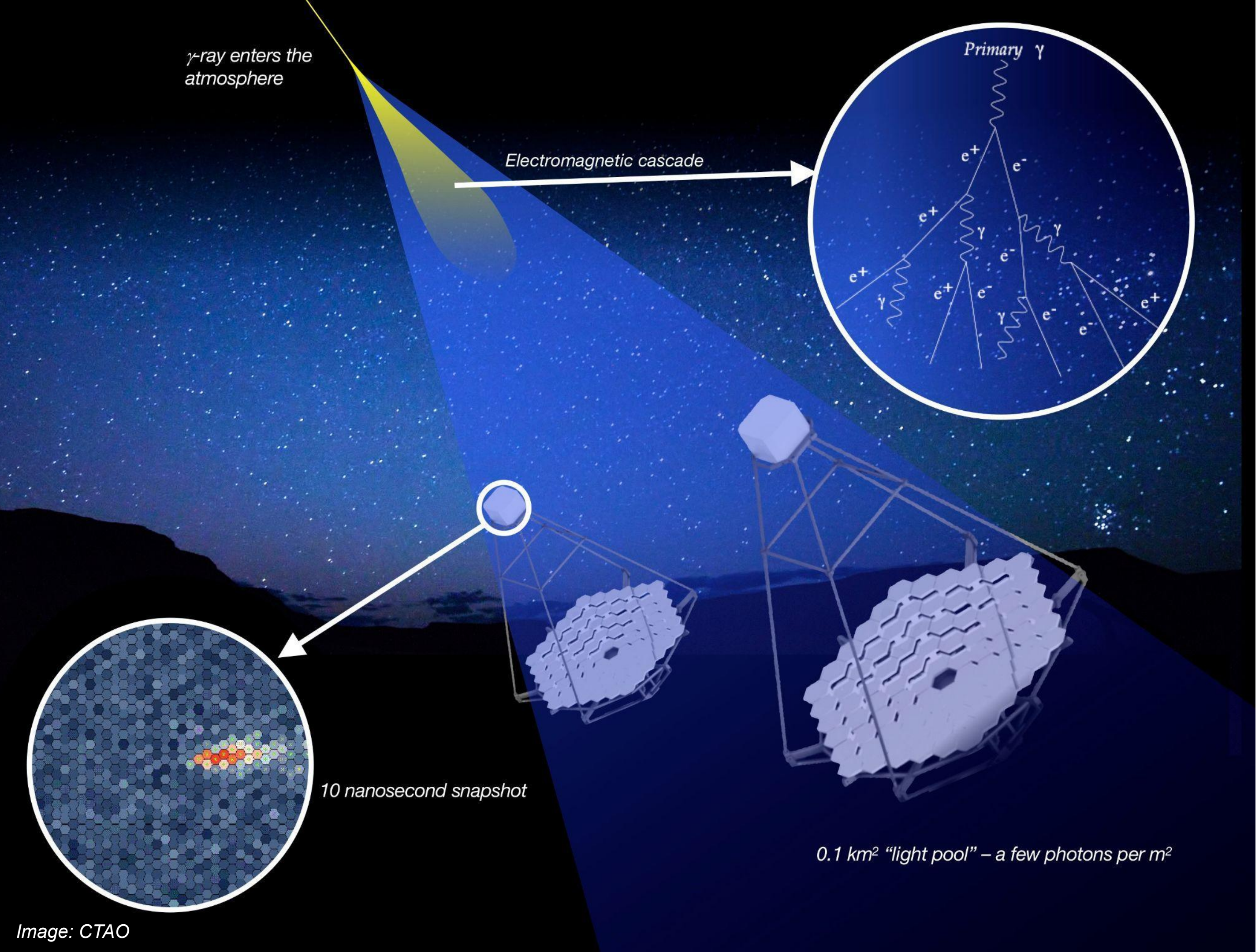
Ambra Di Piano

@UNIMORE @INAF/OAS Bologna

Photo: Daniel López
Overlay: The Matrix

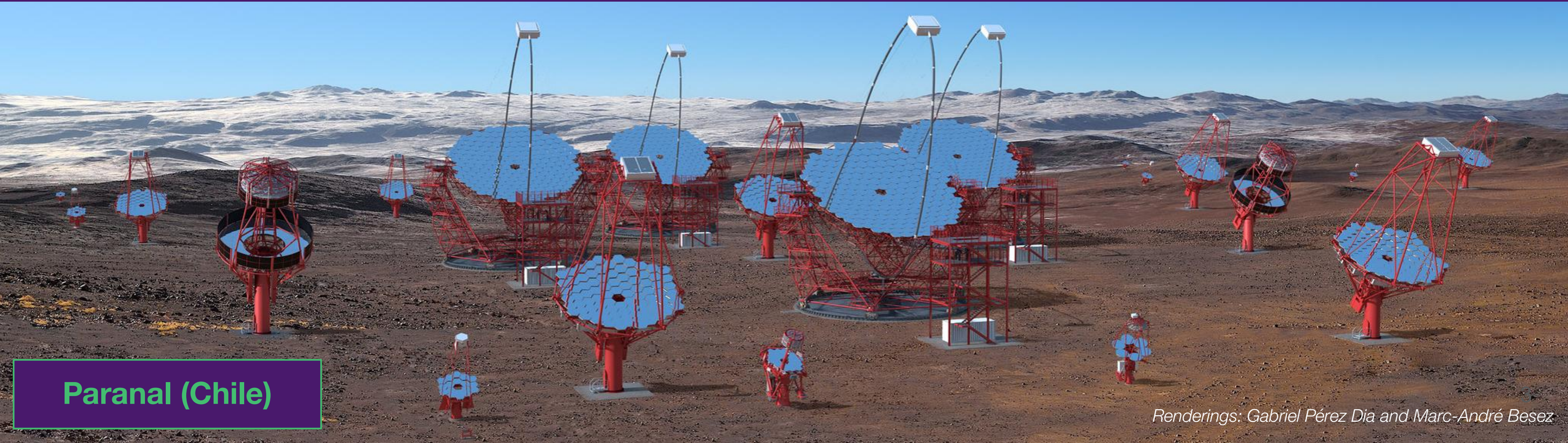


Cherenkov Light

- Visible
- Brief
- Faint

Background-dominated:
only 1 photon
every ~1000 hadrons


Roque De Los Muchachos (La Palma)


Paranal (Chile)
Renderings: Gabriel Pérez Dia and Marc-André Besez

overview

OTHER DEEP LEARNING EFFORTS

>> focus on reconstruction pipelines <<

CTLearn

γlearn

Photo: CTAO

CTLearn: Brill et al. 2023, GammaLearn: Vuillaume et al. 2024

overview

OTHER DEEP LEARNING EFFORTS

>> focus on reconstruction pipelines <<

CHEREN-ZOO talk by Angelo Adamo

Photo: CTAO

CTLearn: Brill et al. 2023, GammaLearn: Vuillaume et al. 2024

overview

Photo: Moritz Huetten

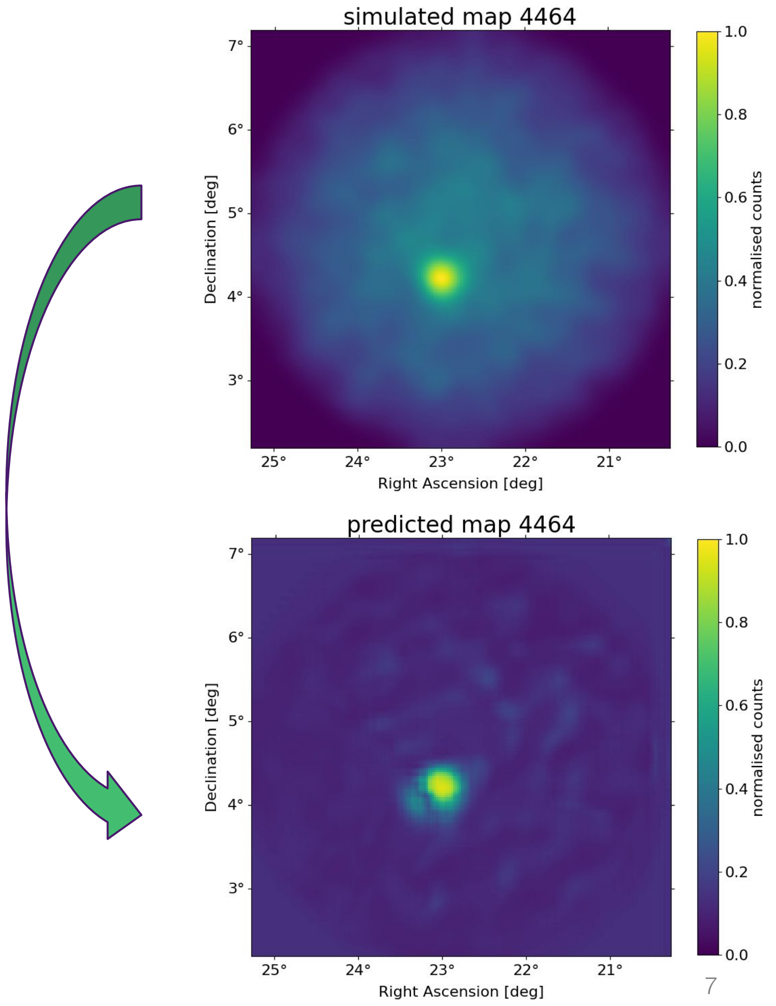
Context:
- High-level data (DL3)
- Real-time analysis
  - [1]ACADA - [2]Science Alert Generation
- Machine learning enhancements

Goals:
- **background-subtraction**
  - no background template
  - unknown target coordinates
- **candidates localisation**
  - no background template

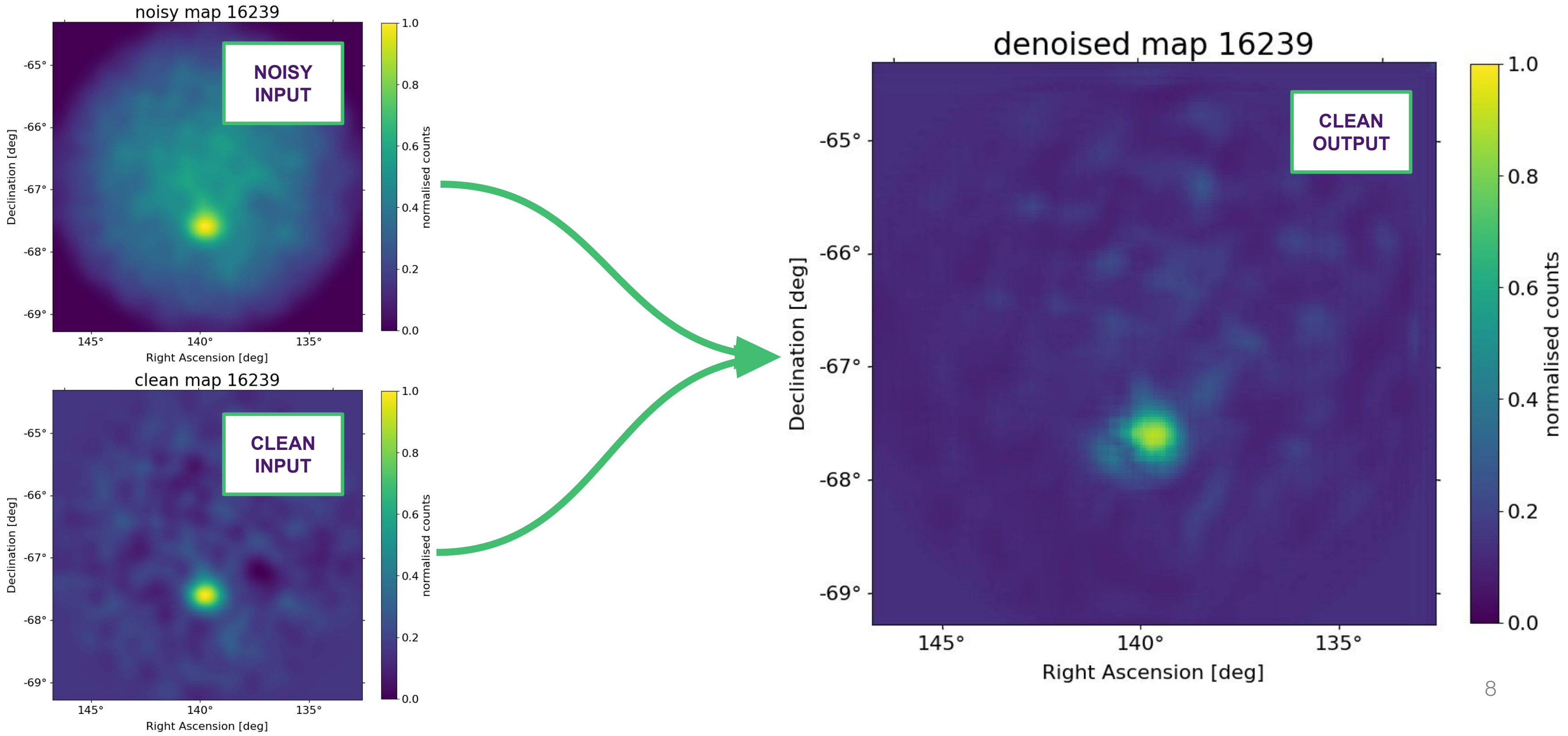
[1]Oya et al. 2019, [2]Bulgarelli et al. 2022

CNN-Cleaner

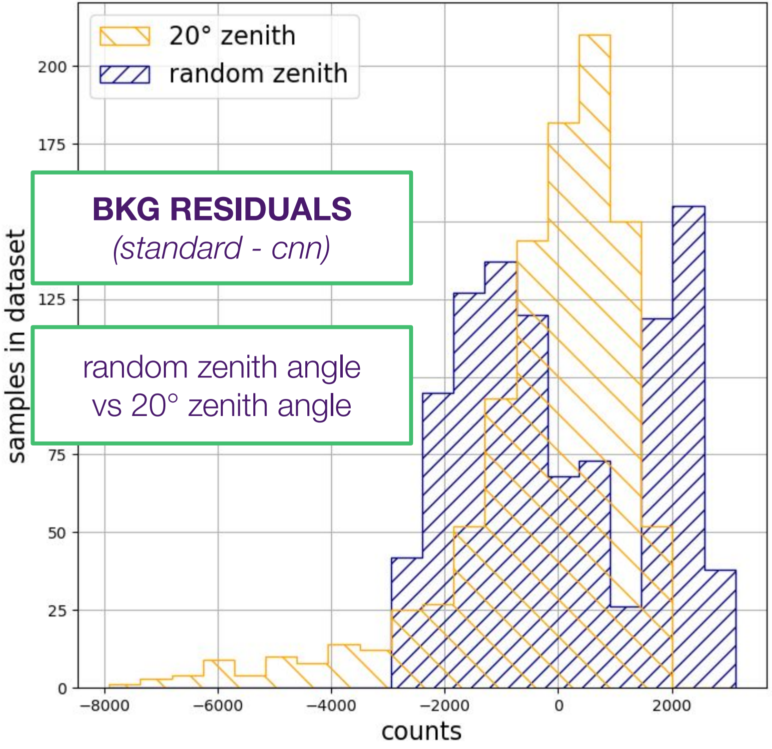
a DL model for background subtraction in real-time



datasets

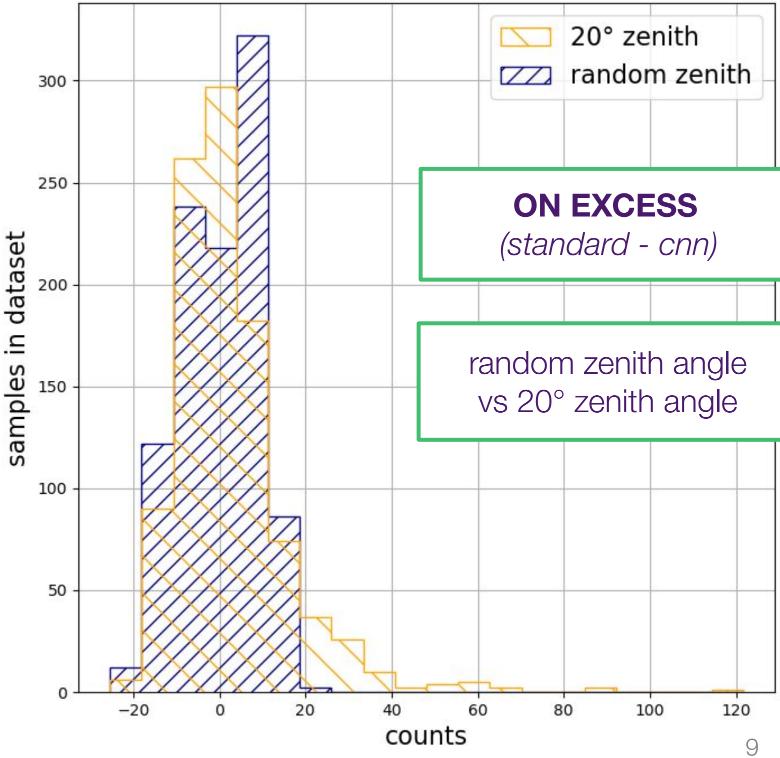

noisy map 16239
NOISY INPUT
Declination [deg]
-65°
-66°
-67°
-68°
-69°
145°
140°
135°
Right Ascension [deg]
normalised counts
1.0
0.8
0.6
0.4
0.2
0.0
clean map 16239
CLEAN INPUT
Declination [deg]
Right Ascension [deg]
normalised counts
denoised map 16239
CLEAN OUTPUT
Declination [deg]
Right Ascension [deg]
normalised counts

denoising



BKG RESIDUALS
(standard - cnn)

random zenith angle
vs 20° zenith angle

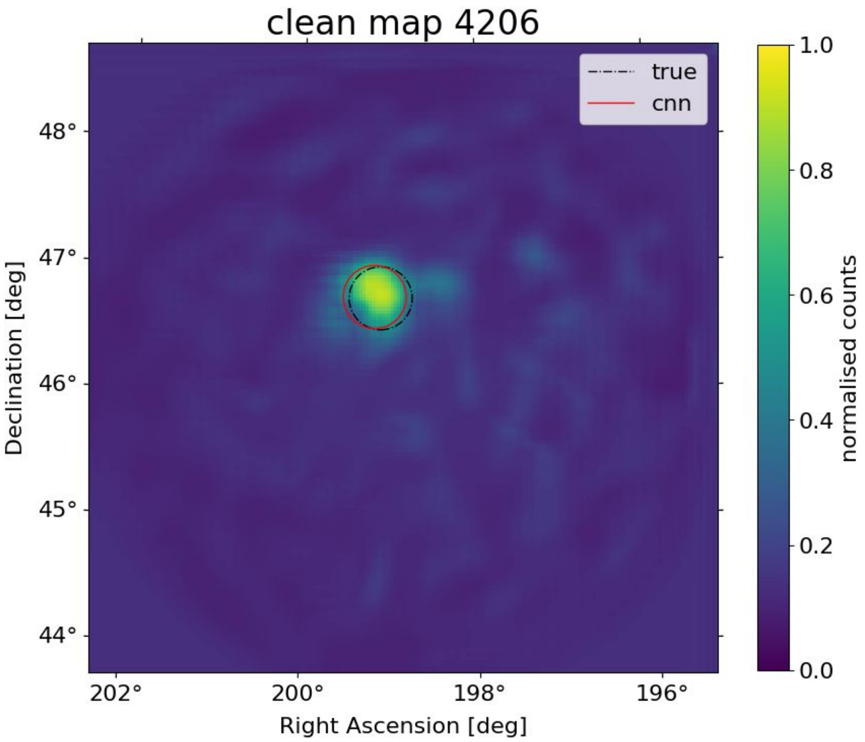


ON EXCESS
(standard - cnn)

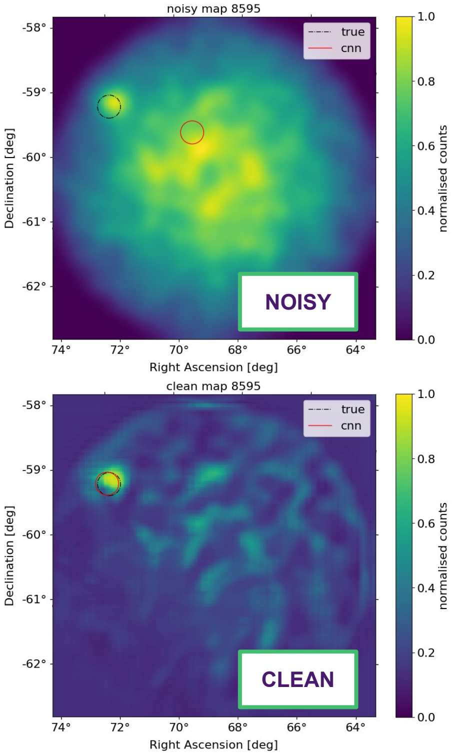
random zenith angle
vs 20° zenith angle

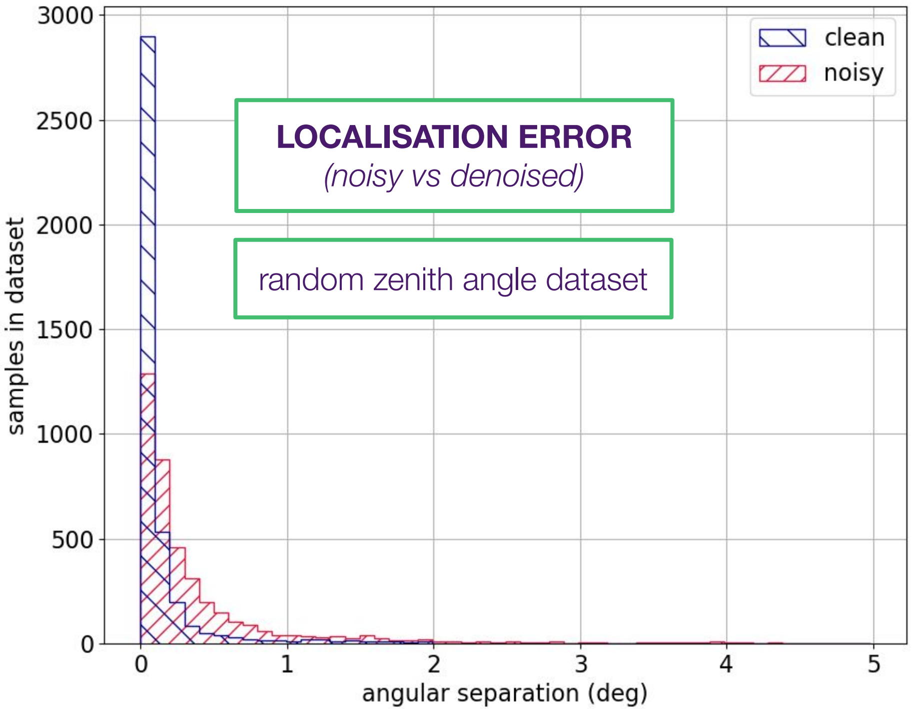
CNN-Regressor

a DL model for candidates localisation in real-time



localisation

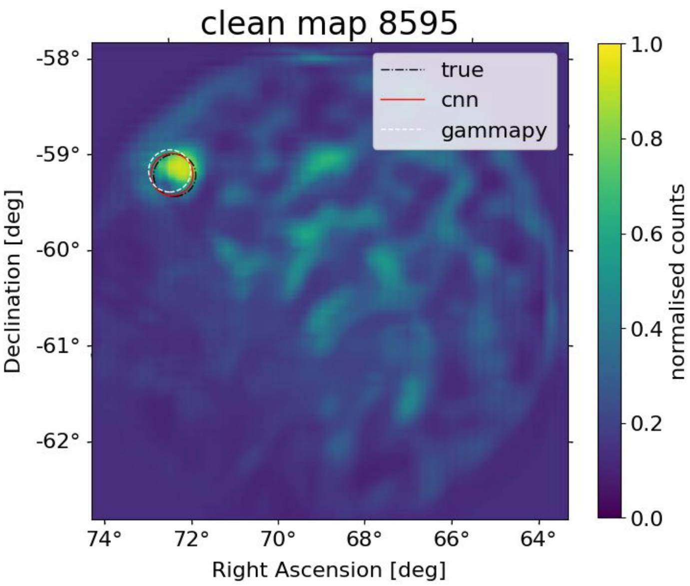


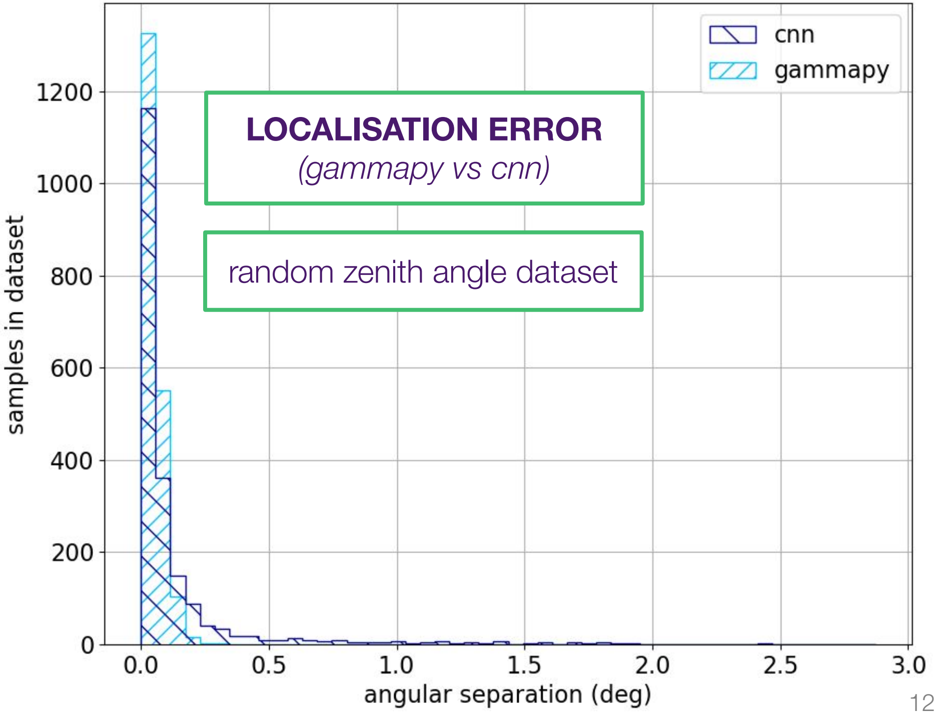


LOCALISATION ERROR
(noisy vs denoised)

random zenith angle dataset

localisation





conclusions

- Application to high level data
- 2 new real-time analysis tools

CNN-cleaner:

- Improved signal-to-noise ratio
- Target independent
- Background/IRF independent
- Improvement on localisation
- Fast (1000 predictions in ~70μs)

CNN-regressor:

- Background/IRF independent
- Fast (1000 predictions in ~65μs)

Known issues:

- Simulations
  → varying exposure, flux and zenith angle
  → refactor with gammapy
- CNN-cleaner
  → uneven source counts loss (zenith 20°)
  → double peaked residuals (random zenith)
- CNN-regressor
  → only one candidate

What comes next…?

- Real data
  → transfer learning
  → incremental learning
- Data quality applications

Ambra Di Piano

Suggestions on …?

- improvements in architecture
- online training with limited samples
- systematic uncertainties / real data
- any/other improvements

commented slides

github page

… contact me at
ambra.dipiano@inaf.it

Bologna - Area CNR

0000-0002-9894-7491

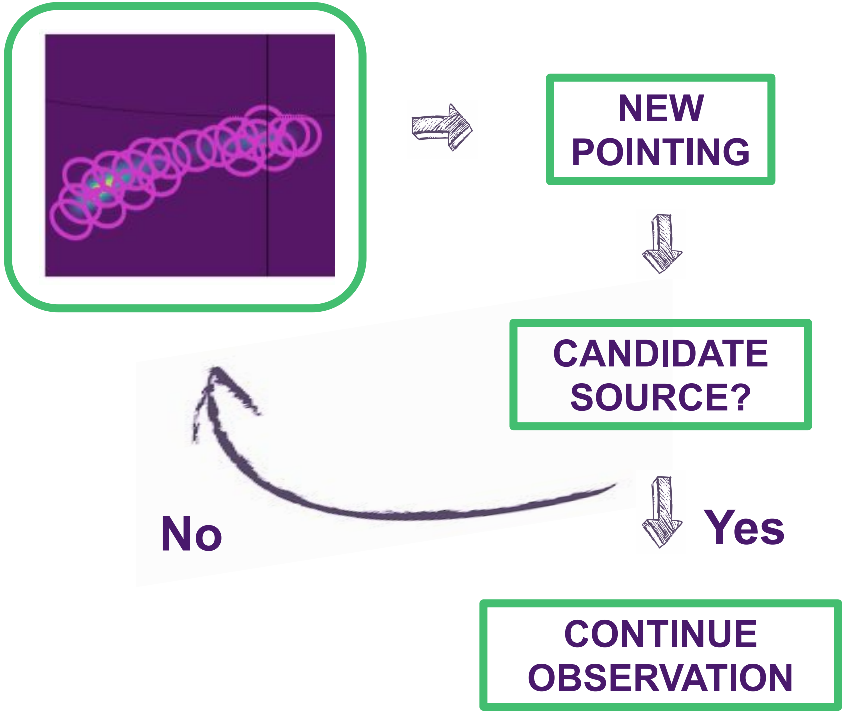
real-time analysis

Context:

- high level Cherenkov data (DL3)
- deep learning enhancements
- real-time analysis
  - [1]ACADA/[2]Science Alert Generation

Possible limitations of Real-Time:

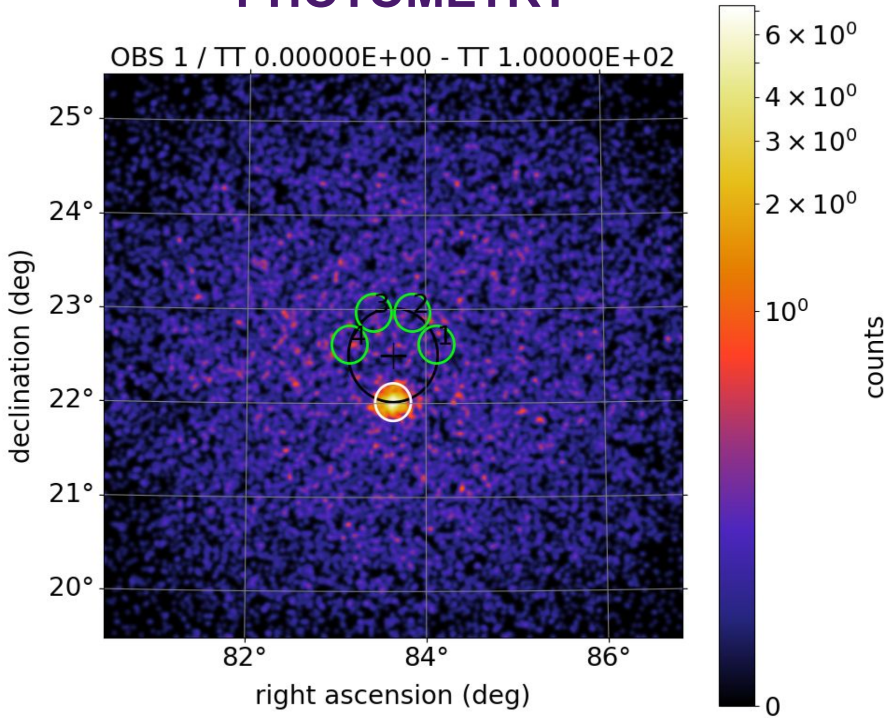
- weather variability
- unknown target
- unknown background / IRFs
- sensitivity degradation
- computation time



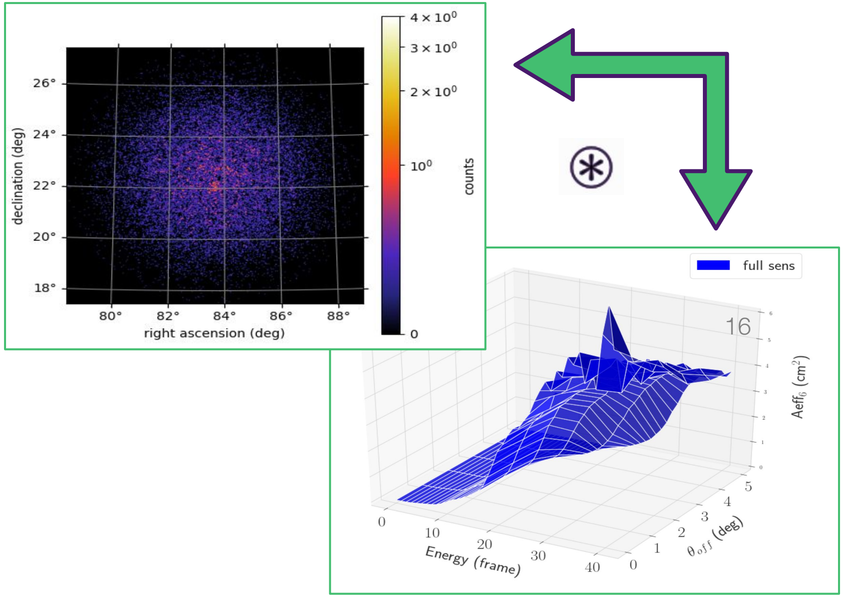
[1]Oya et al. 2019, [2]Bulgarelli et al. 2022

standard techniques

PHOTOMETRY



FULL FIELD MAXIMUM LIKELIHOOD



science tools and frameworks

SIMULATION

[3]Currently in use:

- DL3 dataset
- DL4 dataset

[4]Will replace ctools:

- DL3 dataset
- DL4 dataset

ANALYSIS

[5]Deep learning models:

- background subtraction
- candidate localisation

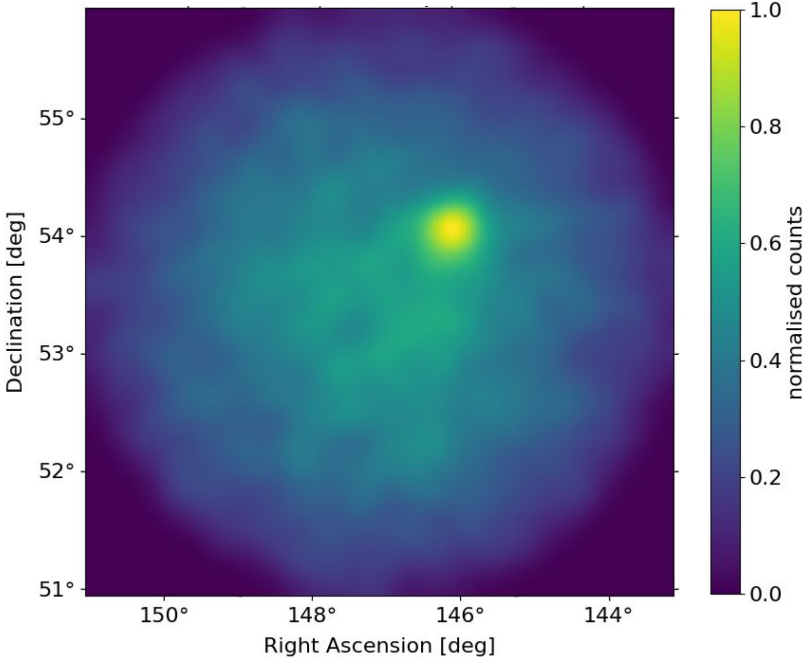
[4]Reference tool:

- candidate localisation
- aperture photometry

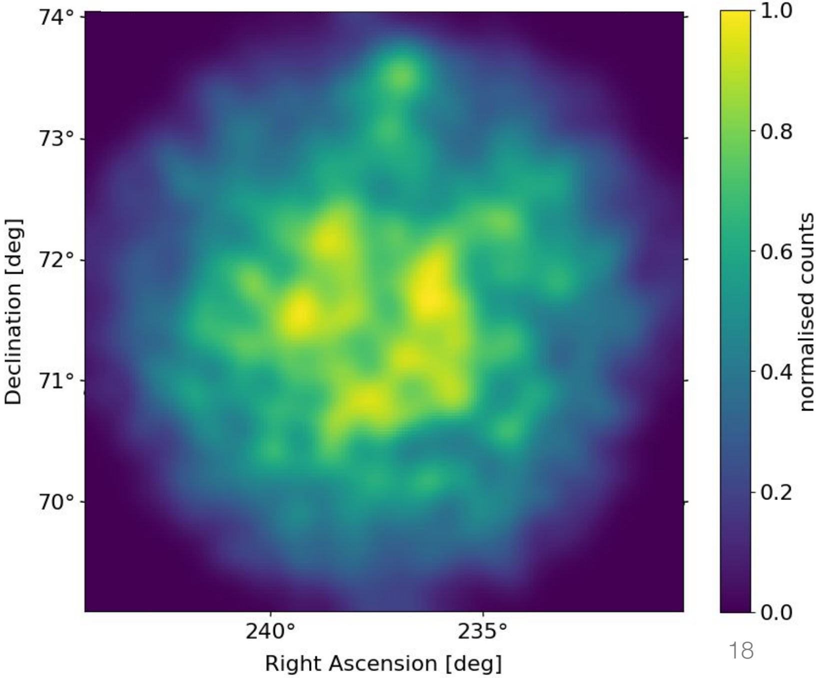
[3]Knoedlseder et al. 2016, [4]Donath et al. 2023, [5]TensorFlow Developers, 2024

simulations

HIGH SNR



LOW SNR



simulations

Fixed parameters:

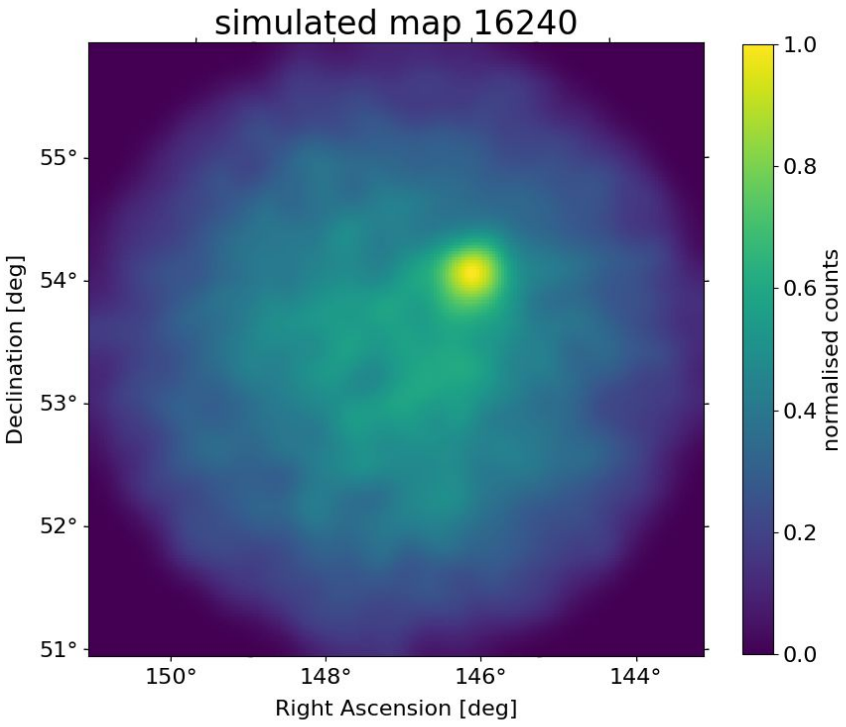
- One source in the FoV
- Flux and spectral model (crab-like)
- Array (4LST) from prod5-v0.1
- Binning (200x200) and pixel (0.025 deg)
- Exposure (100 s) and smoothing (5σ)
- IRF background-subtraction

Random parameters:

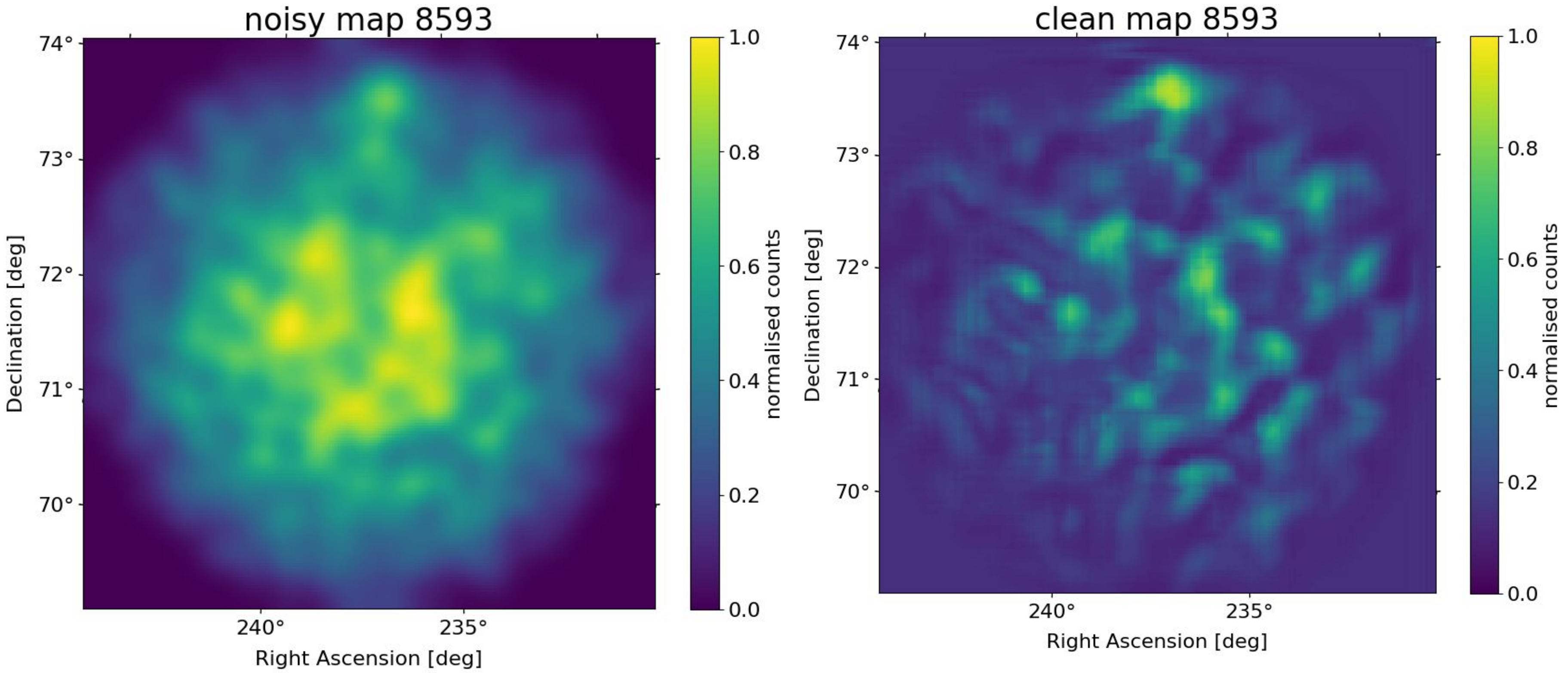
- Background/IRFs (zenith and NSB)
- Pointing and source coordinates
- Source offset (within FoV)

Planned variations (coming next):

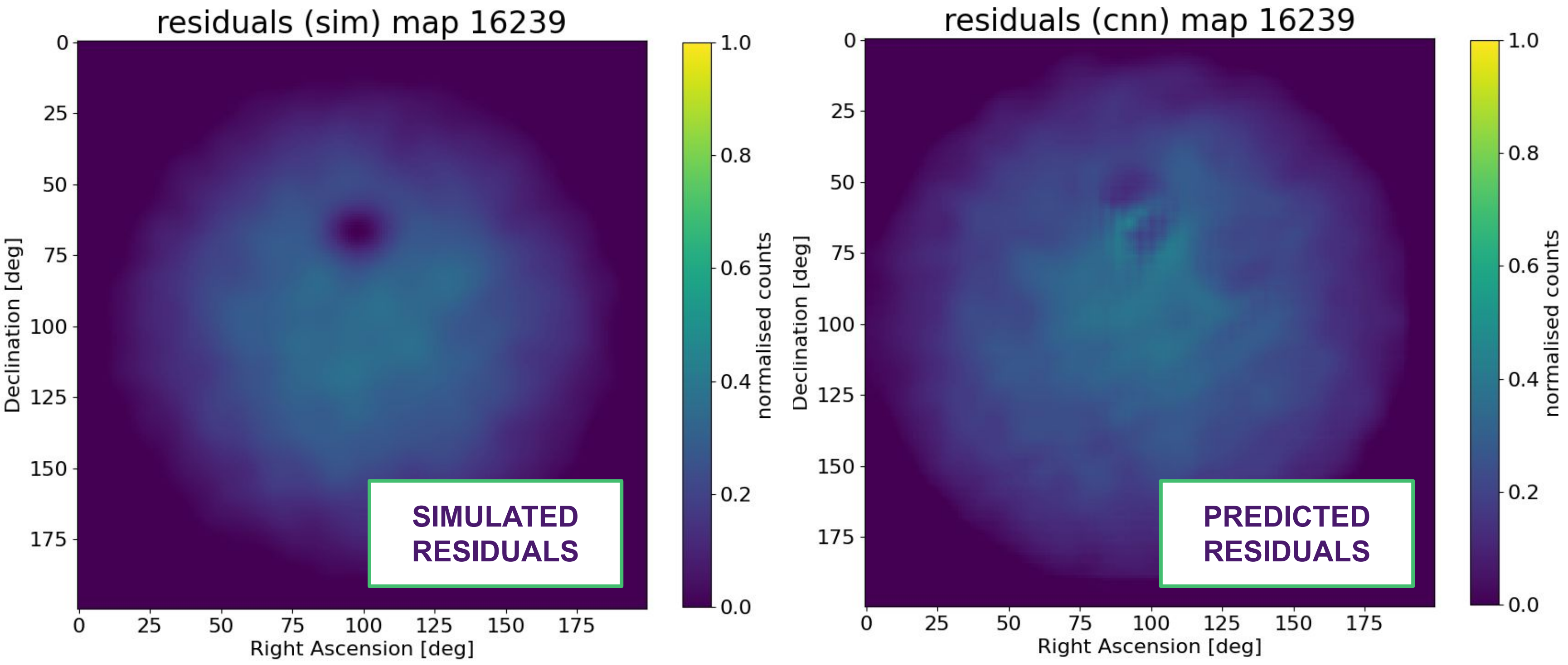
- Vary exposure and flux
- Multiple sources in FoV
- Ring background-subtraction (known targets)



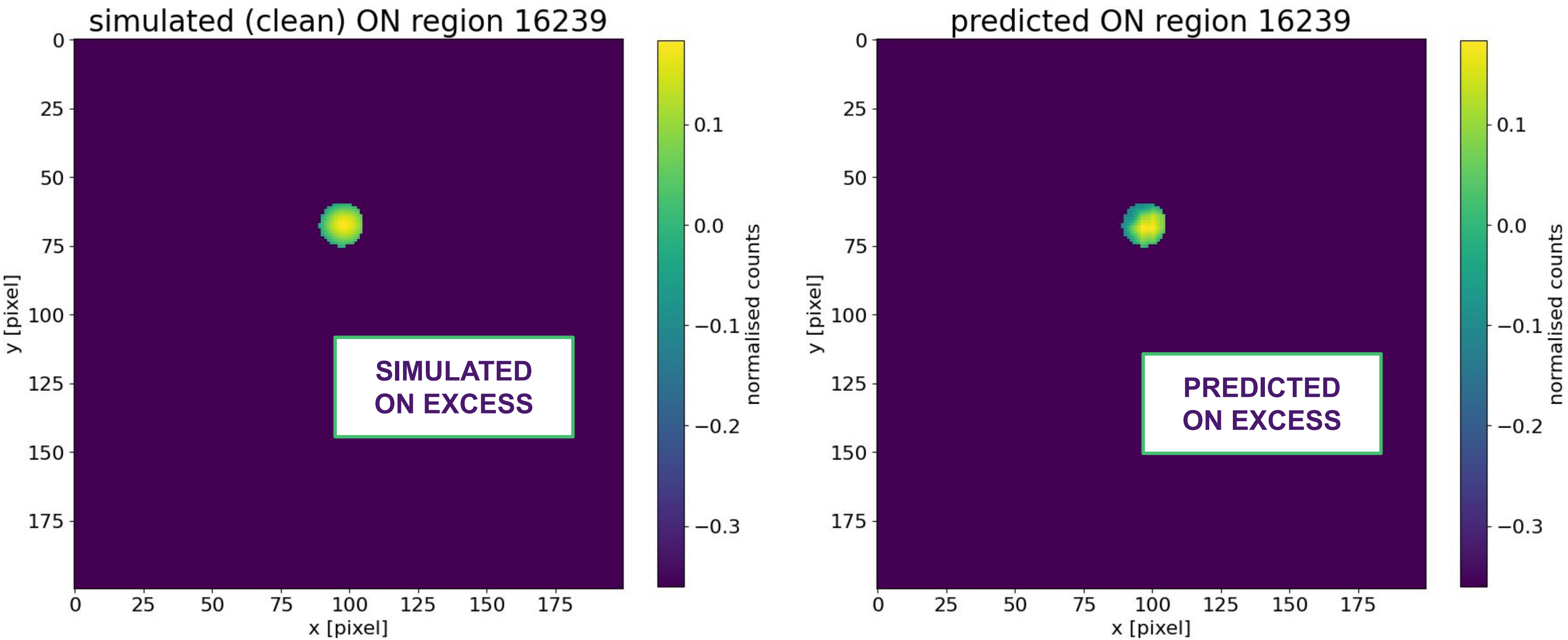
denoising


noisy map 8593
74°
73°
72°
71°
70°
Declination [deg]
240°
235°
Right Ascension [deg]
1.0
0.8
0.6
0.4
0.2
0.0
normalised counts
clean map 8593
74°
73°
72°
71°
70°
Declination [deg]
240°
235°
Right Ascension [deg]
1.0
0.8
0.6
0.4
0.2
0.0
normalised counts

denoising


residuals (sim) map 16239
Declination [deg]
Right Ascension [deg]
0
25
50
75
100
125
150
175
normalised counts
0.0
0.2
0.4
0.6
0.8
1.0
SIMULATED RESIDUALS
residuals (cnn) map 16239
Declination [deg]
Right Ascension [deg]
normalised counts
PREDICTED RESIDUALS

denoising


simulated (clean) ON region 16239
predicted ON region 16239
y [pixel]
x [pixel]
normalised counts
SIMULATED ON EXCESS
PREDICTED ON EXCESS
0
25
50
75
100
125
150
175
0.1
0.0
−0.1
−0.2
−0.3

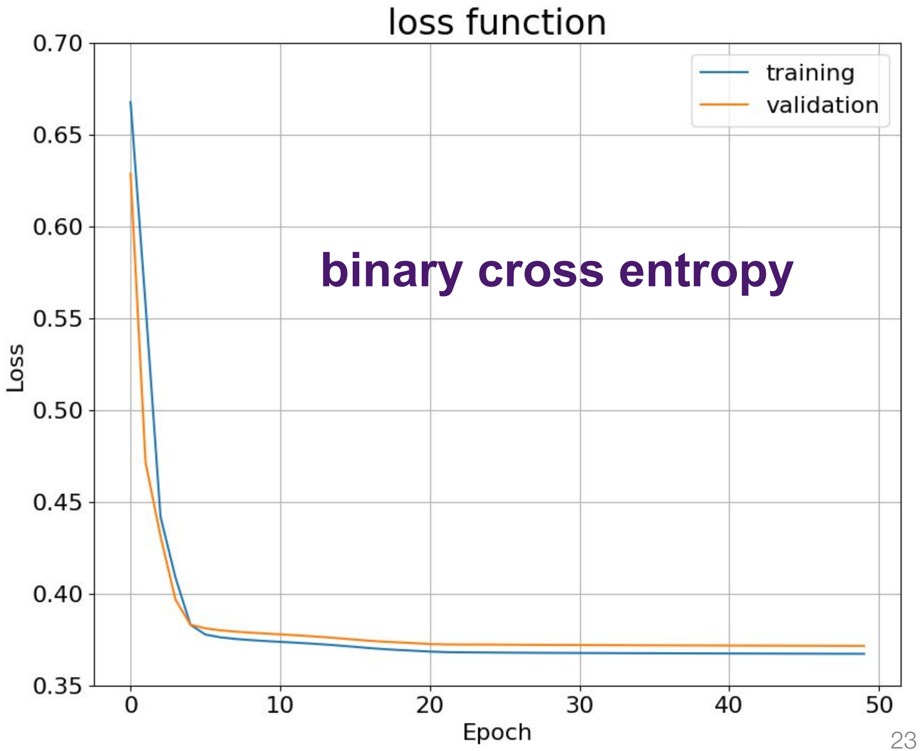
architecture

Input layer:
- Input (200, 200, 1)

Encoder layers:
- Conv2D (12, (2x2), ReLu)
- AvgPooling2D (2x2)
- Conv2D (12, (2x2), ReLu)
- AvgPooling2D (2x2)

Decoder layers:
- Conv2D (12, (2x2), ReLu)
- UpSampling2D (2x2)
- Conv2D (12, (2x2), ReLu)
- UpSampling2D (2x2)

Output layer:
- Conv2D (1, (2x2), Sigmoid)



datasets

Two sets of data:

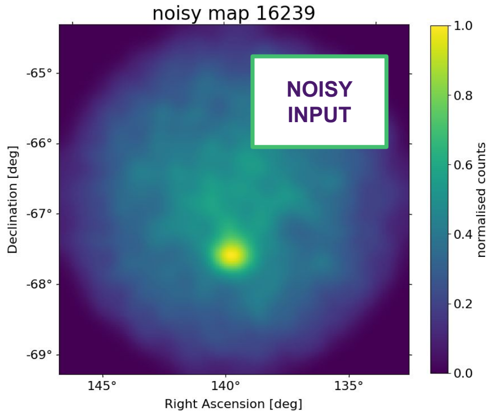
- raw counts maps (noisy)
- background-subtracted counts maps (clean)

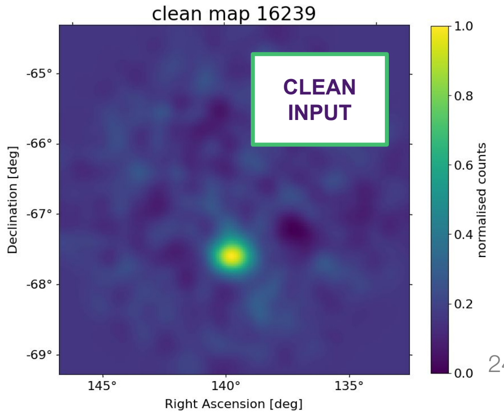
Datasets:

- training dataset → 8k noisy maps and 8k clean maps
- validation dataset → 2k noisy maps + 2k clean maps
- test dataset → 2k noisy maps

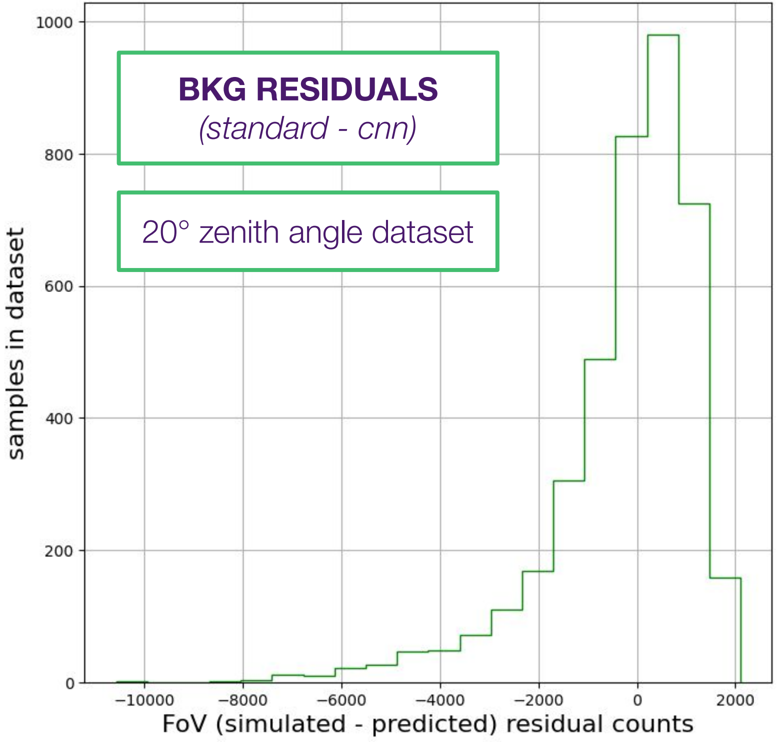
Preprocessing:

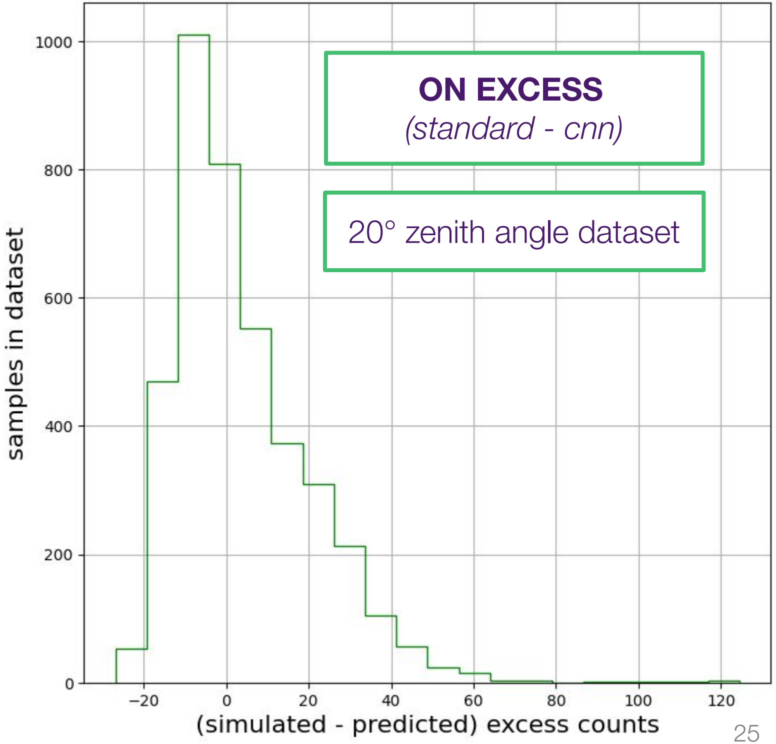
- extract counts maps from photon list
- subtract background for target (clean) datsets
- normalise each counts map with a linear stretch [0, 1]



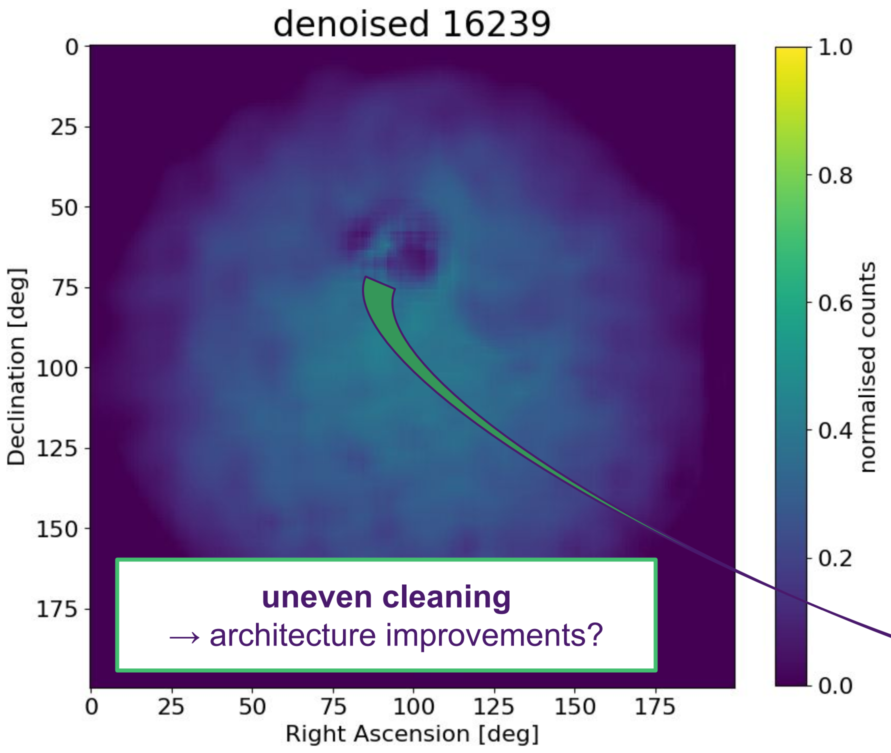


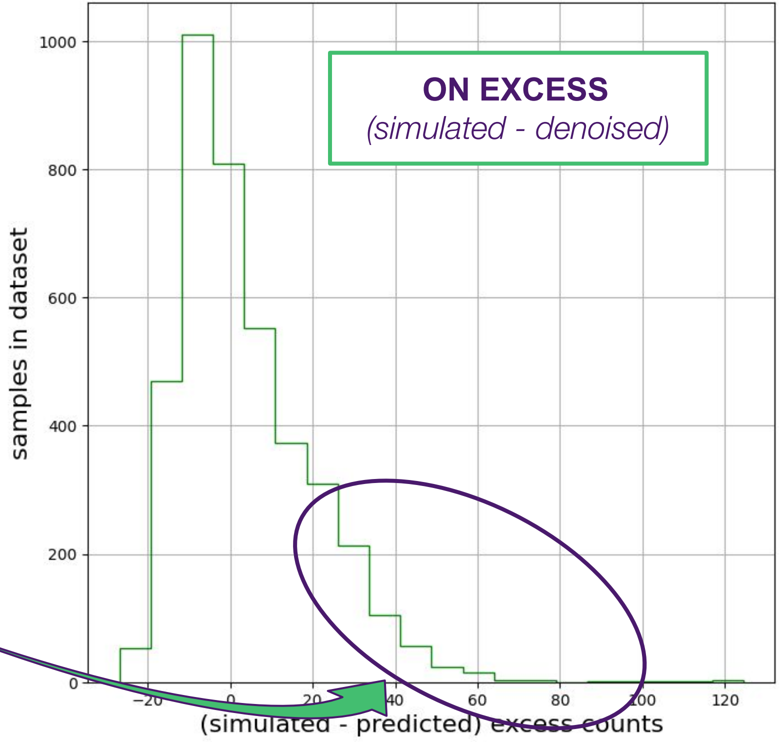
denoising


BKG RESIDUALS
(standard - cnn)
20° zenith angle dataset
samples in dataset
FoV (simulated - predicted) residual counts
1000
800
600
400
200
0
−10000
−8000
−6000
−4000
−2000
0
2000

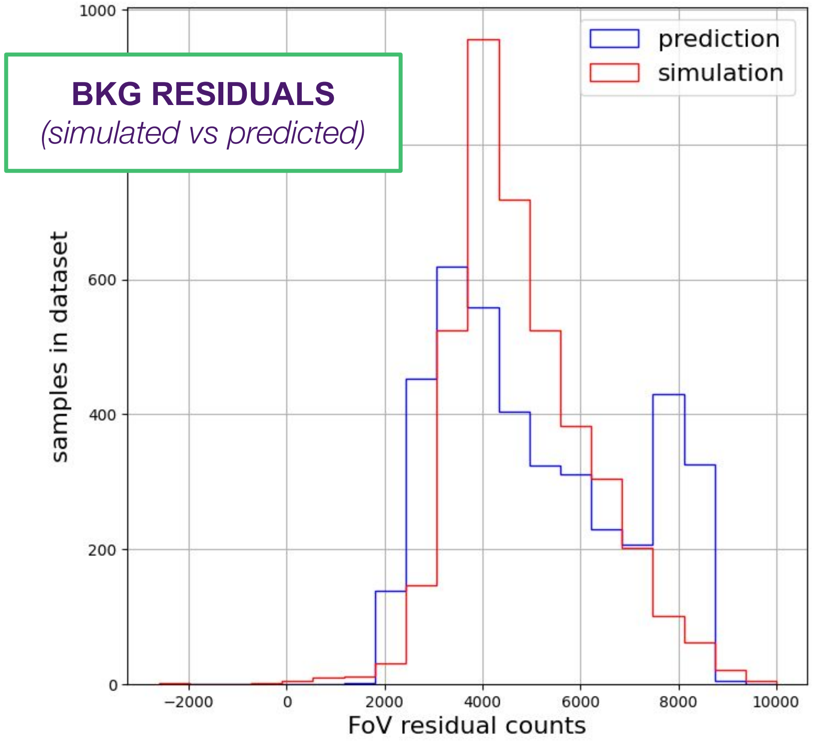

ON EXCESS
(standard - cnn)
20° zenith angle dataset
samples in dataset
(simulated - predicted) excess counts
1000
800
600
400
200
0
−20
0
20
40
60
80
100
120

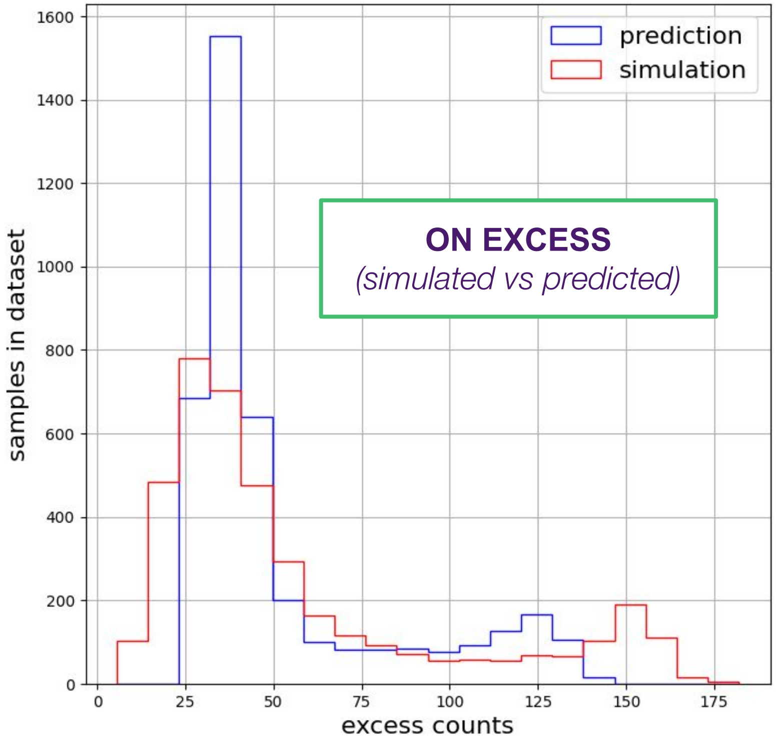
denoising


denoised 16239
Declination [deg]
Right Ascension [deg]
normalised counts
0
25
50
75
100
125
150
175
1.0
0.8
0.6
0.4
0.2
0.0
uneven cleaning
→ architecture improvements?

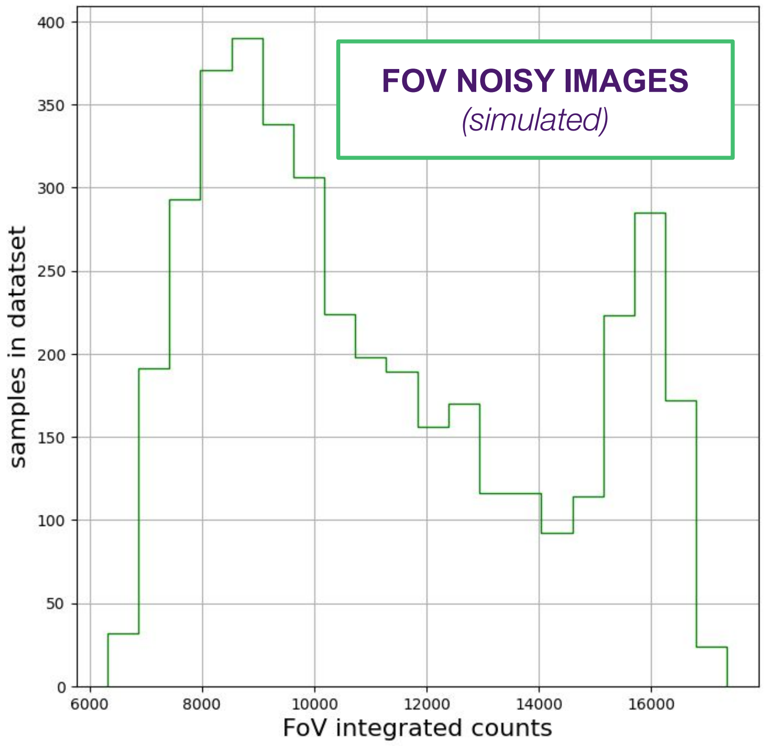

ON EXCESS
(simulated - denoised)
samples in dataset
(simulated - predicted) excess counts
1000
800
600
400
200
0
−20
20
40
60
80
100
120

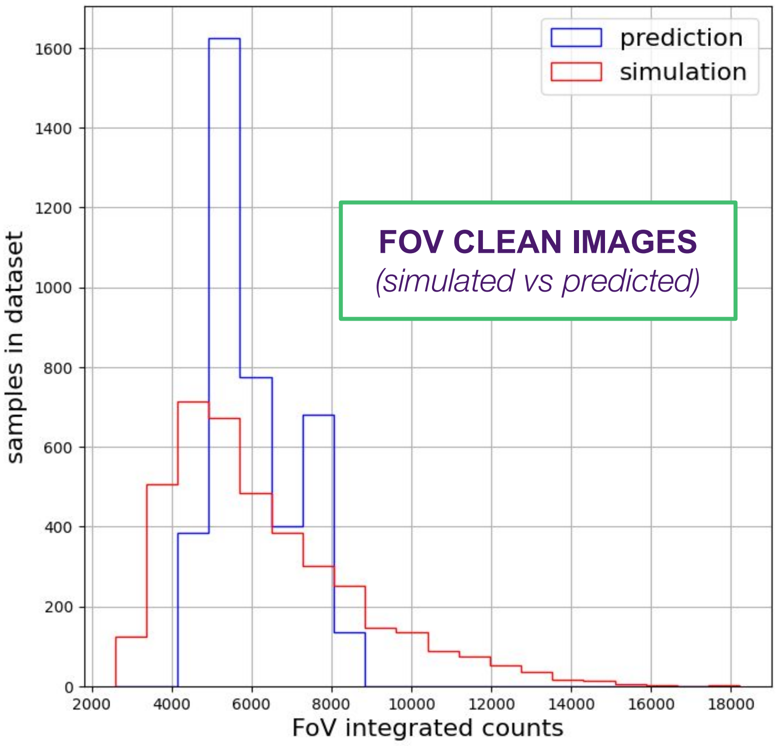
denoising


BKG RESIDUALS
(simulated vs predicted)
prediction
simulation
samples in dataset
FoV residual counts
1000
600
400
200
0
−2000
0
2000
4000
6000
8000
10000


ON EXCESS
(simulated vs predicted)
prediction
simulation
samples in dataset
excess counts
1600
1400
1200
1000
800
600
400
200
0
0
25
50
75
100
125
150
175

denoising


FOV NOISY IMAGES
(simulated)
samples in datatset
FoV integrated counts
0
50
100
150
200
250
300
350
400
6000
8000
10000
12000
14000
16000


prediction
simulation
FOV CLEAN IMAGES
(simulated vs predicted)
samples in dataset
FoV integrated counts
0
200
400
600
800
1000
1200
1400
1600
2000
4000
6000
8000
10000
12000
14000
16000
18000

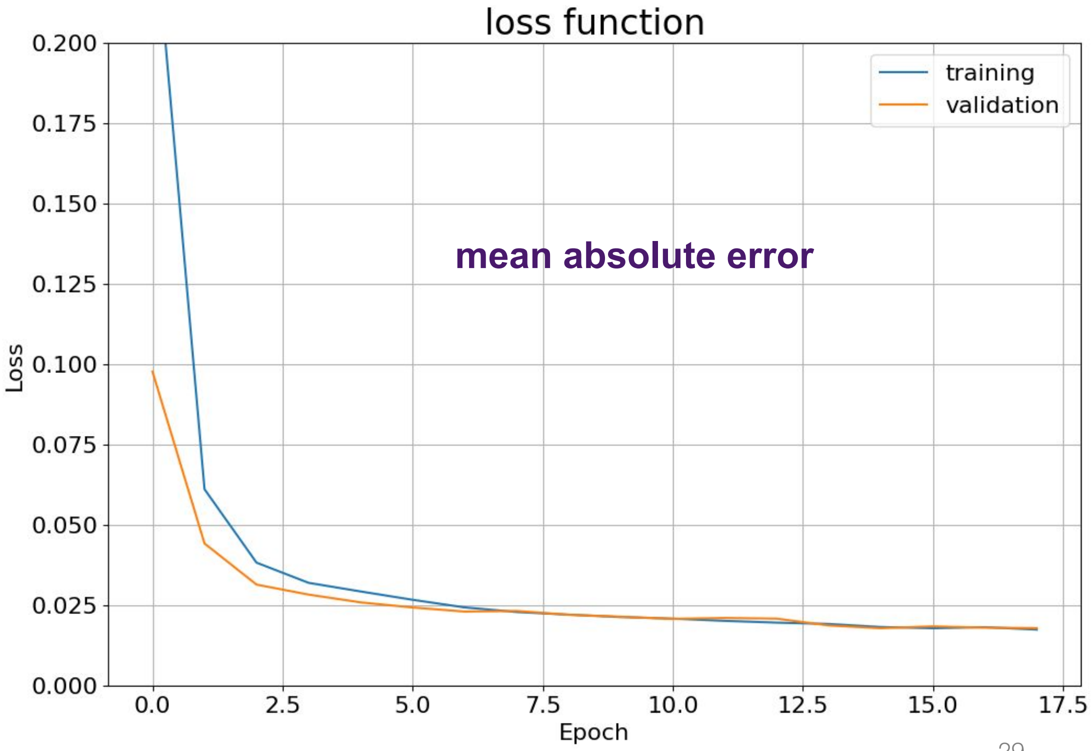
architecture

Input layer:
- Sequential (200, 200, 1)

Layers:

4x{
- Conv2D (6, (4x4), ReLu)
- MaxPool2D (2x2)
- Conv2D (12, (2x2), ReLu)
- MaxPool2D (2x2)
- Dropout (20%)
- Flatten
- Dense (1e4, ReLu)
- Dropout (20%)

Output layer:
- Dense (2, Sigmoid)



datasets

Datasets:

- training dataset → 8k maps
- validation dataset → 2k maps
- test dataset → 2k maps

Preprocessing:

- extract counts maps from photon list
- normalise each counts map with a linear stretch [0, 1]
- denoise counts maps with CNN-cleaner

Preprocessing labels:

- convert WCS to pixel coordinates
- normalise labels by number of bins (x[0, 1], y[0, 1])