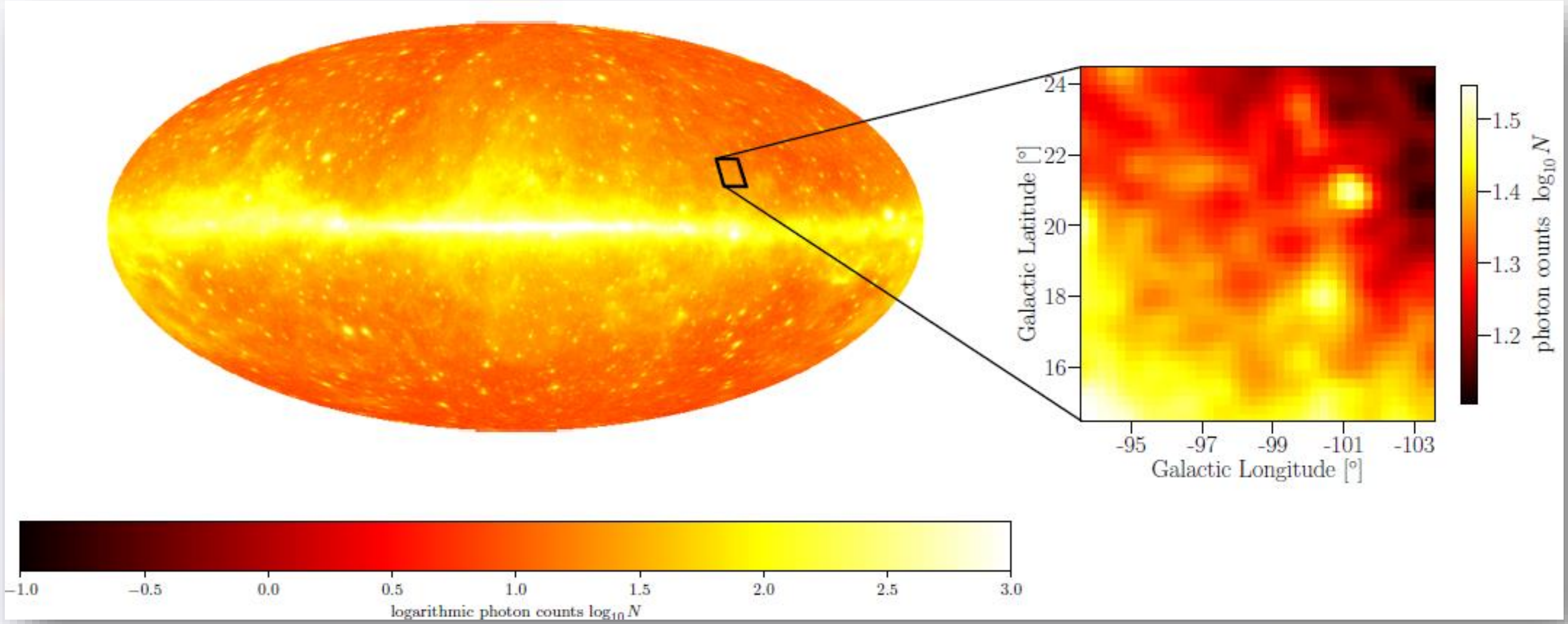
- 10 years of observation period [2008-2018].
  - Energy range 300 MeV to 1 TeV; 6 energy bins;

- Spatial resolution of the sky-maps increases with increasing energy.
  - From $0.8°$ at 0.3 GeV to $0.1° \geq 7$ GeV.

Mock Data Preparation:

- Generating skymaps: Use a simulator [Fermitools]: Convolve astrophysical source models and detector response.

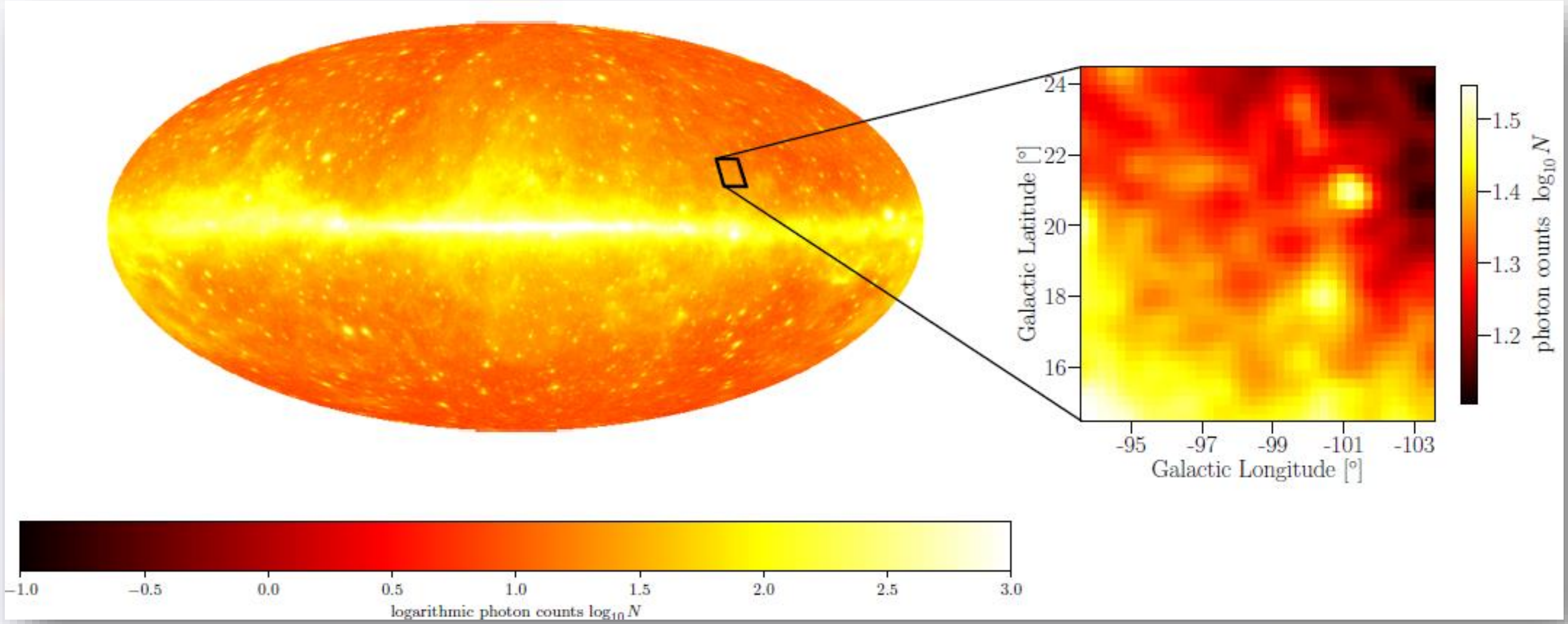
- Random patches (locations of sky) are used for training data. Reduces the possibility of localization network ‘learning’ the background and not the source.



Mock Data Preparation:

- Generating skymaps: Use a simulator [Fermitools]: Convolve astrophysical source models and detector response.

- Random patches (locations of sky) are used for training data. Reduces the possibility of localization network 'learning' the background and not the source.
  - Trained using one Interstellar Emission Model (IEM) and tested with a different IEM.
  - Some faint sources may be hidden in the IEM itself;



IEM small scale structures ⇄ Misidentification of Faint Sources

Why Use Deep Learning for this Task?

- Develop a complementary method to the likelihood method (detection, localization, flux estimation)
  - Detection Likelihood: $TS = \log\left(\frac{L}{L_0}\right)$; L_0: Likelihood without the source, L: Likelihood with the source.
- Machine learning including Deep Neural Net has been used to classify sources based on the catalog itself, including searching for various source classes from unidentified objects.
  - 'Multi-class classification of γ-ray sources & excess of GeV γ-rays near GC'; D. Malyshev Poster Id: 67, Explainable AI.

Why Use Deep Learning for this Task?

- Develop a complementary method to the likelihood method (detection, localization, flux estimation)
    - Detection Likelihood: $TS = \log\left(\frac{L}{L_0}\right)$; L_0: Likelihood without the source, L: Likelihood with the source.

- Machine learning including Deep Neural Net has been used to classify sources based on the catalog itself, including searching for various source classes from unidentified objects.
    - 'Multi-class classification of γ-ray sources & excess of GeV γ-rays near GC'; D. Malyshev Poster Id: 67, Explainable AI.

- Detecting point sources using the traditional likelihood method depends on modeling the background.
    - Possibility of IEM Model independent results?

- Possibility of extending the pipeline to test its capability at other wavelengths.

Data Analysis Pipeline:

- Detection + Localization
  - Segmenting source pixels from background pixels.
  - Find the center of the source pixels;

- U-Net (Modified)
- Laplacian of Gaussian/K-Means

Data Analysis Pipeline:

- Detection + Localization
  - Segmenting source pixels from background pixels.
  - Find the center of the source pixels;

- U-Net (Modified)
- Laplacian of Gaussian/K-Means

- Location Uncertainty Estimation
  - Regression network; Refined location + Uncertainties.
- Flux Estimation (+ Uncertainties):
  - Same as above; Estimate the flux with uncertainties.

- Deep Ensembles

Data Analysis Pipeline:

- Detection + Localization
  - Segmenting source pixels from background pixels.
  - Find the center of the source pixels;

- U-Net (Modified)
- Laplacian of Gaussian/K-Means

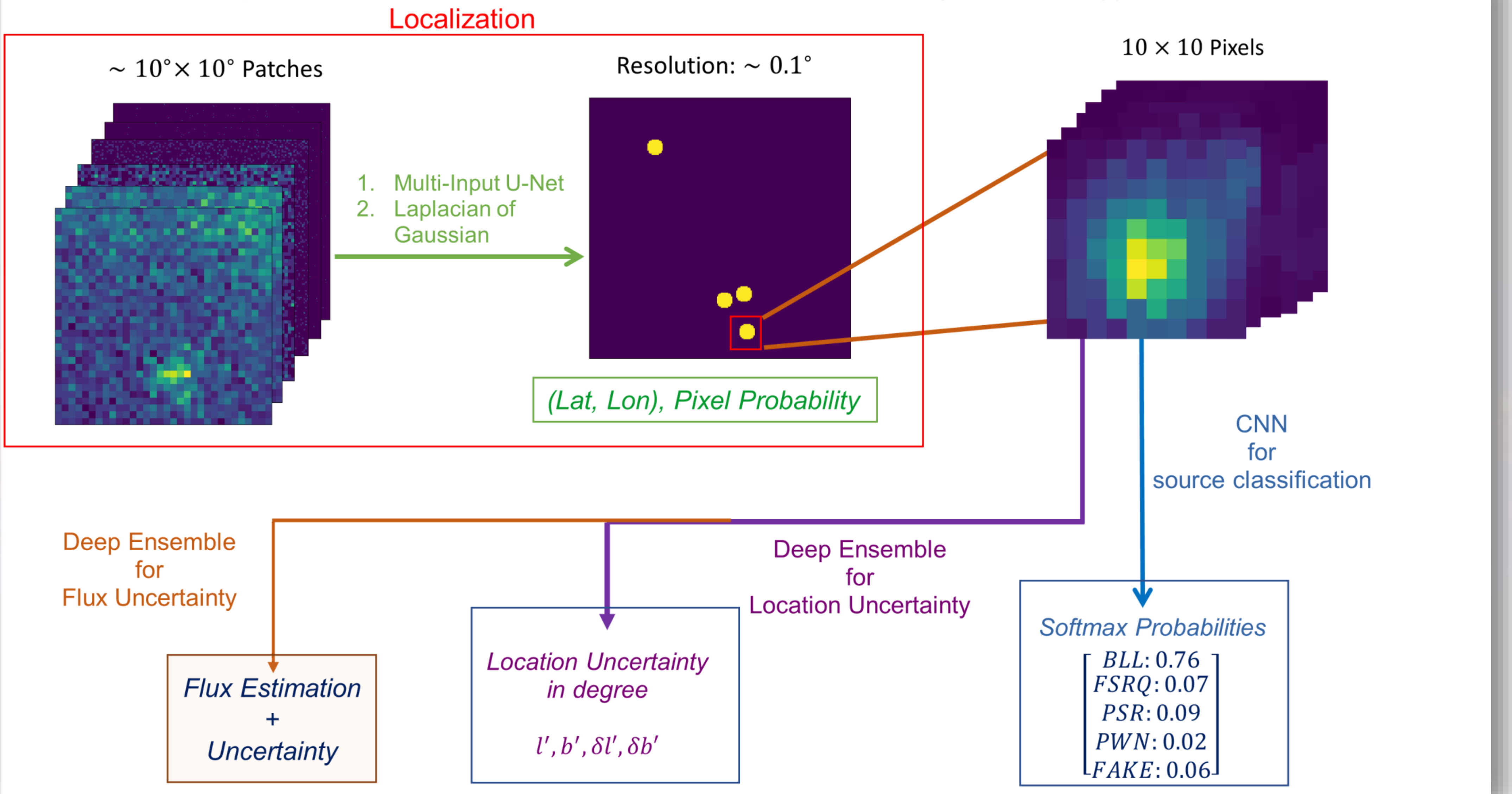
- Location Uncertainty Estimation
  - Regression network; Refined location + Uncertainties.
- Flux Estimation (+ Uncertainties):
  - Same as above; Estimate the flux with uncertainties.

- Deep Ensembles

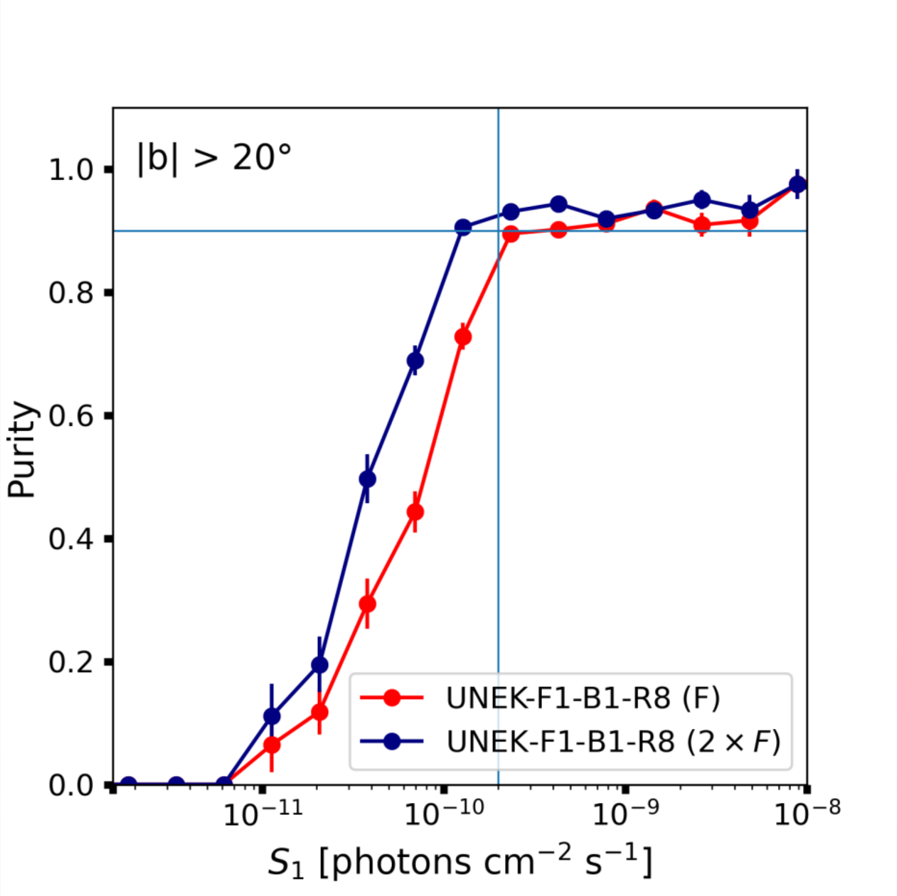
- Classification:
  - Binary/Multi-class classification.

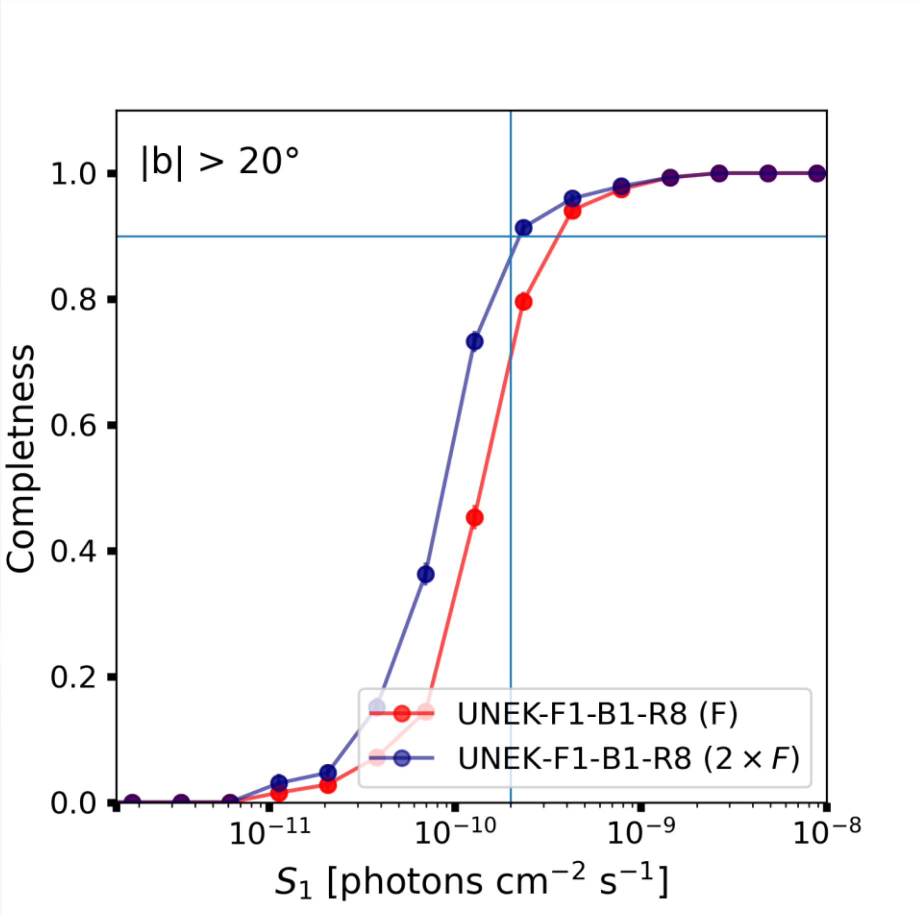
- VGG Like

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



Performance Evaluation on Simulated Data: Precision (Purity) and Recall (Completeness)





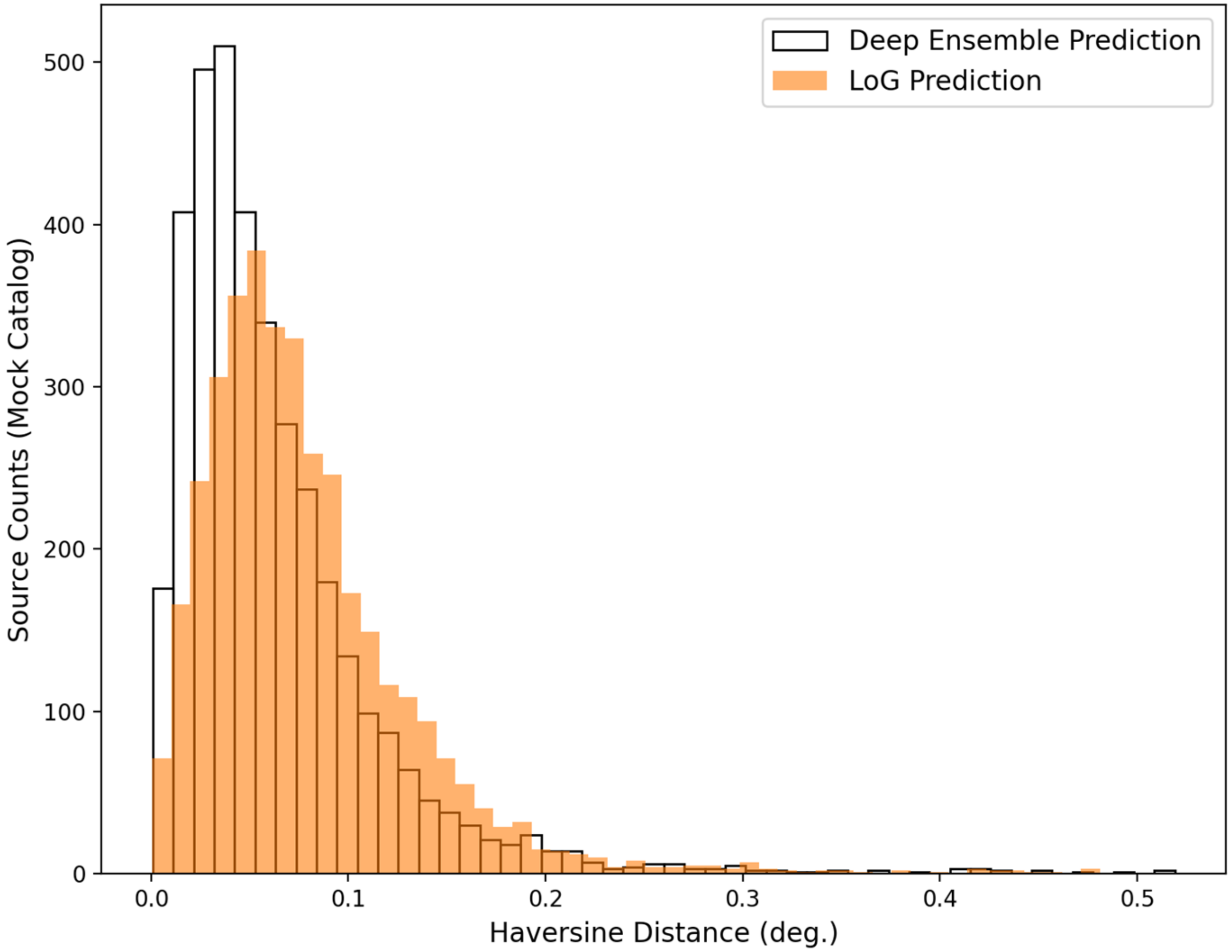
Comparison of network performance with Front Only (F) and 2 times Front Data. (2F)

'Front': Photons converted in the Front part of the detector (thicker calorimeter==better reconstruction).

Vertical Blue Line: LAT 4FGL catalog threshold.

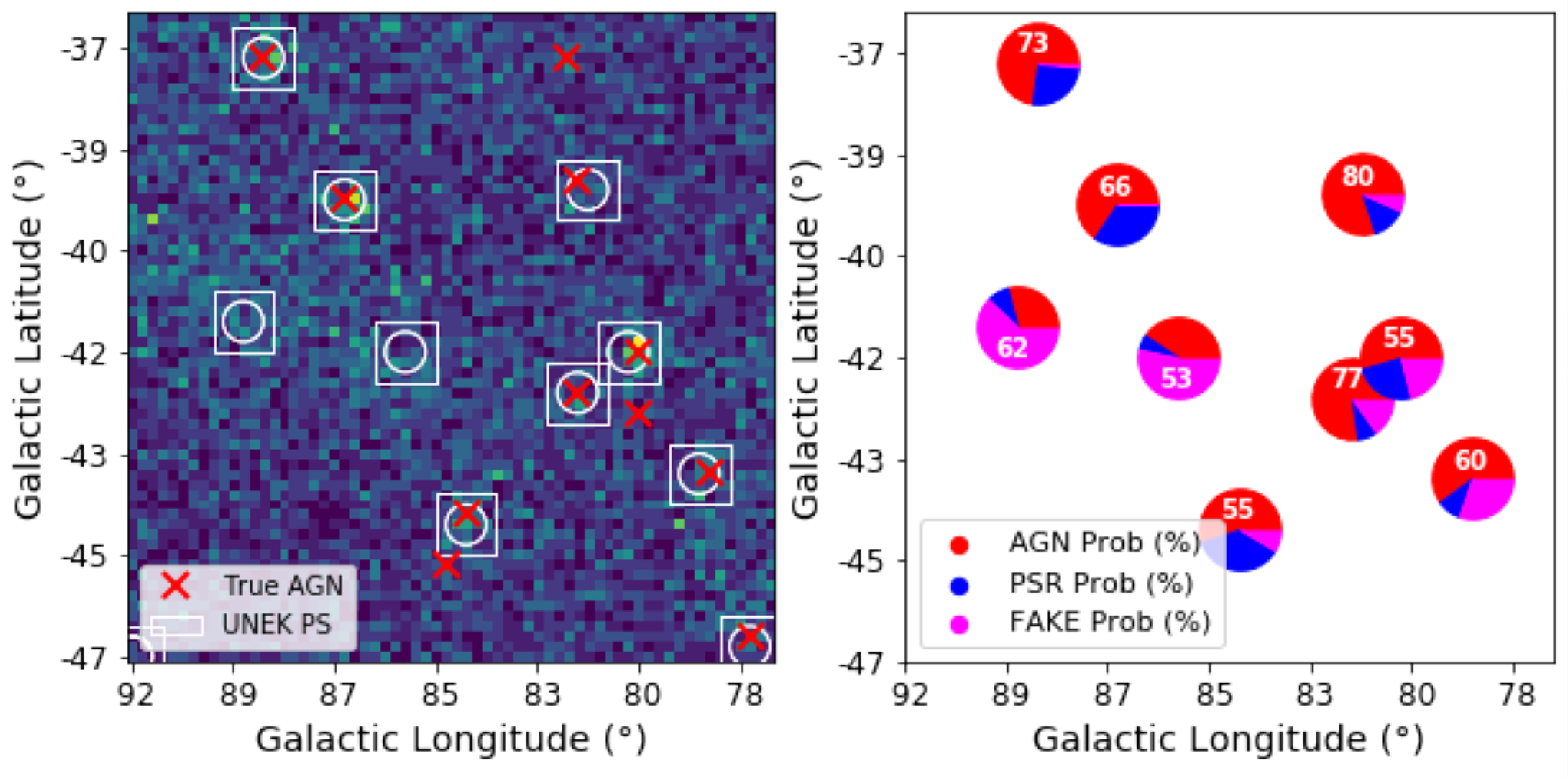
Location Reconstruction with Deep Neural Nets

- After the initial location prediction (LoG), we further refine the location using deep ensemble.
  - A regression network; Ensemble of 15 different networks; Aggregate and average for location uncertainty prediction.

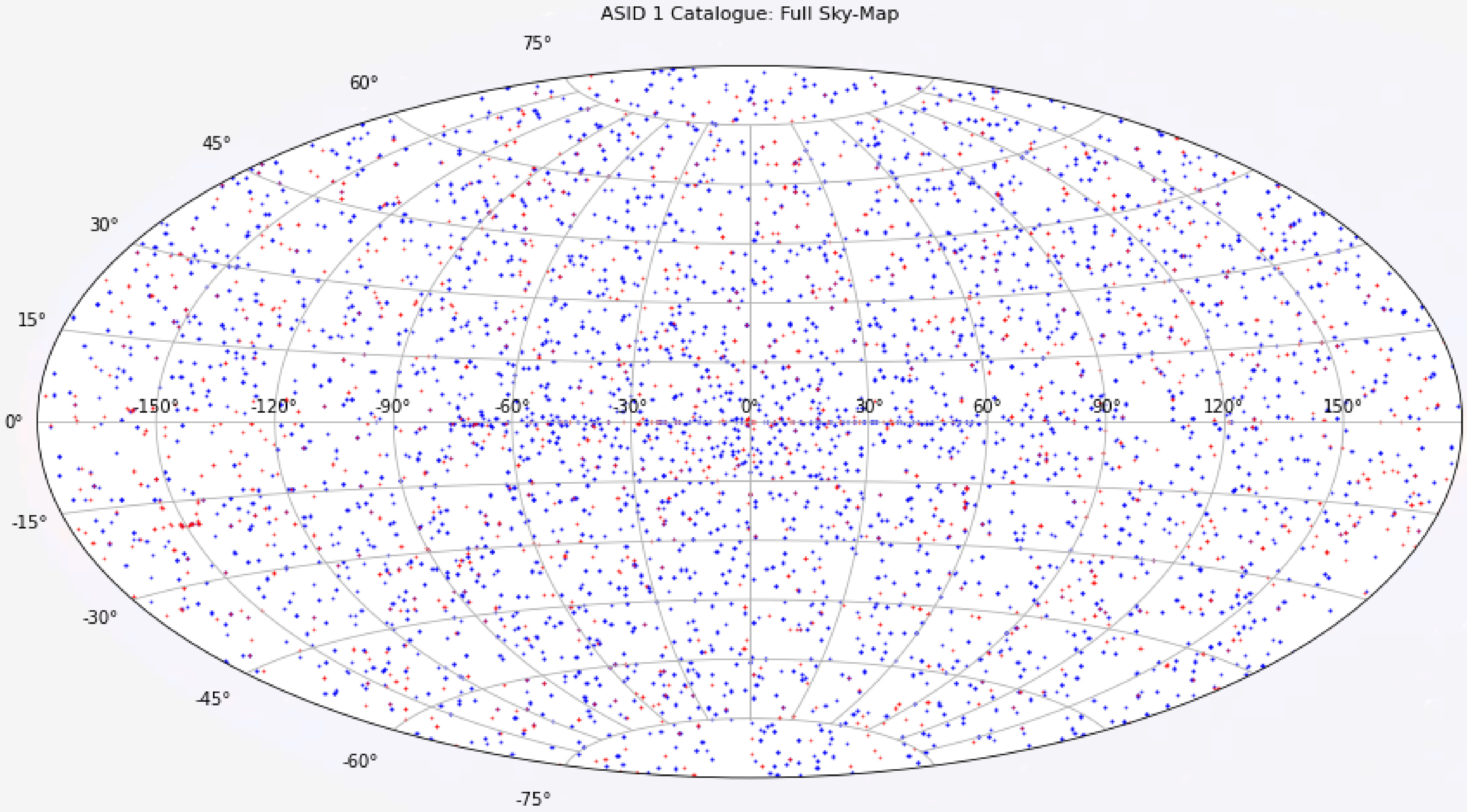


Building a Complementary Gamma-Ray Catalog

- Long-term target: Apply our algorithm on the real data & Build a complementary gamma-ray catalog.
- Already tested for simulated data:
  - 'Identification of point sources in gamma rays using U-shaped convolutional neural networks and a data challenge' [arXiv: 2103.11068]; A&A (A62, 2021); B. Panes, S. Caron, R. Austri, G. Zaharijas et.al.



Building a Complementary Gamma-Ray Catalog

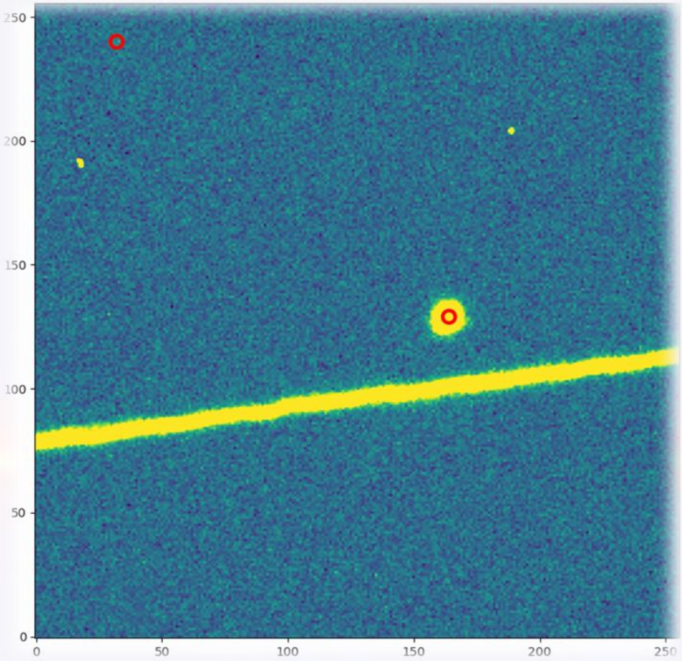


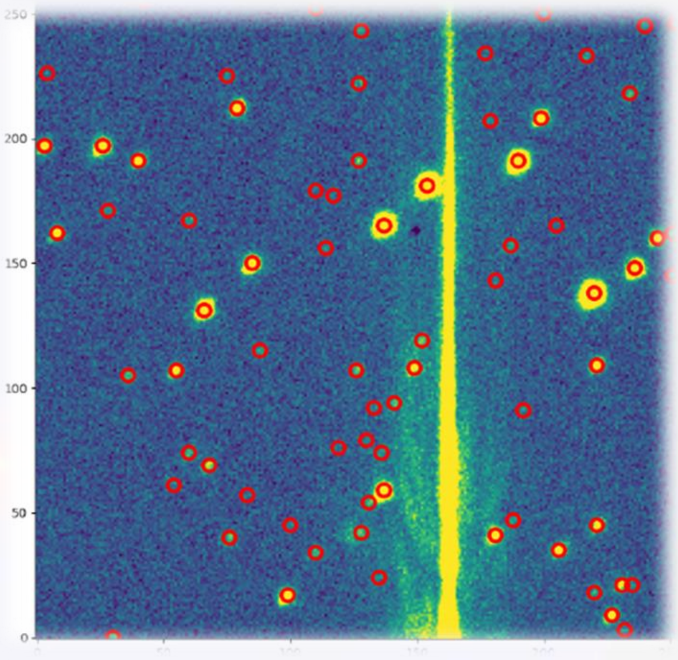
- Application on real data.
  - Ongoing
- Sources found by ASID
  - Blue: True sources
  - Red: False Positives

(After classification)

Can We Extend The Pipeline for Other Wavelengths? (Localization)

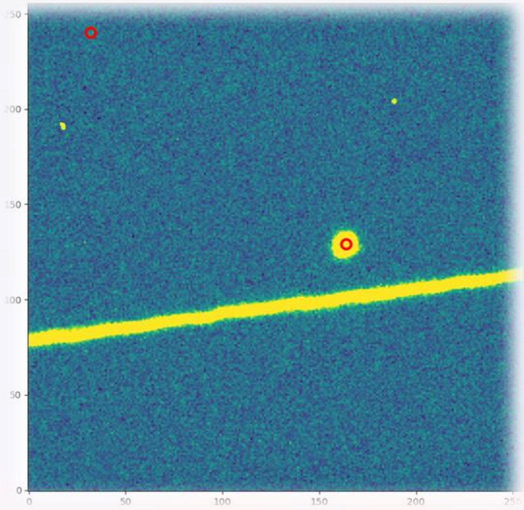
- Trained and tested with MeerLICHT data
  - 'ASID-Light: Fast Optical Source Localization'; [arXiv: 2202.00489]; A&A (A109, 2022); F. Stoppa et.al
- Automatic rejection of CR contaminants, satellite trail.

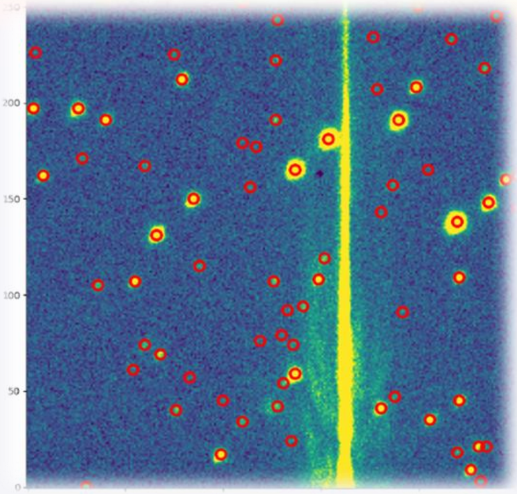


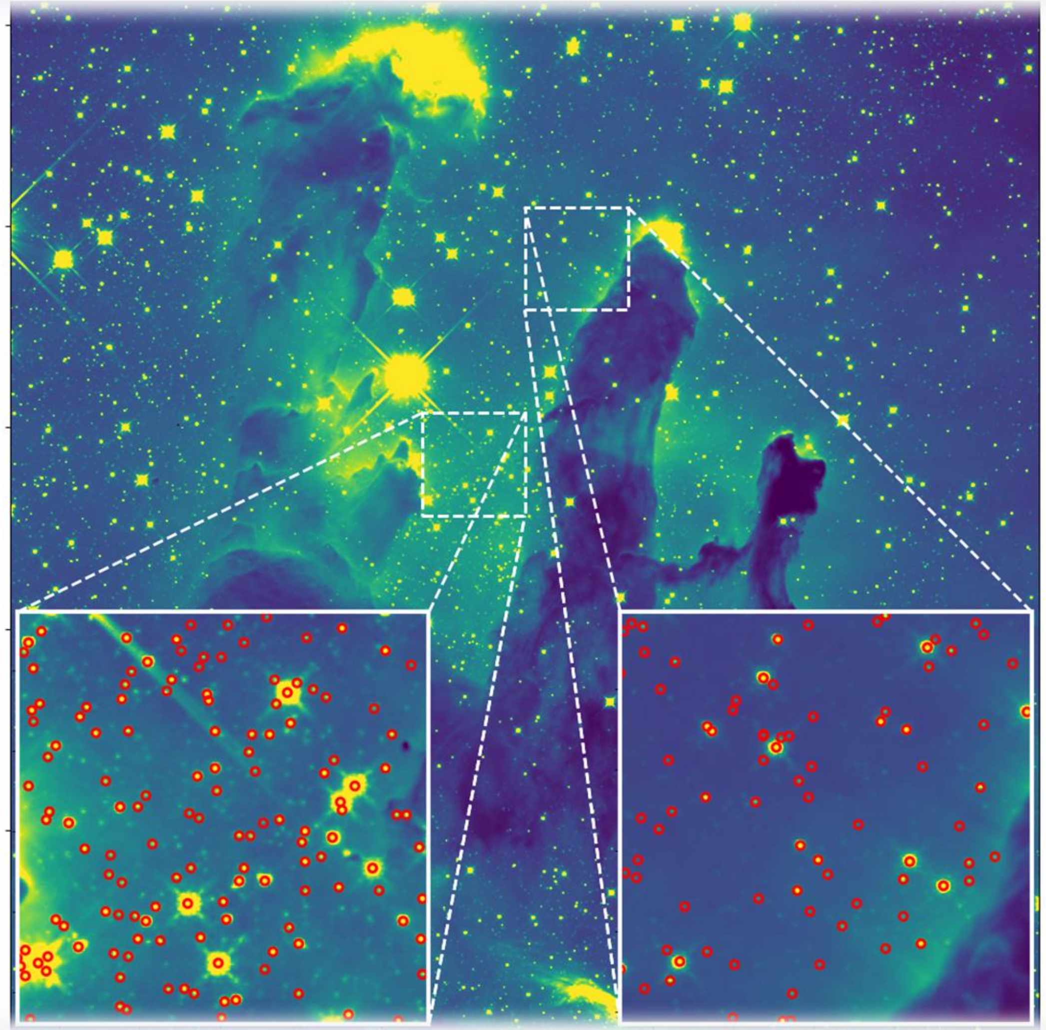


Can We Extend The Pipeline for Other Wavelengths? (Localization)

- Trained and tested with MeerLICHT data
  - 'ASID-Light: Fast Optical Source Localization'; [arXiv: 2202.00489]; A&A (A109, 2022); F. Stoppa et.al
- Automatic rejection of CR contaminants, satellite trail.
- Try transfer learning with Hubble data
  - Hubble PSF: 0.11 arcsec, MeerLICHT telescope PSF: 2-3 arcssec.
- Also tested for WISE data.

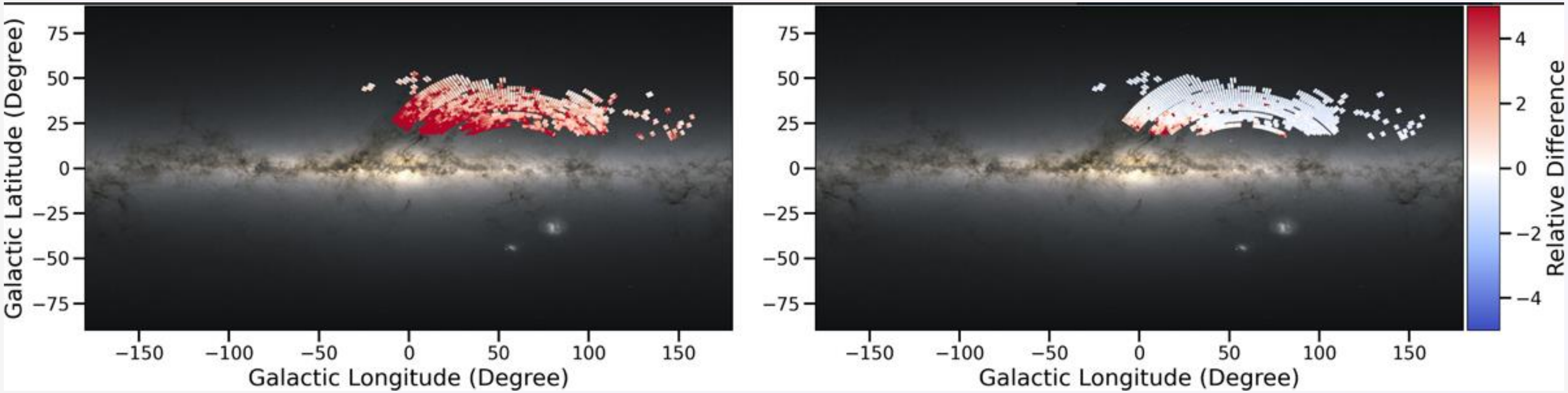






Can We Extend The Pipeline for Other Wavelengths? (Characterization)

- Once detected proceed to classify stars and galaxies;
- 'ASID-C: Star-Galaxy Classification'; [arXiv: 2307.14456]; A&A (A109, 2023); F.Stoppa et.al.,
- Better performance than SourceExtractor at high stellar dense region
  - Better calibration of classification probability, less overprediction of galaxies



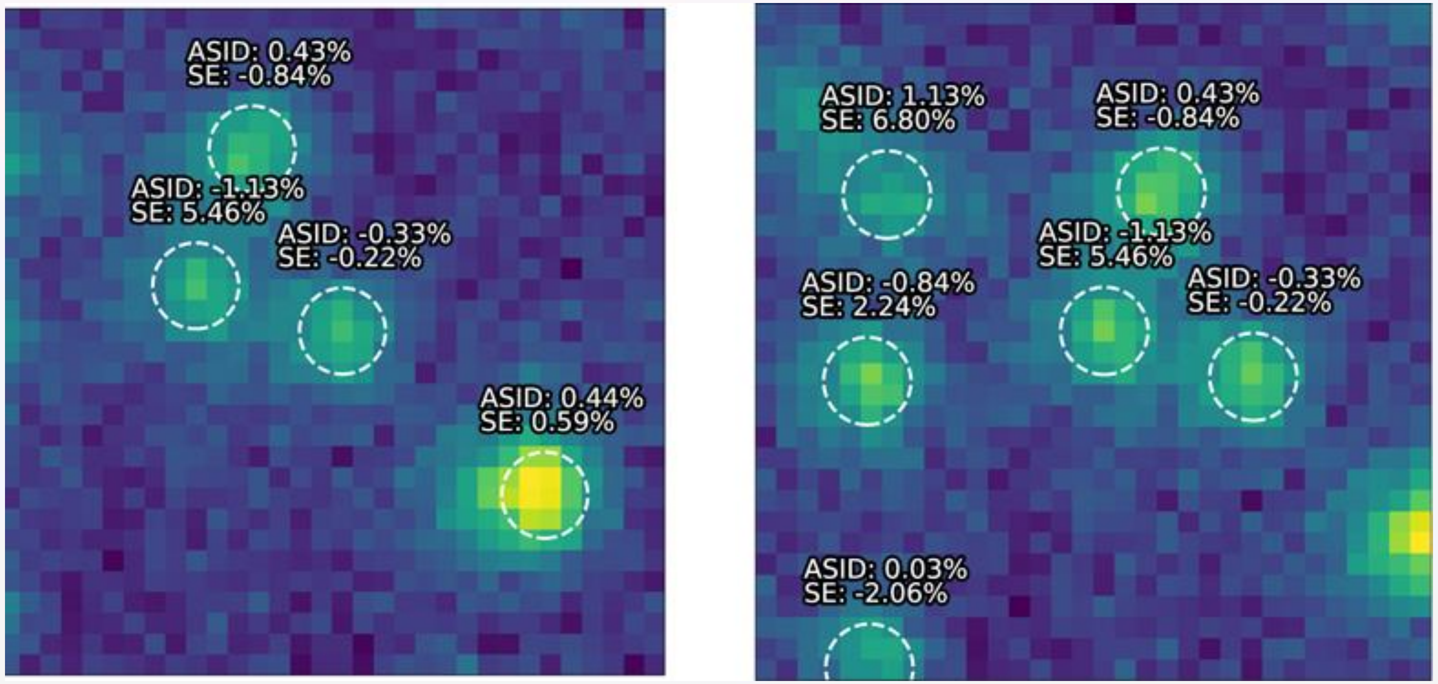
SourceExtractor

ASID-C

Relative difference of actual and estimated number of galaxies.

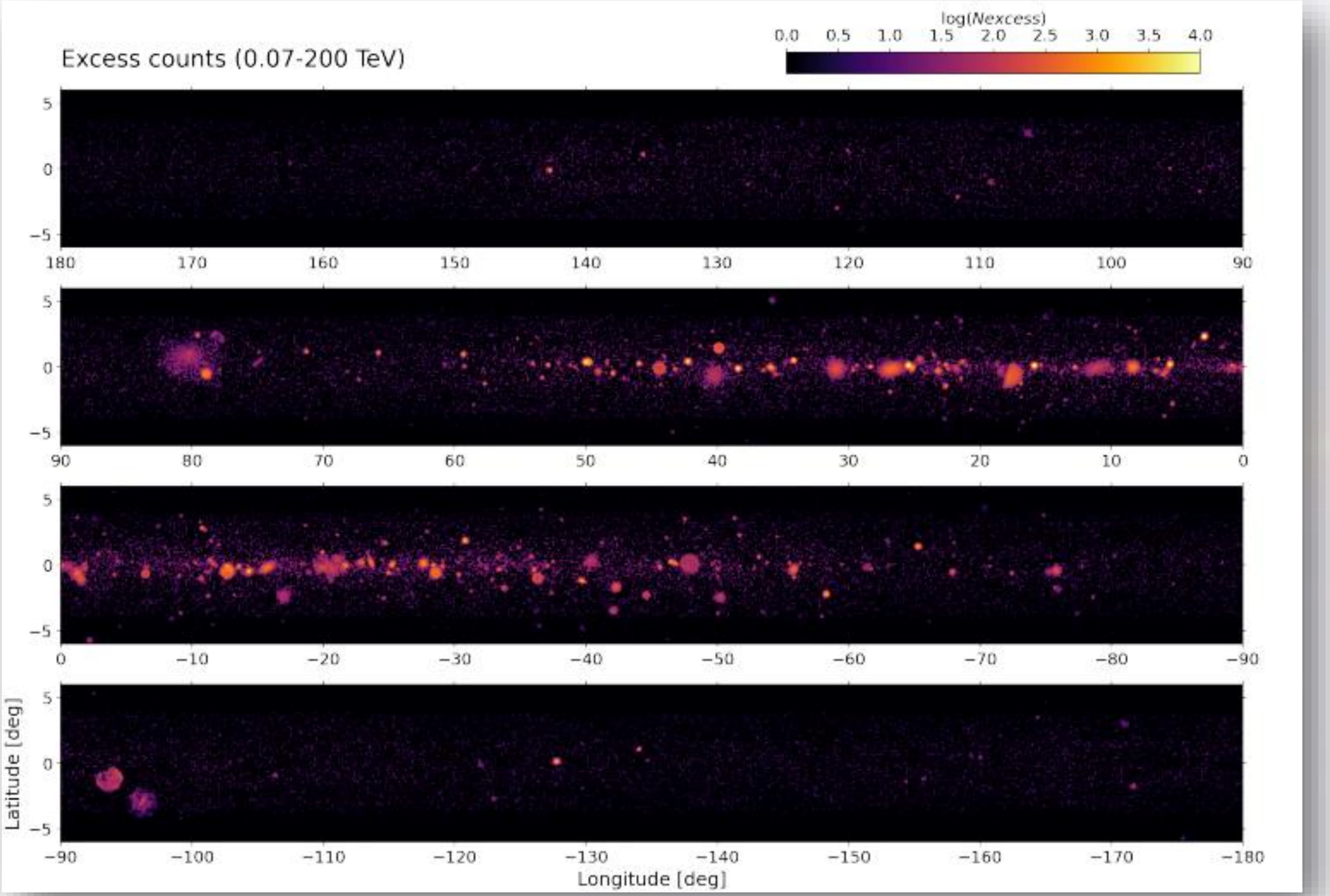
Can We Extend The Pipeline for Other Wavelengths? (Characterization)

- Once localized, estimate flux with uncertainties (single band image cutout).
- Two step network; Mean Variance Estimator Network
- 'ASID-FE: Flux Estimation & Uncertainty Characterization'; [arXiv: 2305.14495]; A&A (A108, 2023); F. Stoppa et.al.,
  - Performs better in crowded field compared to source extractor; Well-calibrated uncertainty



Predicted flux percentage error at two different levels of crowdedness between ASID-C and Source Extractor

Can We Apply This for CTA Simulated Data (Characterization)?



CTA Galactic Plane Survey (GPS);

Observation of the galactic plane with CTA telescope in the inner latitude region $|b| < 6°$

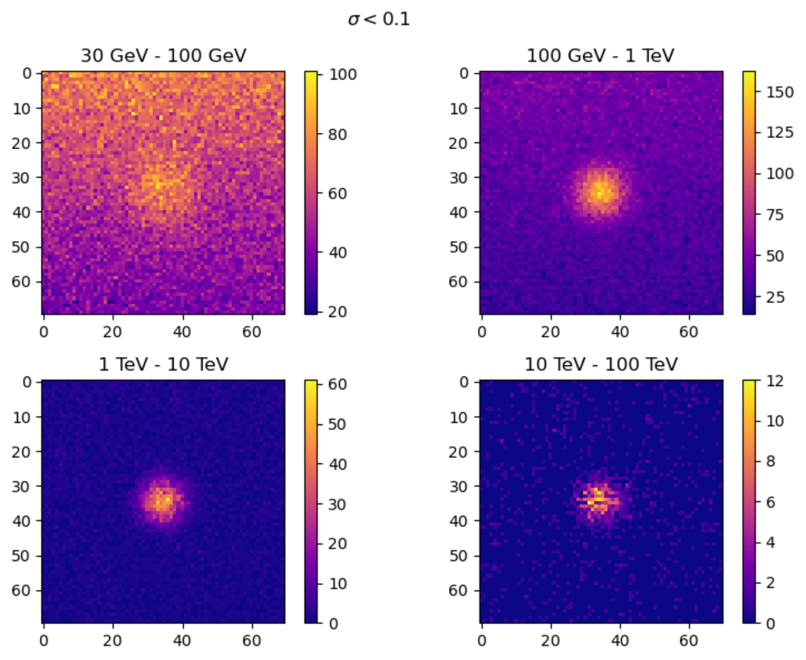
Total observation of 1620 hours over 10 years.

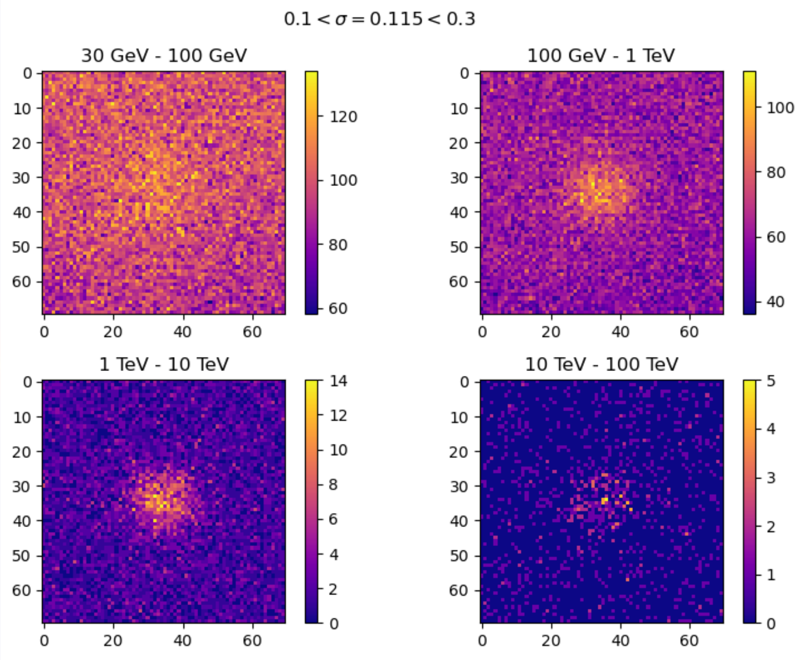
CTA Source Characterization (Example Simulated Sources)

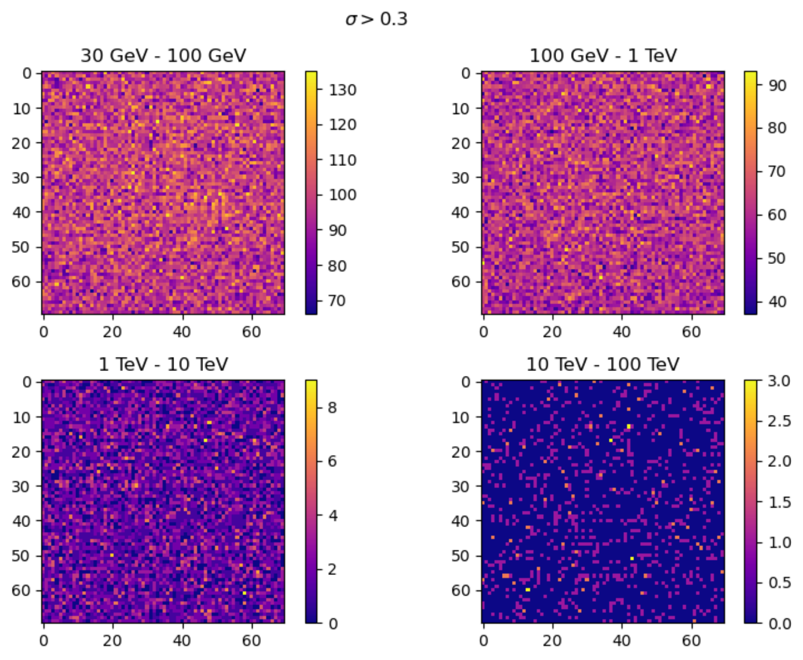
C0: $\sigma < 0.1$

C1: $0.1 < \sigma < 0.3$

C2: $\sigma > 0.3$





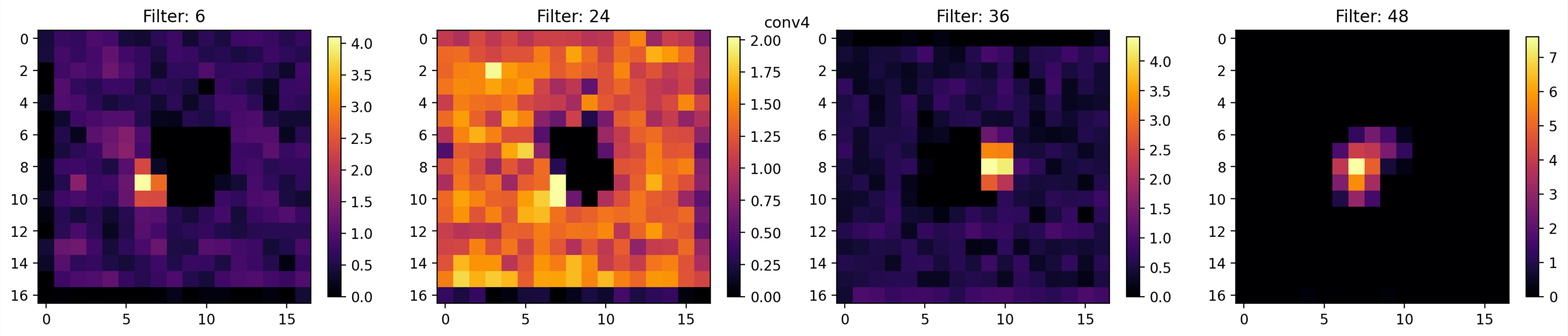


Target: Classify them based on their extensions

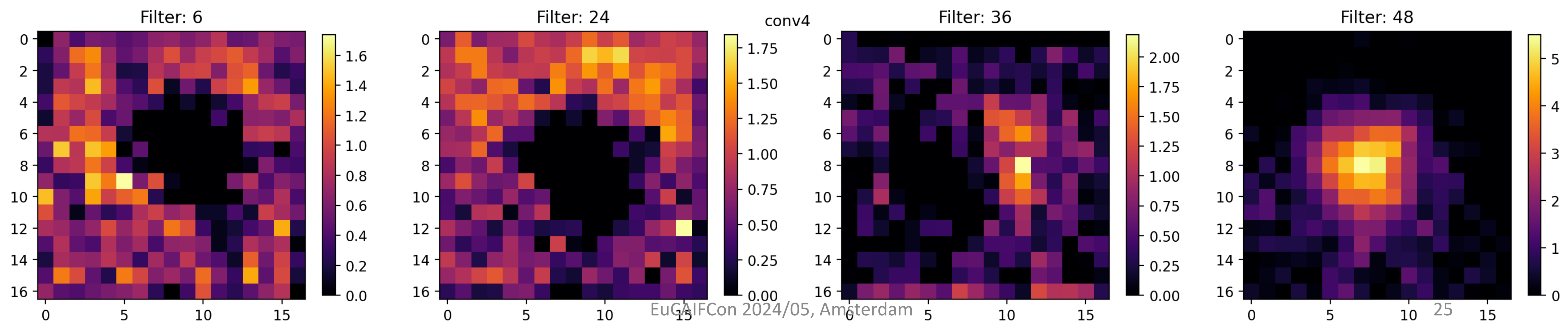
CTA Source Characterization (Network Activation Maps)

What parts of an image were used in different filters?

For a 'C0' Source

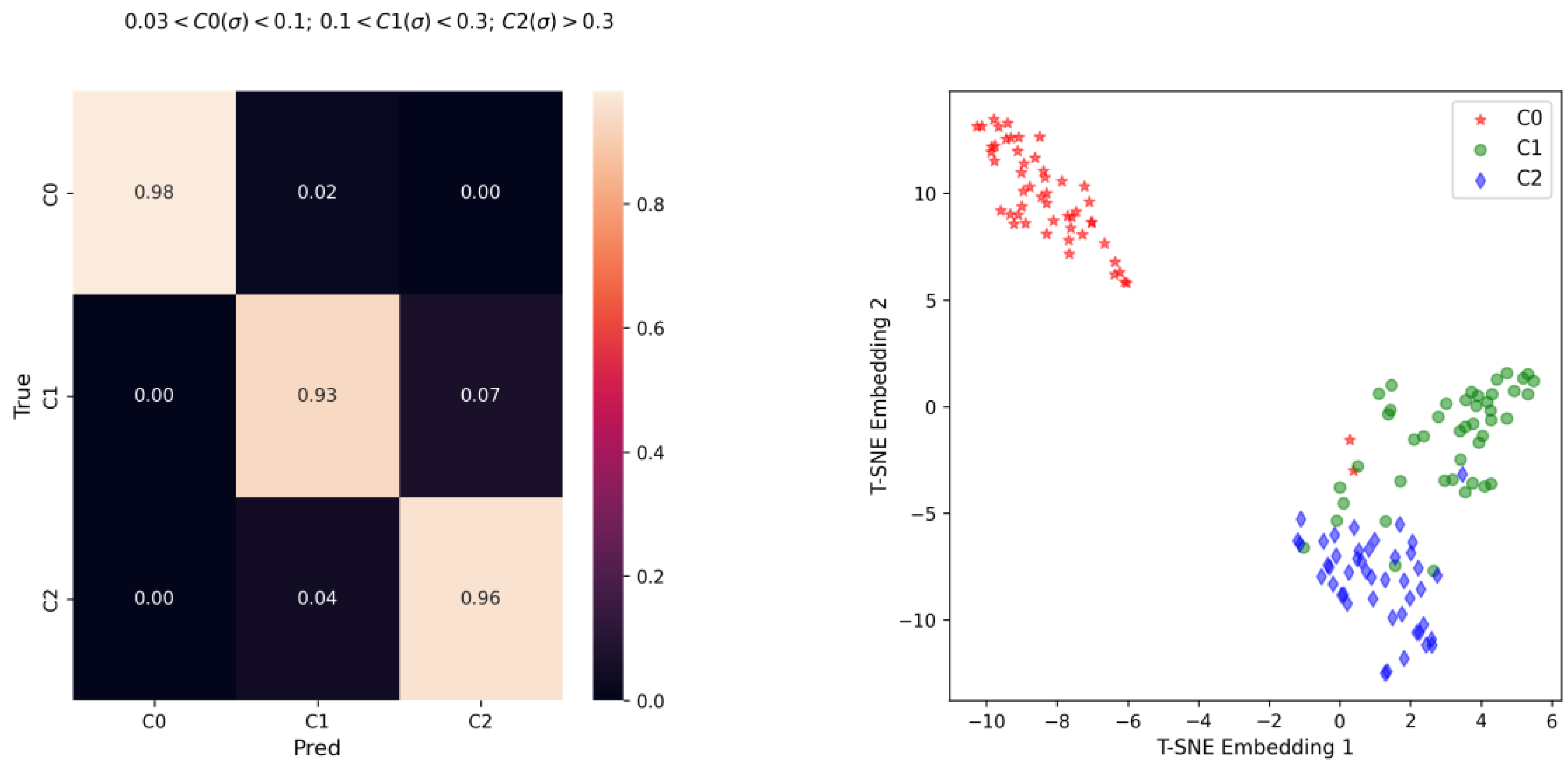


For a 'C1' Source



CTA Source Characterization (Extension Classification Results)

97% classification accuracy between point and extended sources.



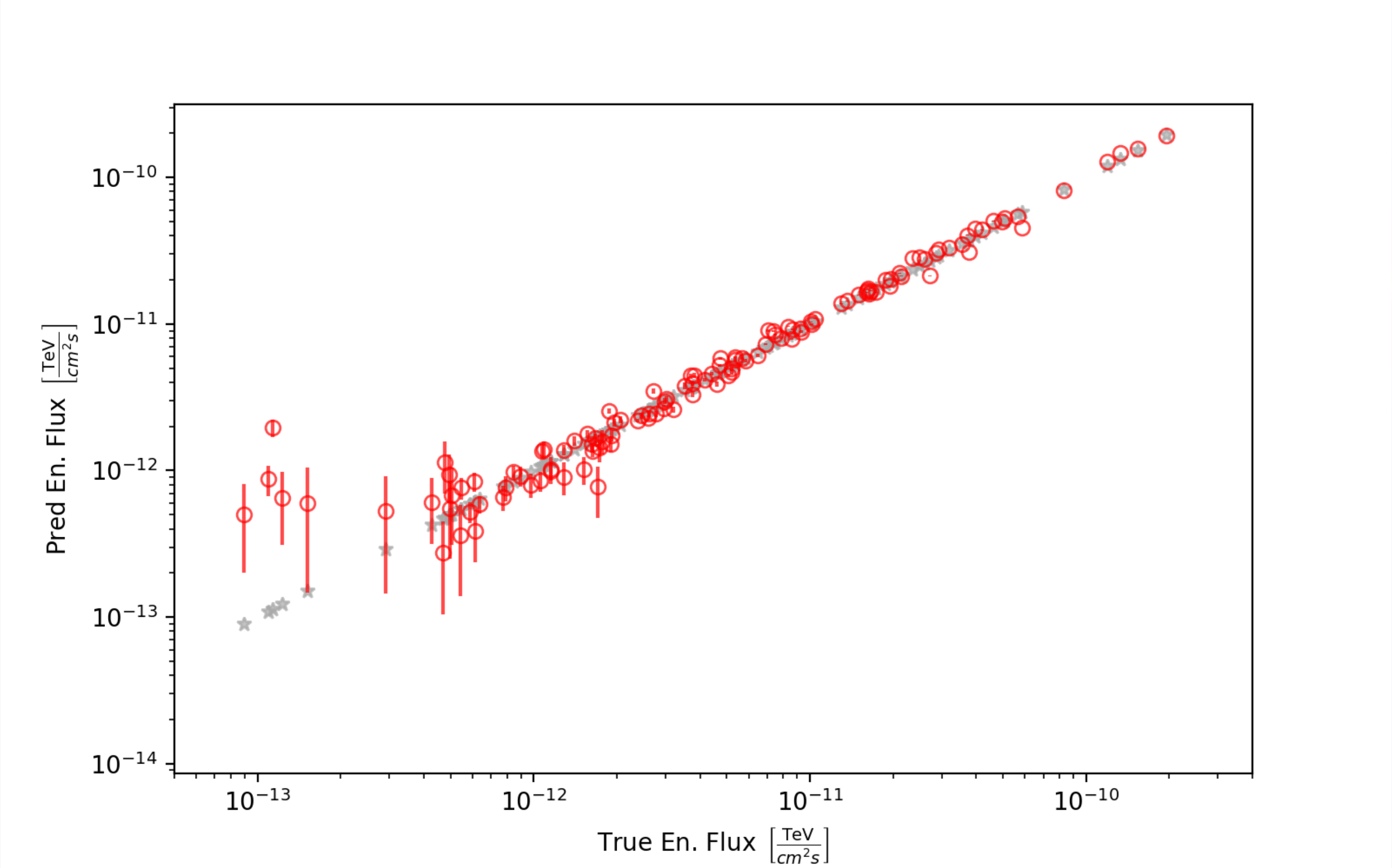
(a) CF Matrix; Rows: True labels; Cols: Predicted labels. (b) 2-D TSNE embeddings from fc1 layer (Fig. 3).

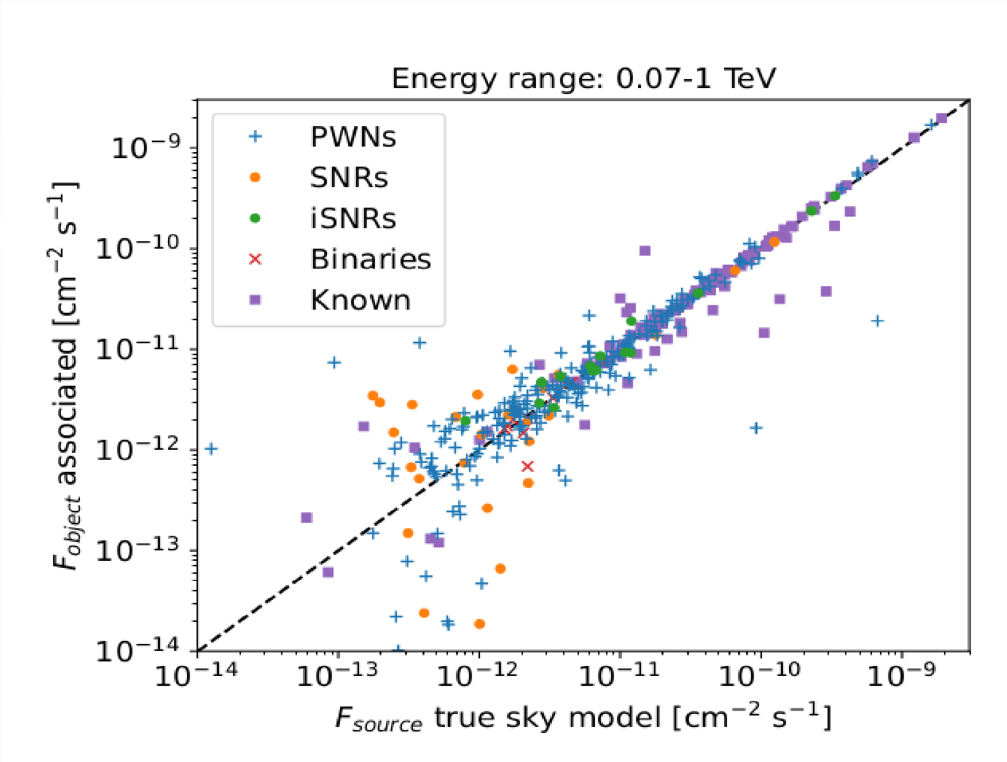
Reported in ICRC 2023,

PoS ICRC 444, 599

CTA Source Characterization (Flux Estimation)

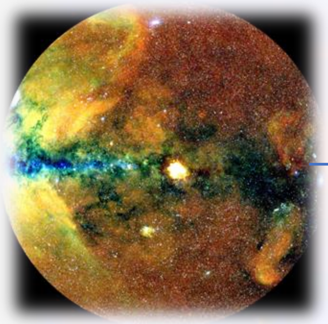
- Preliminary results: Given only point sources we could achieve better performance than published results.
- Left: Results from the test set in our calculation; Integrated flux in the range 70 GeV – 1 TeV.
- Right: Comparison with likelihood calculation [Plot from CTA GPS paper;].





Analysis Pipeline (Current Status & Future Prospects): Broad Overview

Input Data



Detection + Localization

Characterization of Individual Sources

- Flux Prediction
- Class Labels
- Source Extension
- ….
- ….

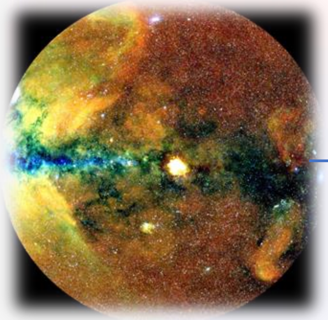
Analysis Pipeline (Current Status & Future Prospects): Broad Overview

Can we remove background?

See Poster (ID 82): 'BGREM: Background Removal of Astronomical Images with Diffusion Model'; Rodney Nicolas et.al.,

Does it help the source detection pipeline?

Input Data



Detection + Localization

Characterization of Individual Sources

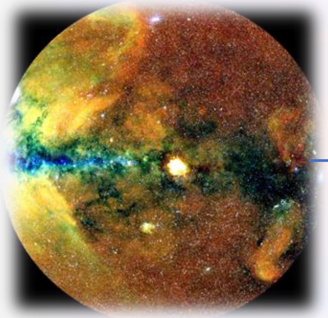
- Flux Prediction
- Class Labels
- Source Extension
- ….
- ….

Can we remove background?

See Poster (ID 82): 'BGREM: Background Removal of Astronomical Images with Diffusion Model'; Rodney Nicolas et.al.,

Does it help the source detection pipeline?

Input Data



Detection + Localization

Characterization of Individual Sources

- Flux Prediction
- Class Labels
- Source Extension
- ….
- ….

Can we generate high-quality data? Replace the simulator with DDPM network.

Little More….

- Possibility of collaboration? If our pipeline helps or if you want to modify for your own data ☺
- Possibility of recent graduates to apply for SMASH Fellowship; Marie-Curie Cofund Fellowship [2023-2028]

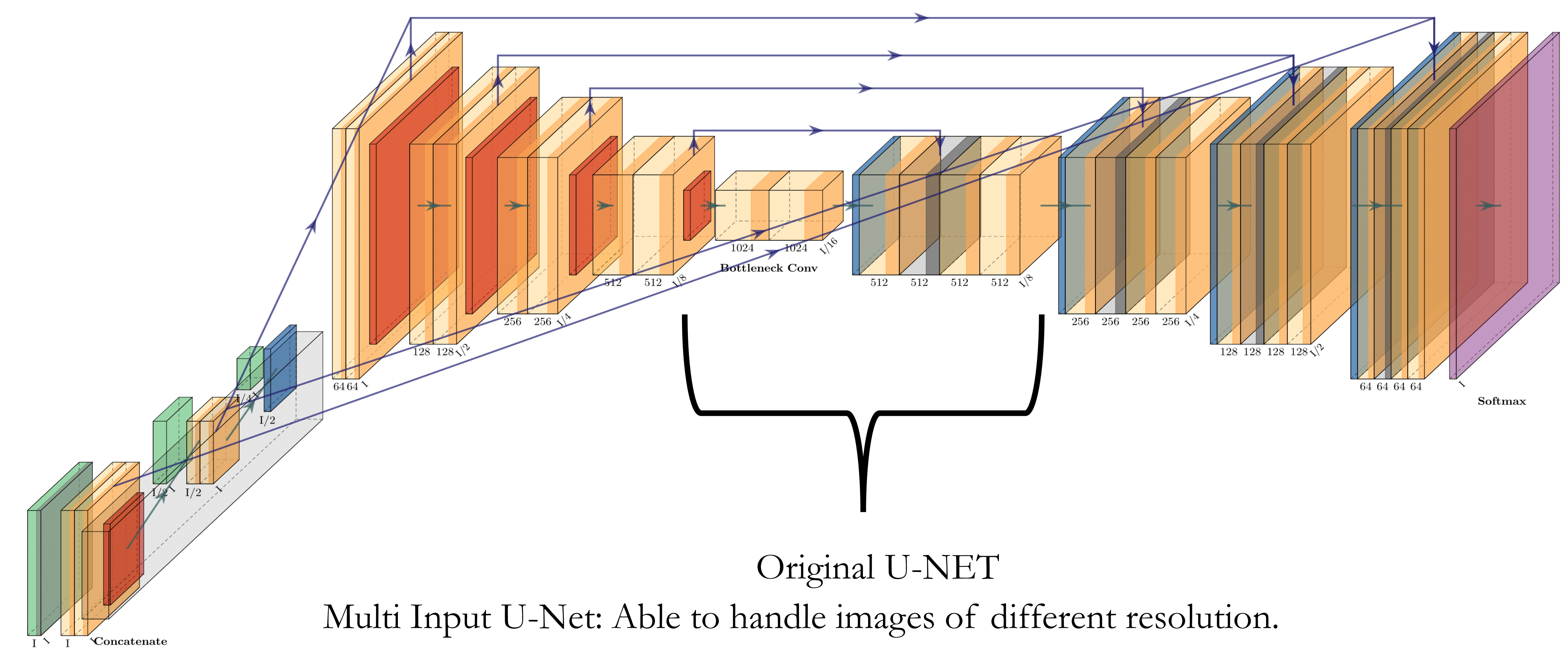
Check Here for
AutoSource-ID

Check Here

SMASH
machine learning for science and humanities postdoctoral program

Backups

Multi-Input U-NET Structure



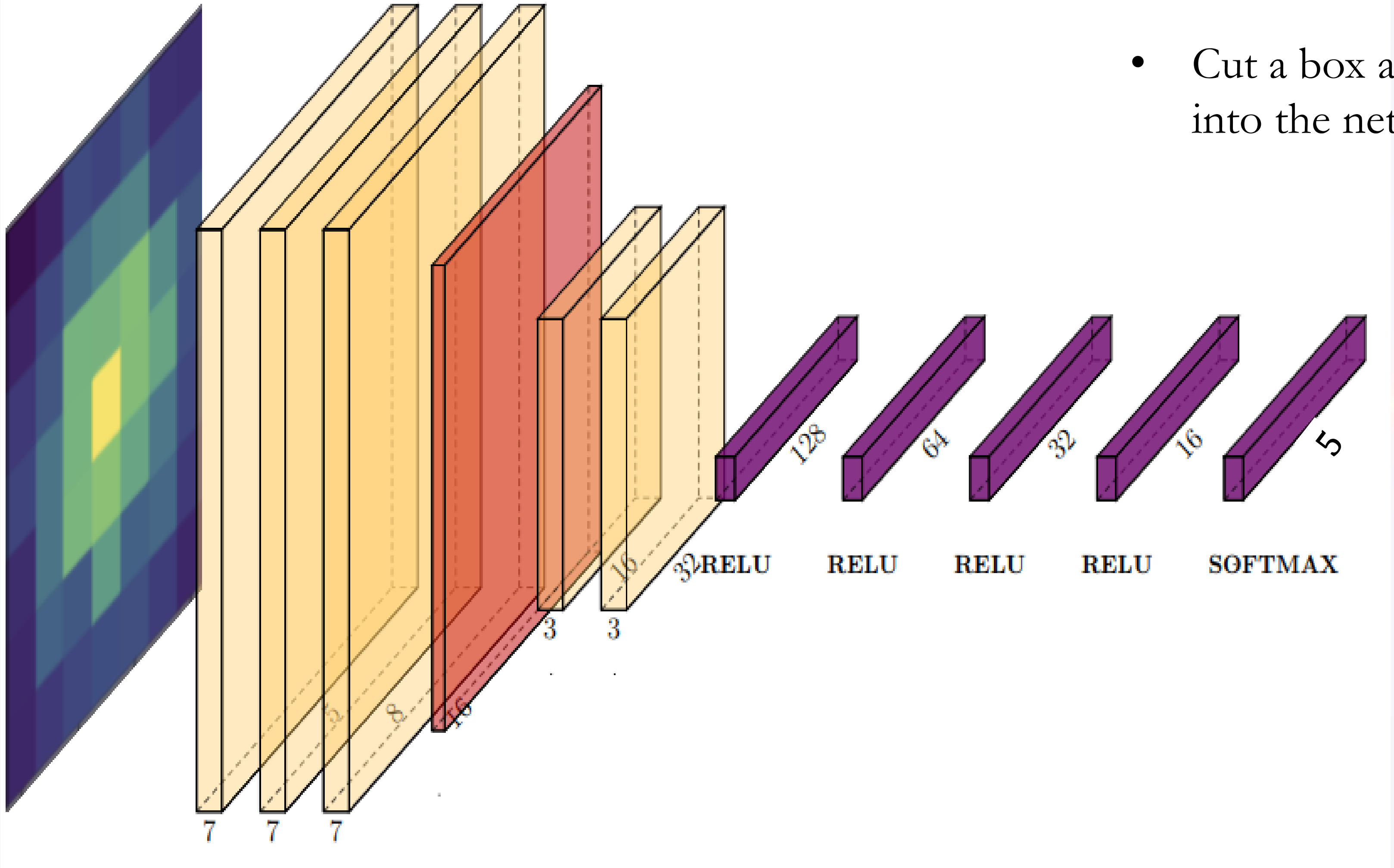
Original U-NET

Multi Input U-Net: Able to handle images of different resolution.

Produces a binary mask (1: Source, 0: Rest), Same resolution as the highest resolution input.

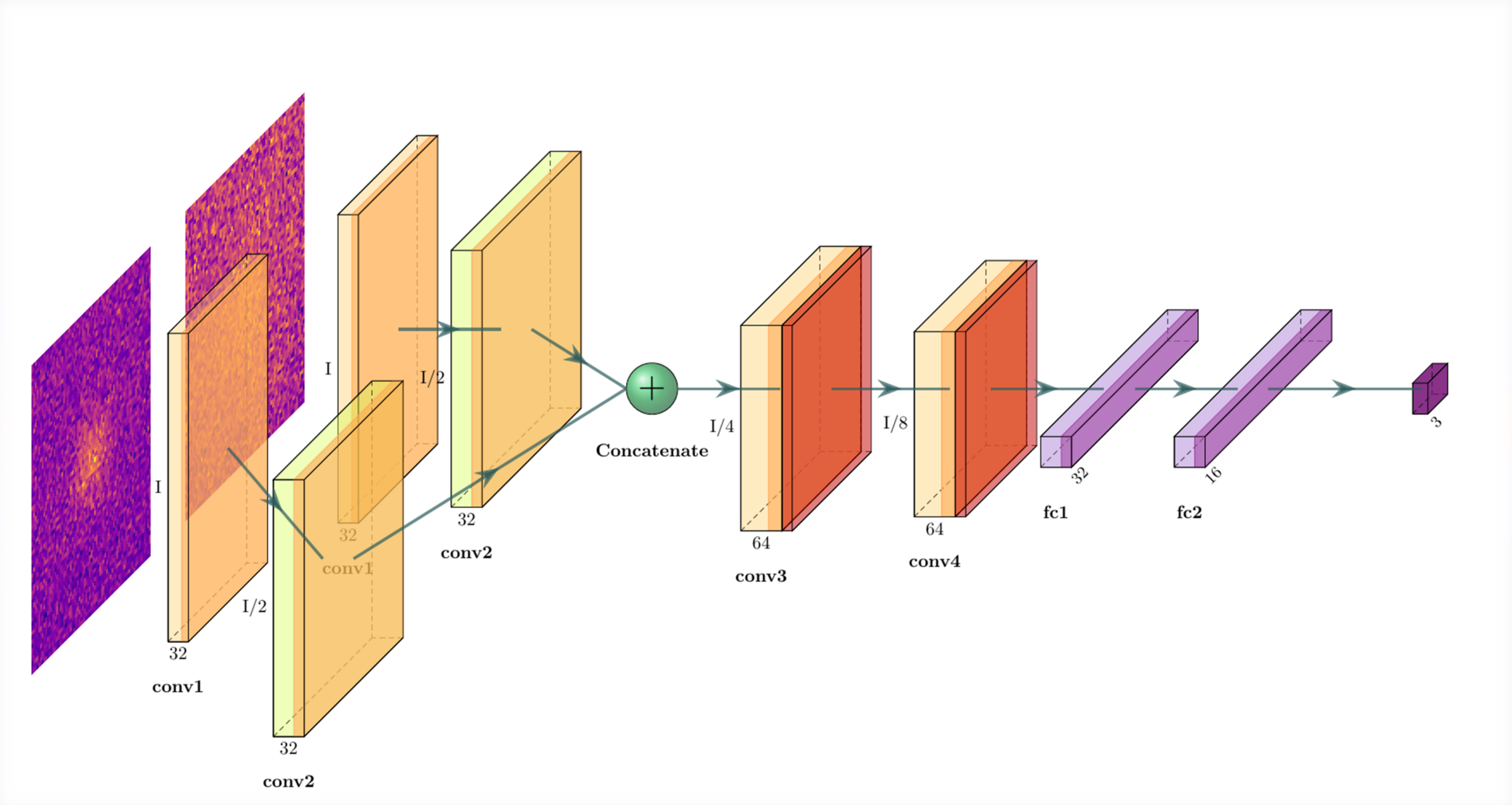
Classification Network

- Localized sources are then acting as inputs for a separate classification network.
- Cut a box around the predicted location and feed into the network.

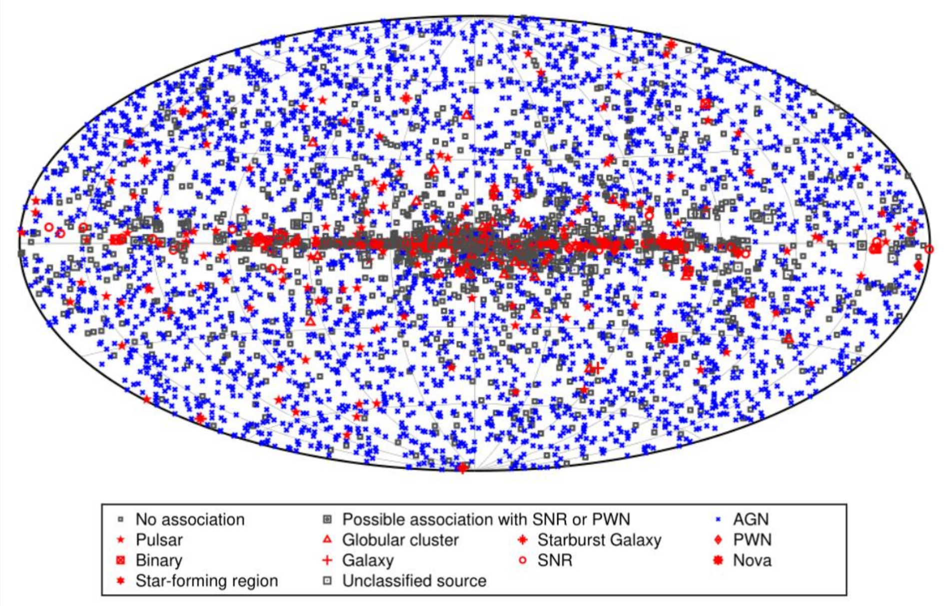


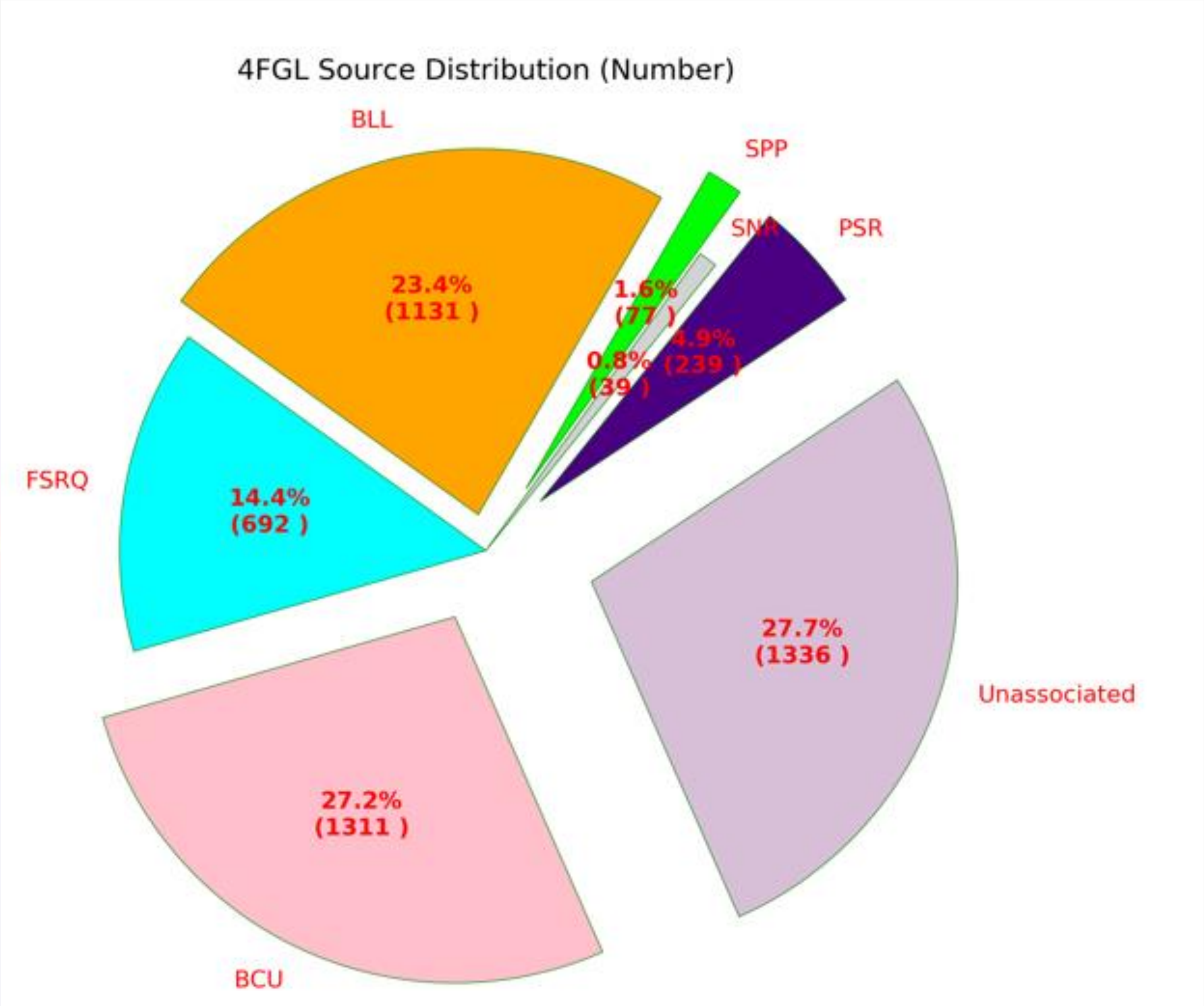
- Classification Network is a 3D CNN.
- Input shape (W, H, 10, 6).
- 10 years, 6 energy bins.
- Class-imbalance problem

CTA Source Classification-Network



4FGL Catalog; 8 years of Data, 5064 Sources

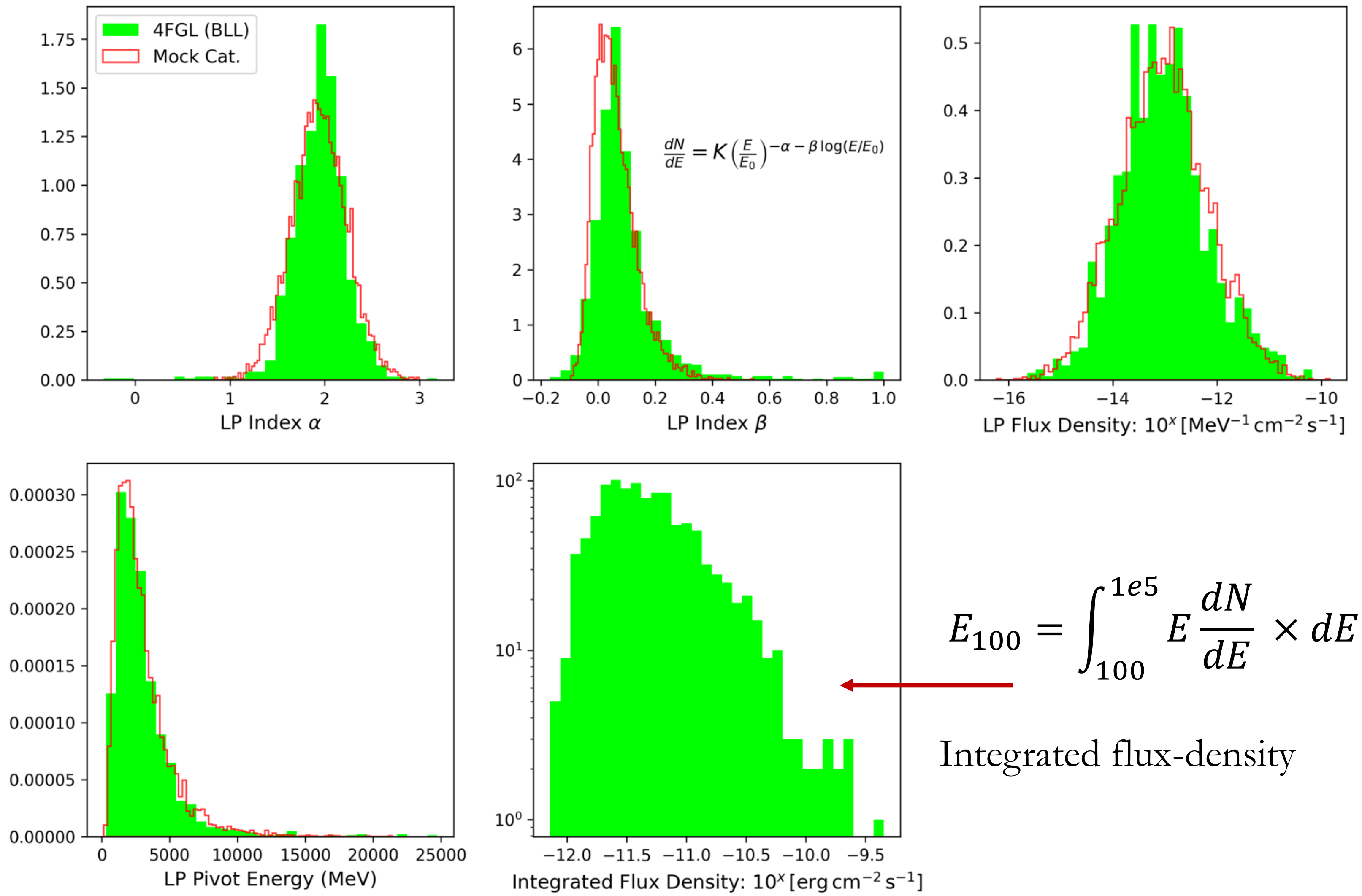




Ref: Fermi-LAT, 4FGL
ApJS 247, 33 (2020)

Simulation: Mock Catalog Generation: Example for BLLac: LP Parametrization

- Spectral shape:
  - Log Parabola
  - $\frac{dN}{dE} = K \left(\frac{E}{E_0}\right)^{-\alpha-\beta \log\left(\frac{E}{E_0}\right)}$
  - AGNs (BLLac, FSRQ, PWN, SPP)
- Distribution in Sky:
  - BLLac, FSRQ : Uniformly distributed over the whole sky.
  - PSR, PWN/SPP : Uniform distribution in longitude
  - Latitude distribution peaks at the plane.

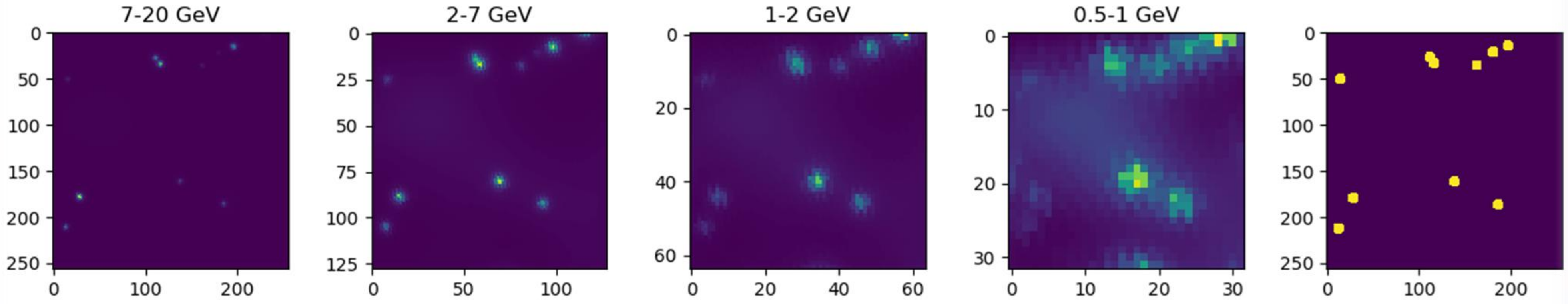


$$E_{100} = \int_{100}^{1e5} E \frac{dN}{dE} \times dE$$

Integrated flux-density

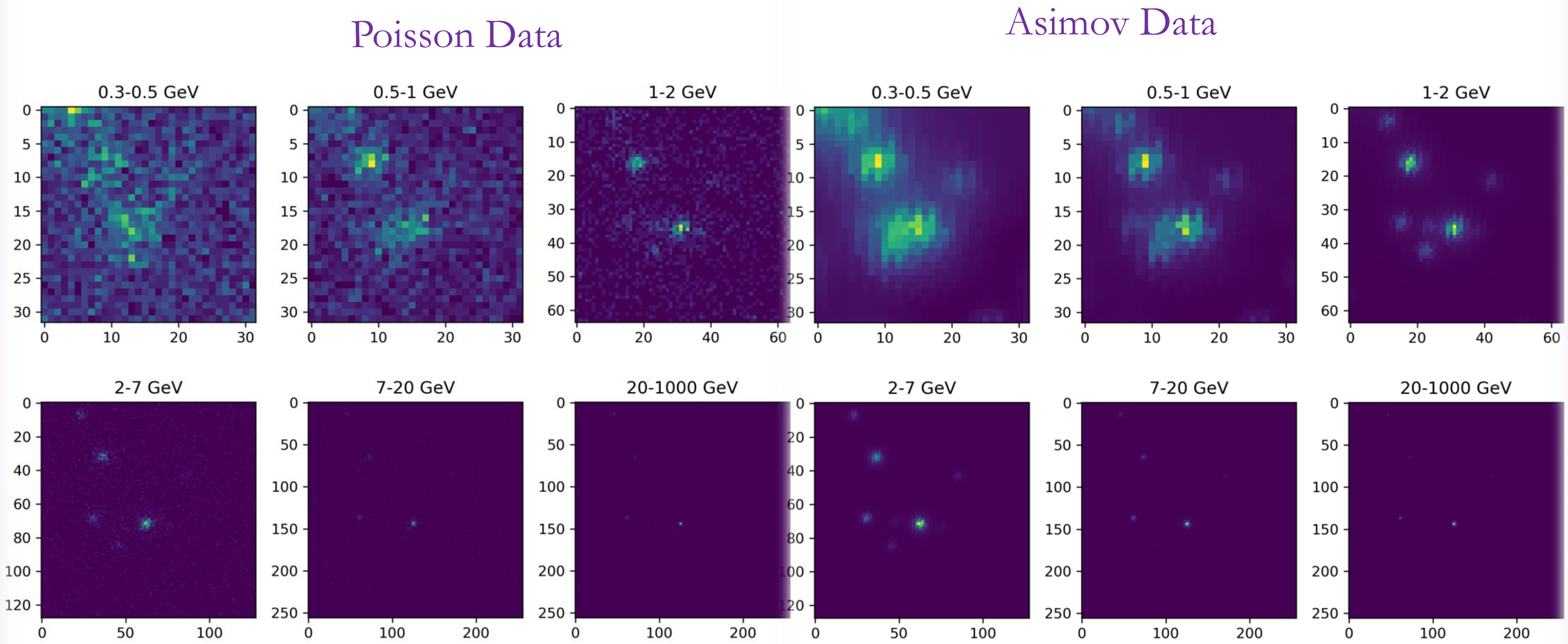
Training Data & Localization Scheme

- Images of full sky data in 6 energy bins [0.3 GeV - 1 TeV].
- Step1: Implement U-Net like algorithm. Segmentation task.
    - Each pixel is assigned with a label score (≈1, source pixel; ≈0, otherwise).

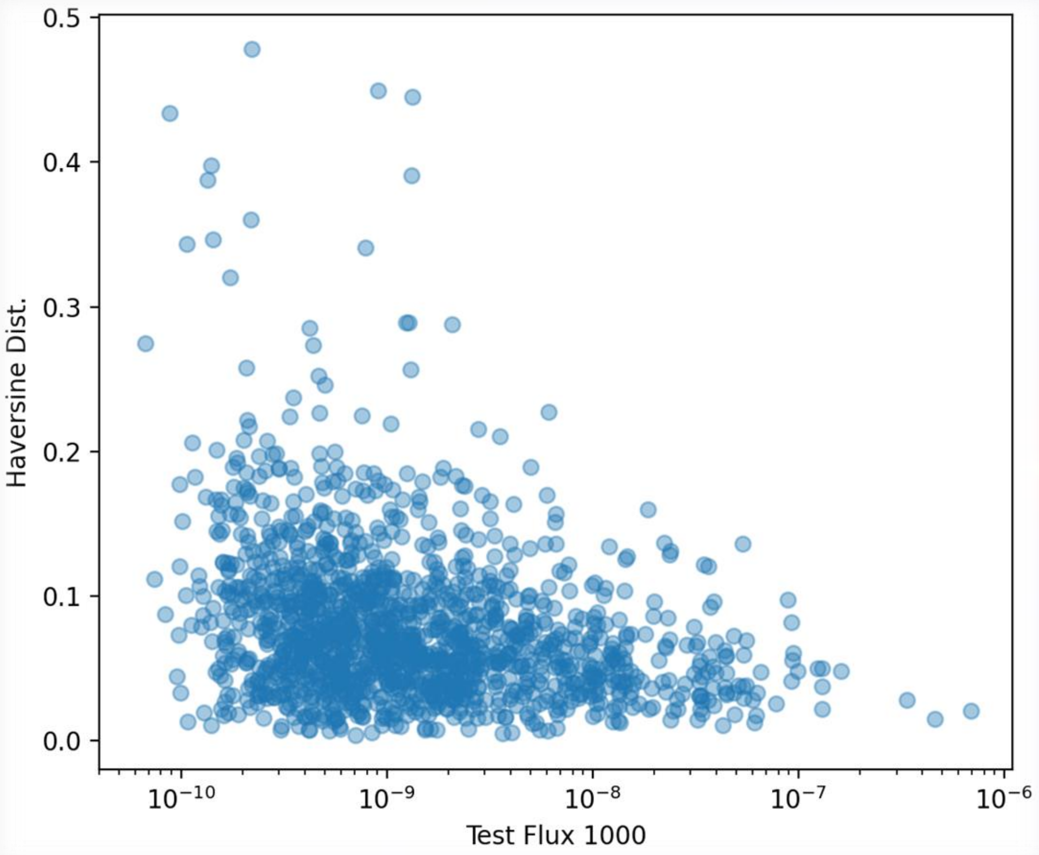


- Step2: Apply Laplacian of Gaussian (LoG)
    - Find the center of source pixels in (X, Y) and convert to (Lon, Lat).

Assimov & Poisson Patches:



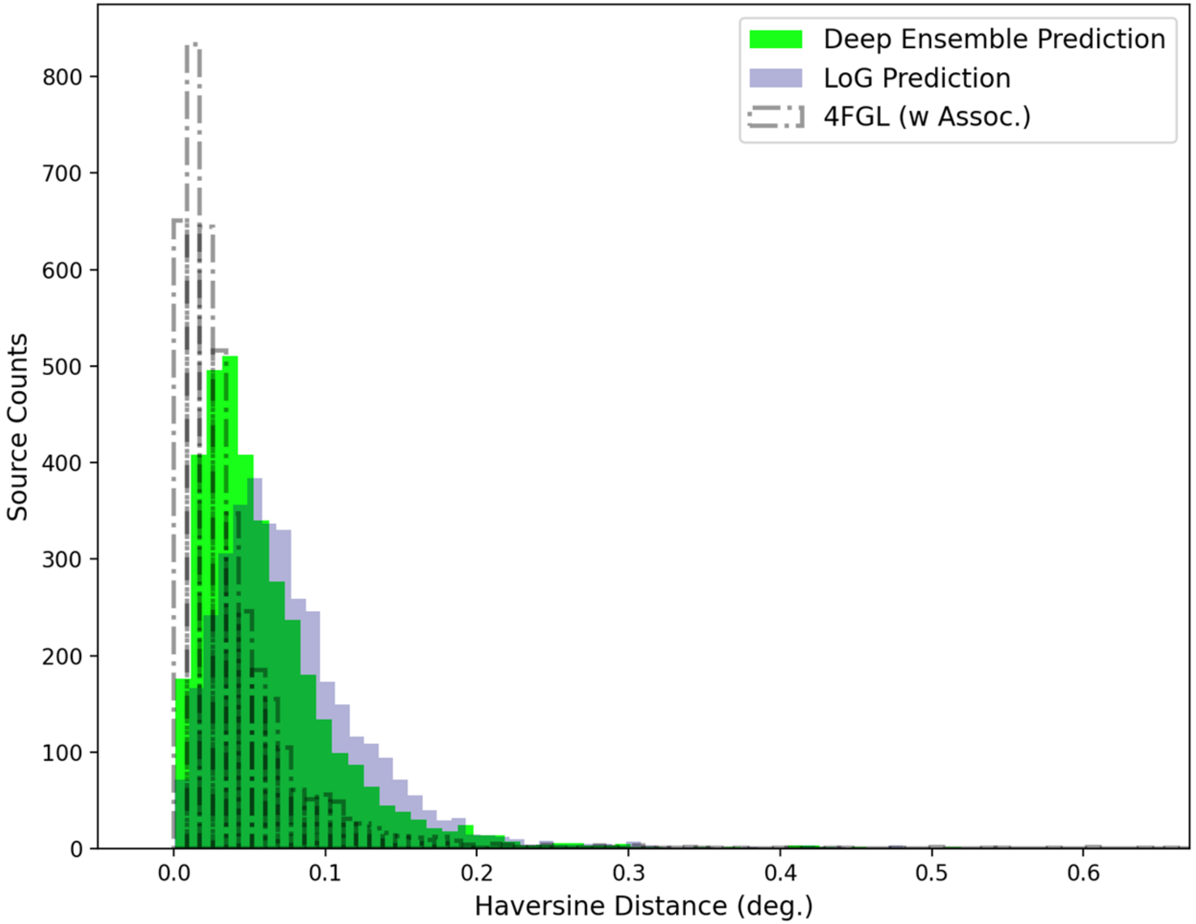
Dependence of Location Reconstruction on Source Flux



On the training set (simulated), we have flux information;

Further check that the brightest sources are indeed predicted with better accuracy. X: Integrated Flux; Y: Hav. Dist.

Comparison of Location Reconstruction



94% of the detected sources' locations are within 0.15° from the true locations

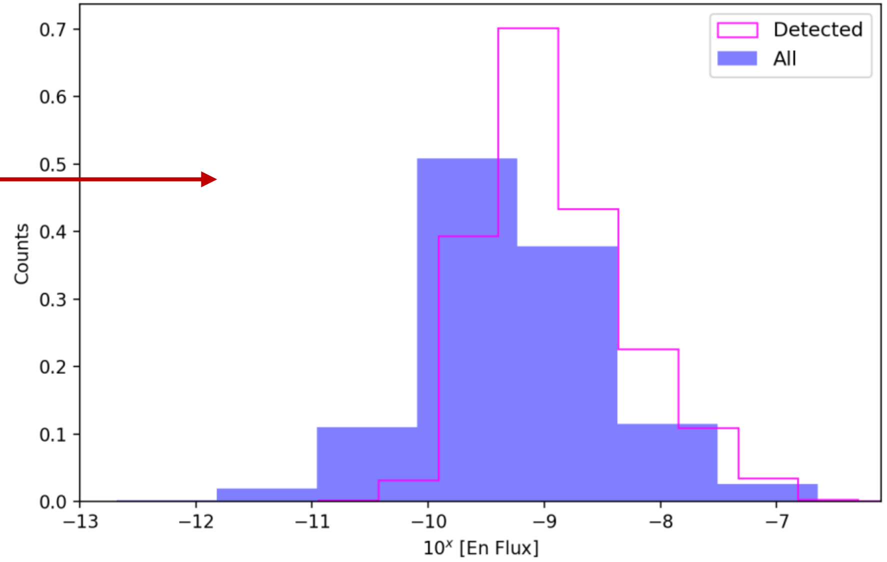
Also for the 4FGL 'associated' sources we show the haversine distance for:

4FGL location vs. Association location

Source characterization: Flux Estimation

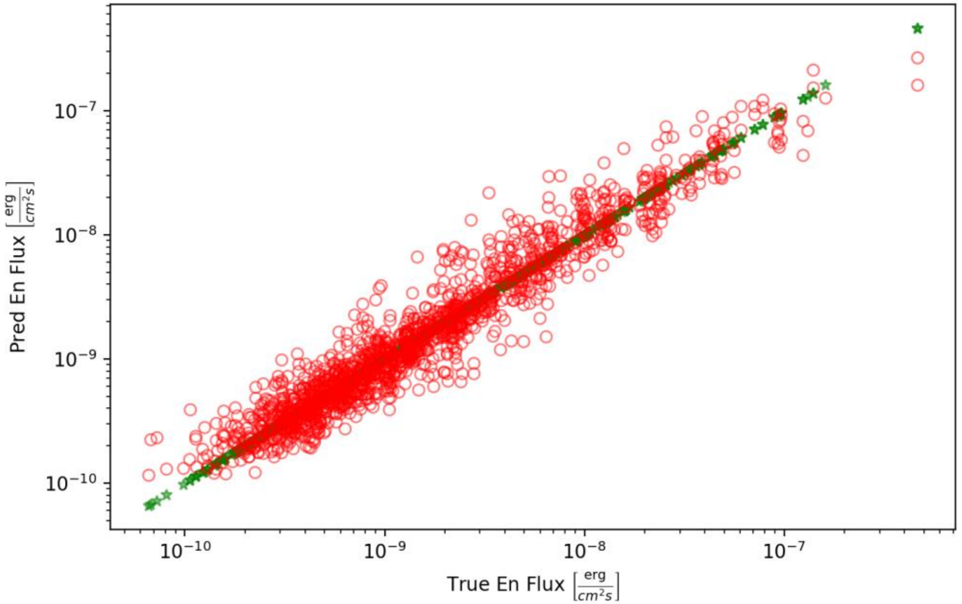
We follow up on the detected sources: Flux Estimation

Flux distribution of all sources vs Detected sources is shown here.



Use a simple CNN (7 layers) for the regression task.

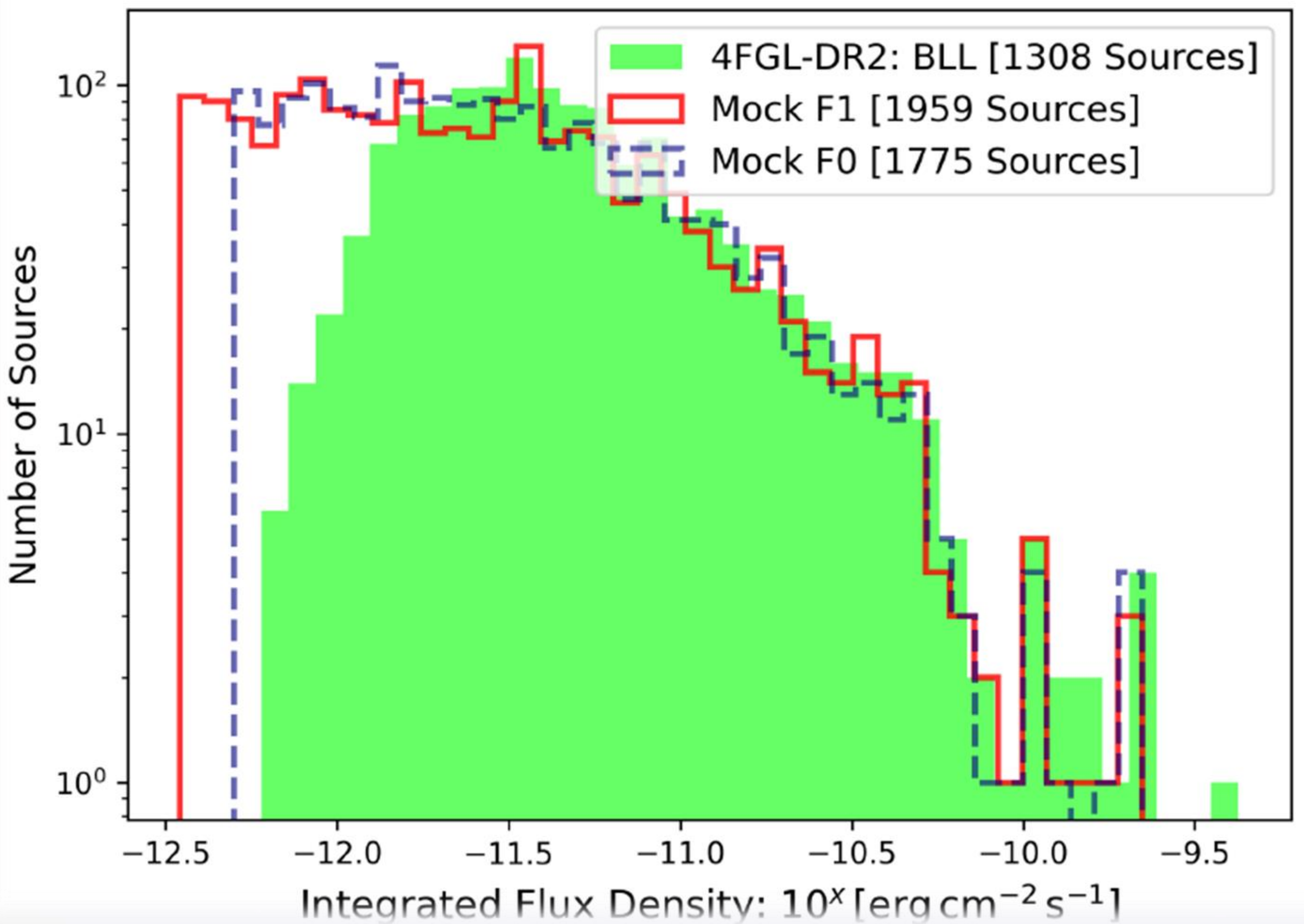
- Given an input image (patch with a source at the center) and corresponding integrated flux;
    - Network learns to regress to mean. Simple 'Mean Squared Error'.
- Outputs from the ensemble of networks are then used for uncertainty prediction.
    - Ongoing; Hyperparameter checks and
    - Uncertainty prediction.



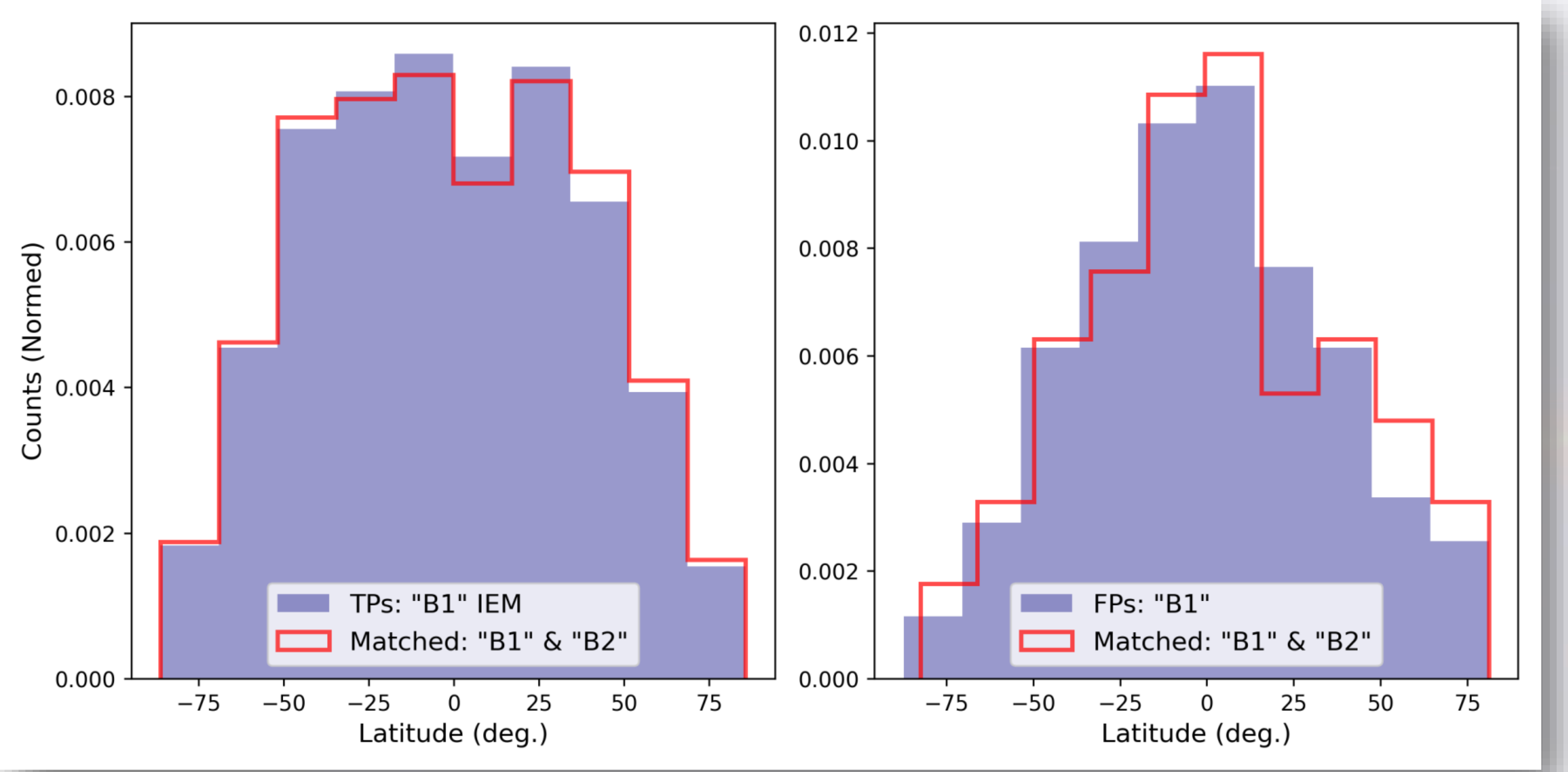
Current Status & Upcoming Tasks:

- Finalize the analysis pipeline and apply it to real data.
- Previously we applied to real data and Benoit helped to check with 4FGL association pipeline.
    - This led to refinement in location uncertainty calculation, SNR calculation, and eventually addition of Flux estimation part.
- Check the significance of the sources at the predicted location with Fermipy; Especially the ones that are 'new'.
    - Giacomo currently helping with this.
- Source separation capability check with MC simulation; If a bright source/faint source close by to an existing source;
    - Dima currently helping with this.
- Whether it is possible to recover the fainter sources (DR2 sources with significance between $5\sigma - 10\sigma$) by lowering the segmentation threshold?
- Create a new catalog and check with the association pipeline (Benoit will help).
- Any comments/suggestions are very welcome ☺

Example Luminosity for Mock Catalog (BLL):



Performance Comparison with Background Models:



Example Numbers:

Trained & Test Data both w 'B1':

Detected sources: 3353 (63%)

Trained with 'B1', Test with 'B2':

3157 out of 3353 sources were found.

Even if we use two different backgrounds, we can recover ~ 93% of the sources.

TP: True Positive, FP: False Positive

Mask Radius vs Network Performance (Detection)