

Application of Machine Learning in Very-High-Energy Gamma-ray Astrophysics




टाटा मूलभूत अनुसंधान संस्थान
ML4HEP 2026
Pre-School

Tata Institute for Fundamental Research

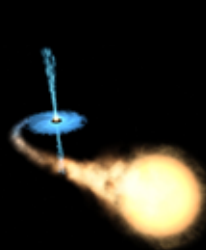
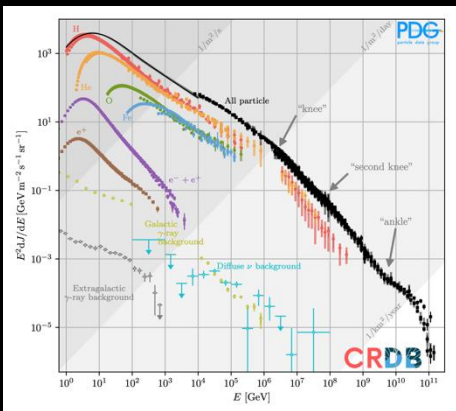
Daniel Nieto

(d.nieto@ucm.es)

Institute for Particle and Cosmos Physics

IPARCOS-UCM





Gamma-ray Binaries



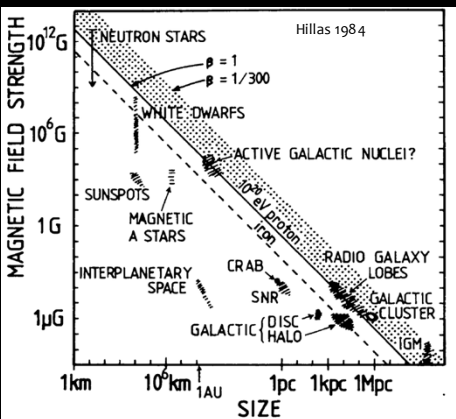
Pulsars



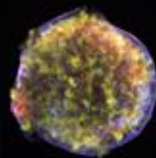
Gamma-ray Bursts



Compact-object mergers



Pulsar Wind Nebulae



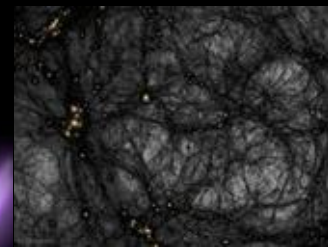
Supernova Remnants



Starburst Galaxies

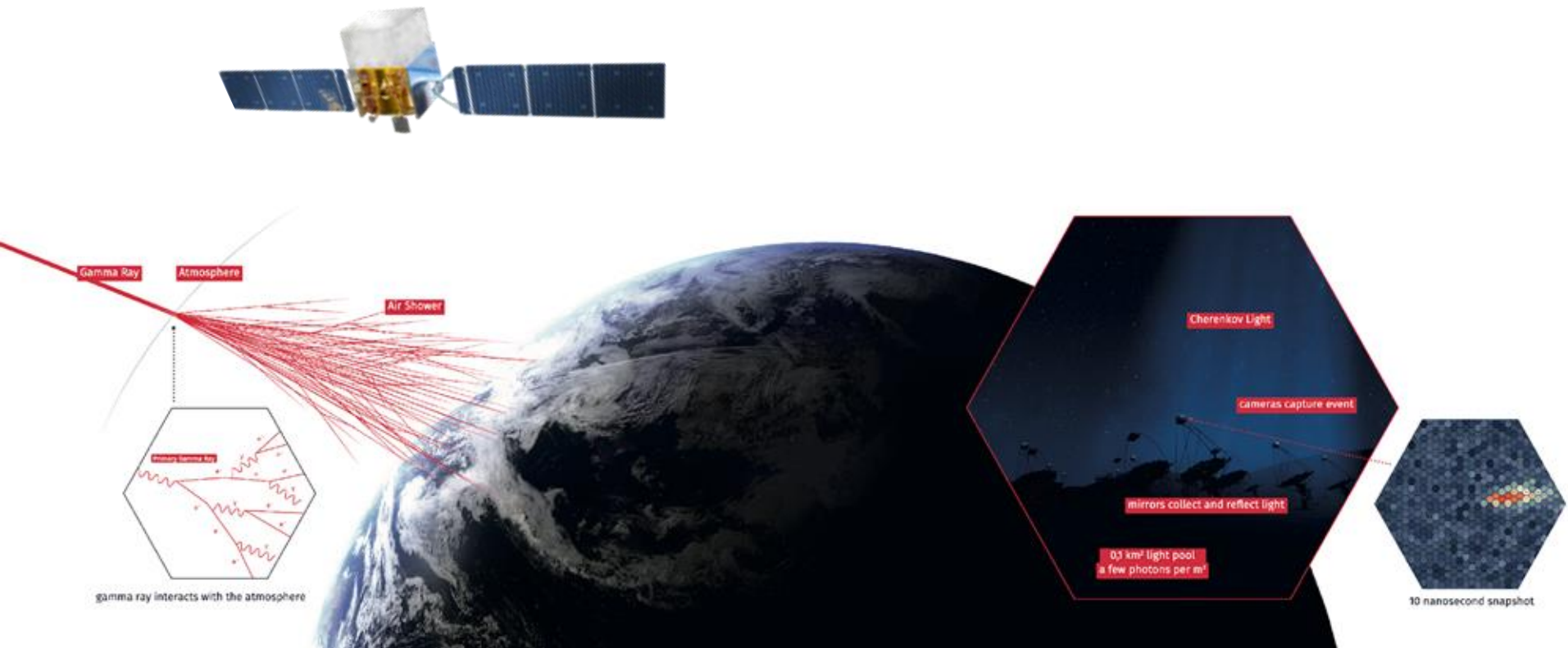


Active Galactic Nuclei



Dark Matter searches
Lorentz Invariance
Cosmology

Fundamental physics

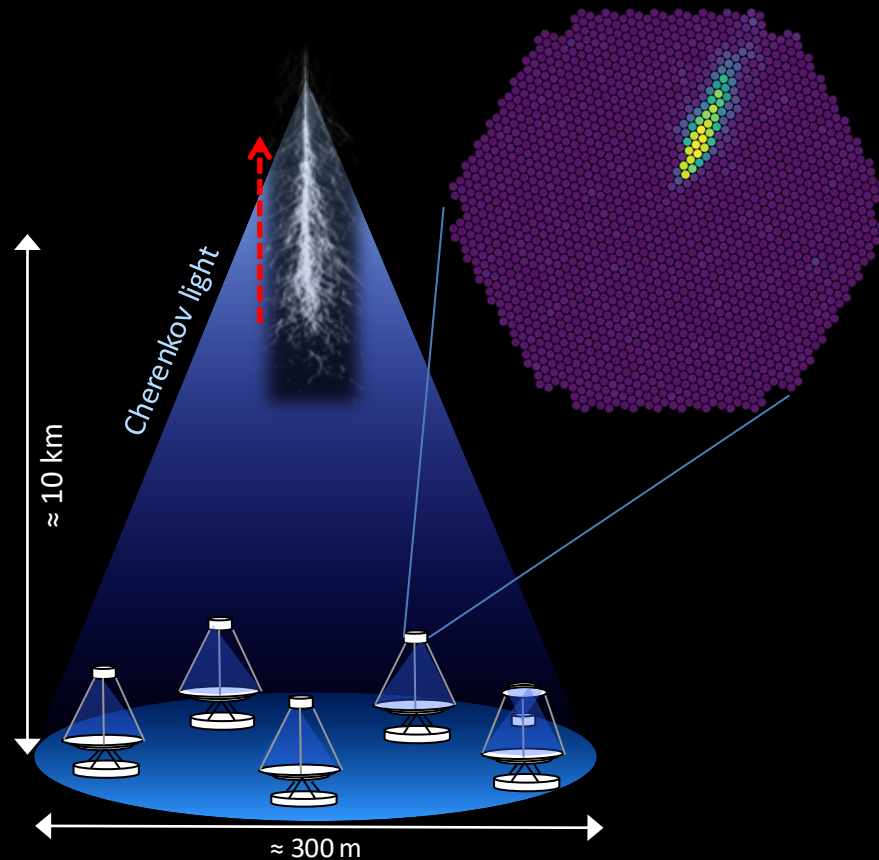


Particle showers produced in Earth's atmosphere by gamma-ray, proton, and carbon-13

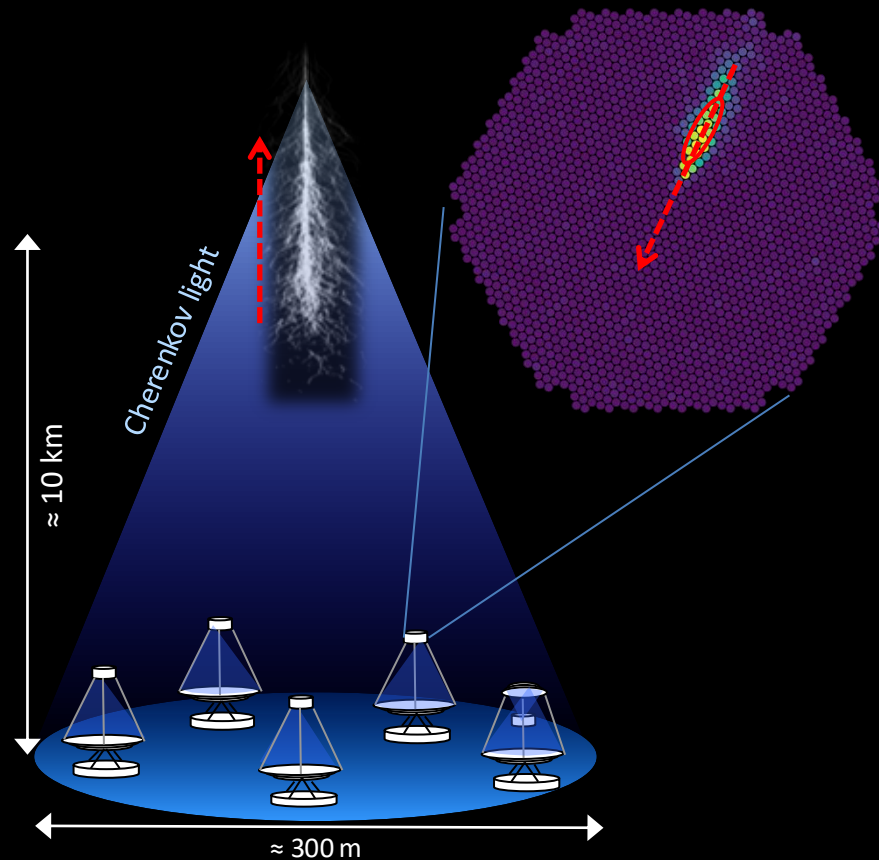
- Initial particle energy: 400 GeV
- Animation time: Shower reaching ground
- **Charged particles: Red dots**
- **Cherenkov light: Blue dots**

Visit <http://veritas.sao.arizona.edu>

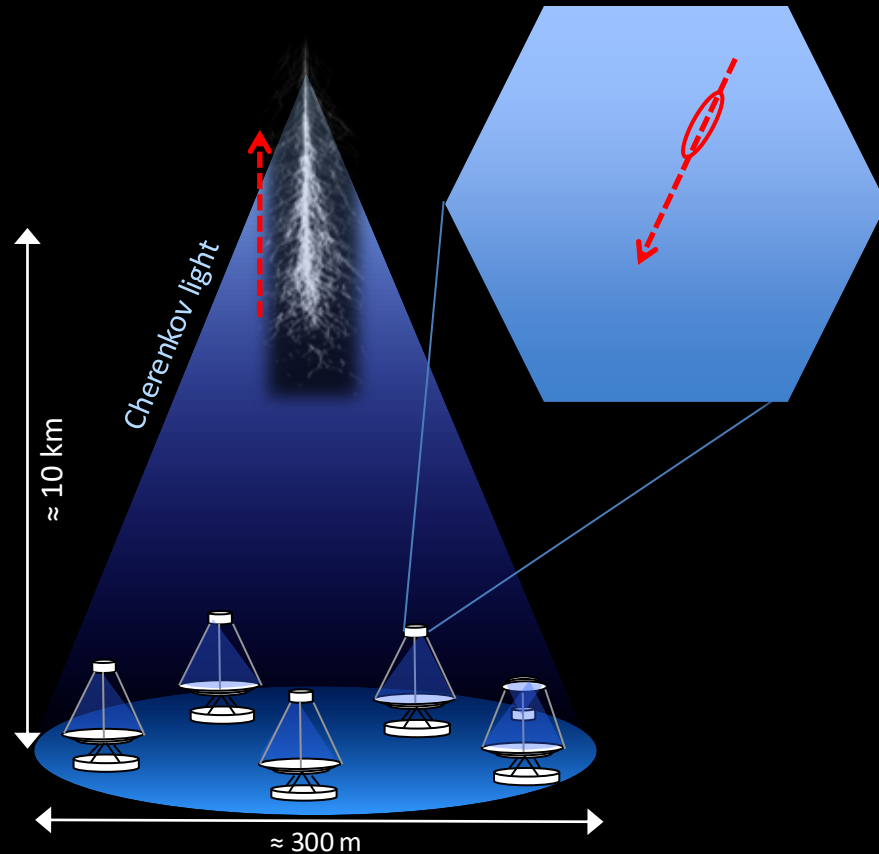
©2012 Martin Schroedter
VERITAS & Harvard Smithsonian Center for Astrophysics



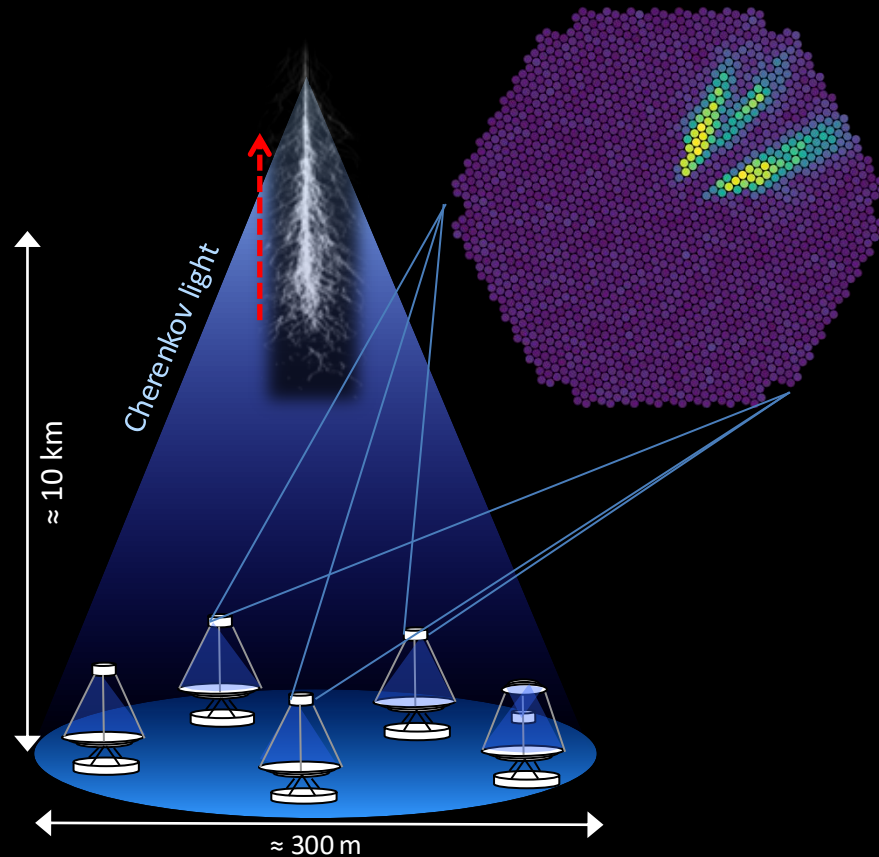
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area ($\sim 10^5 \text{ m}^2$)
- Large background from charged CR
 - Partly irreducible (e^-/e^+ , single-EM, with current methods)
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction



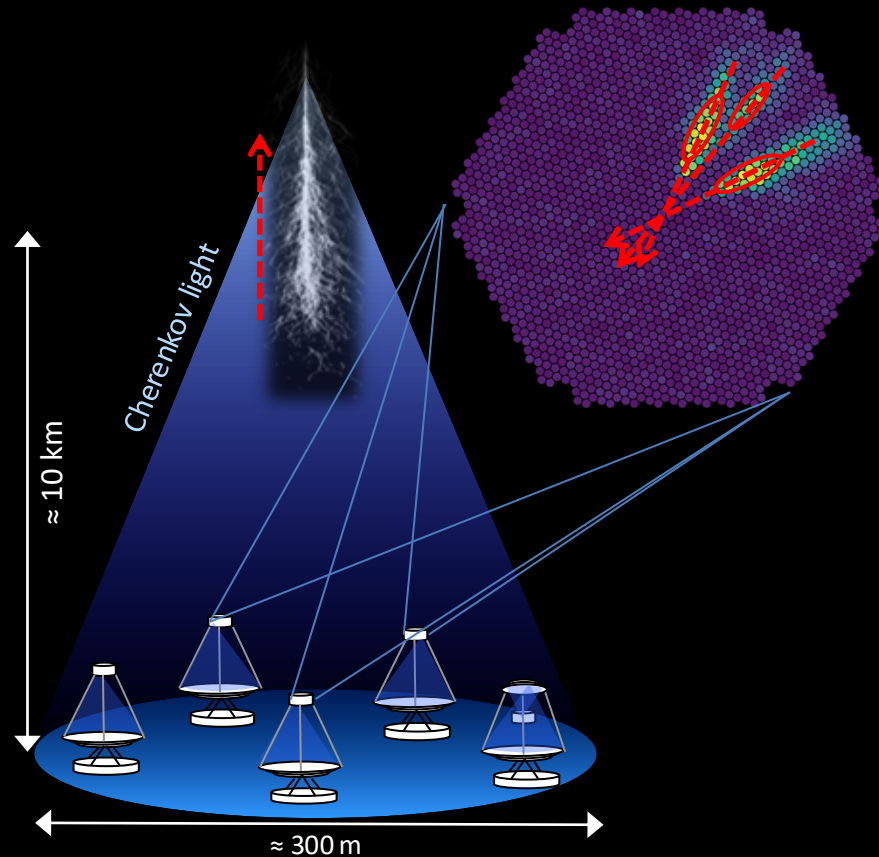
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area ($\sim 10^5 \text{ m}^2$)
- Large background from charged CR
 - Partly irreducible (e^-/e^+ , single-EM, with current methods)
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction



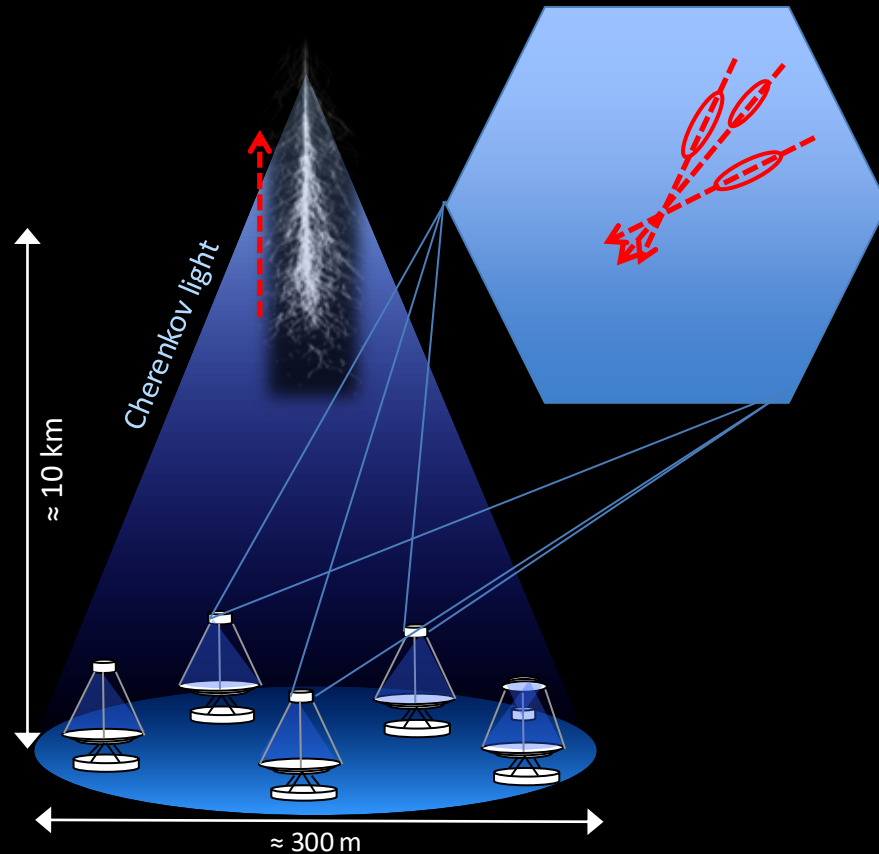
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area ($\sim 10^5 \text{ m}^2$)
- Large background from charged CR
 - Partly irreducible (e^-/e^+ , single-EM, with current methods)
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction



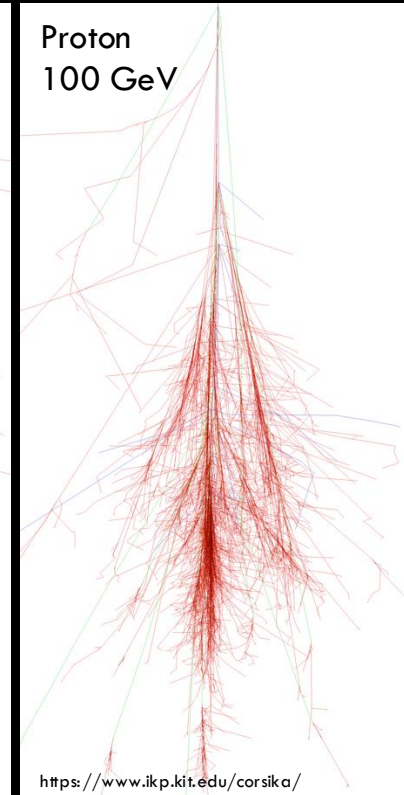
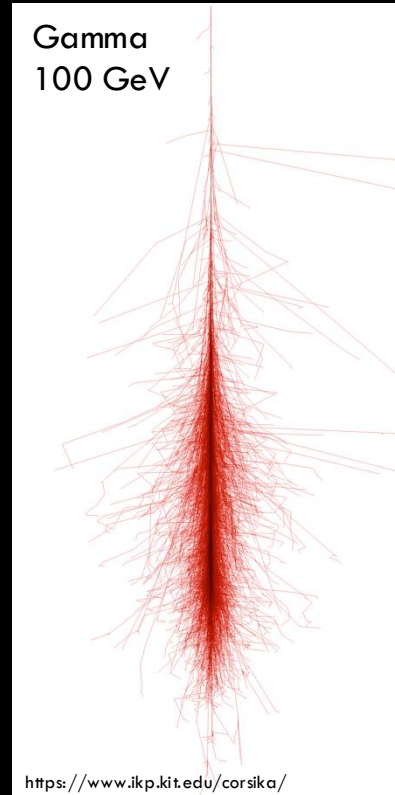
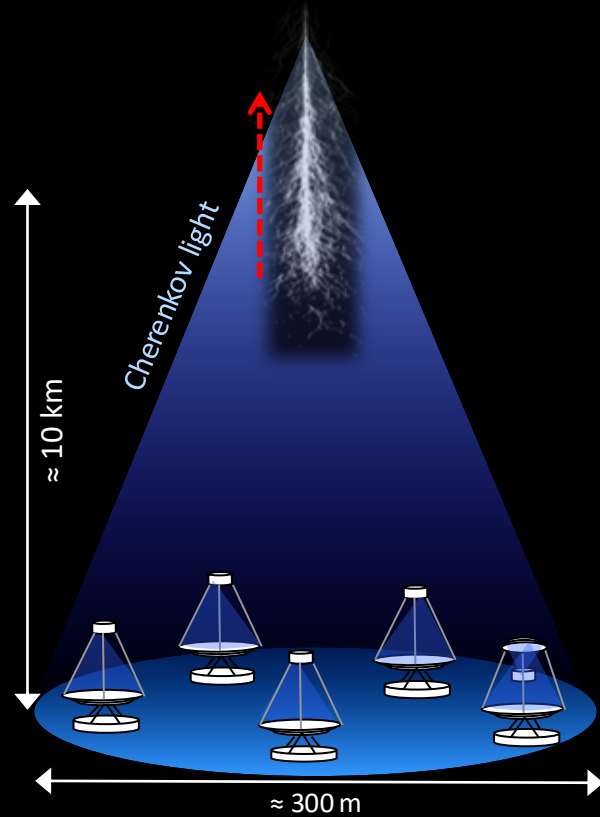
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area ($\sim 10^5 \text{ m}^2$)
- Large background from charged CR
 - Partly irreducible (e^-/e^+ , single-EM, with current methods)
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction

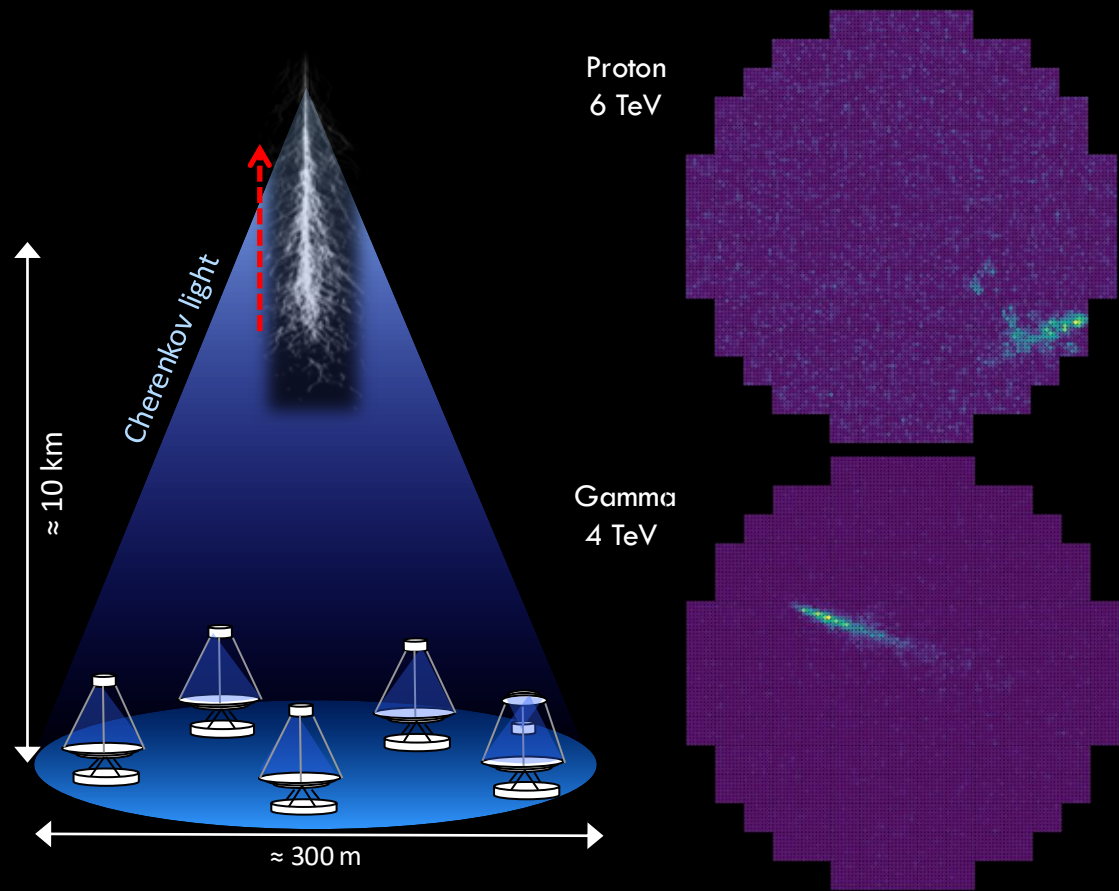


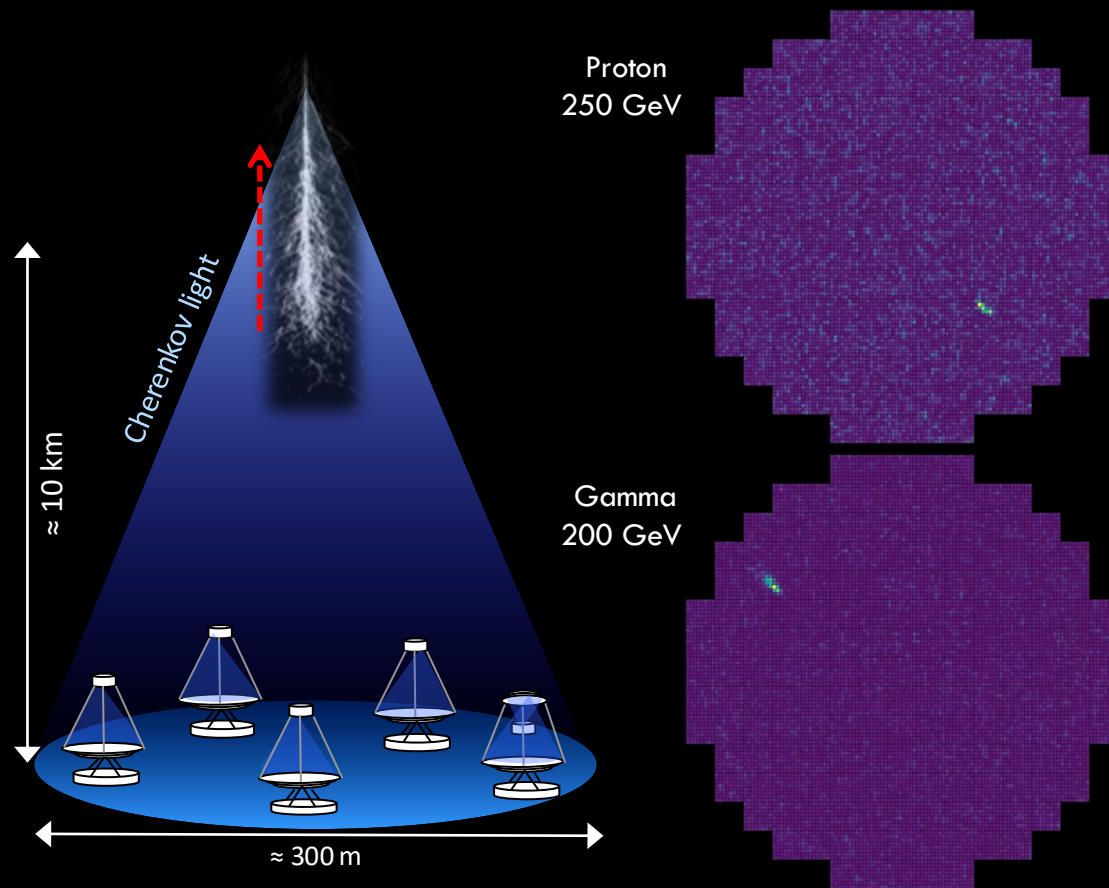
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area ($\sim 10^5 \text{ m}^2$)
- Large background from charged CR
 - Partly irreducible (e^-/e^+ , single-EM, with current methods)
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction

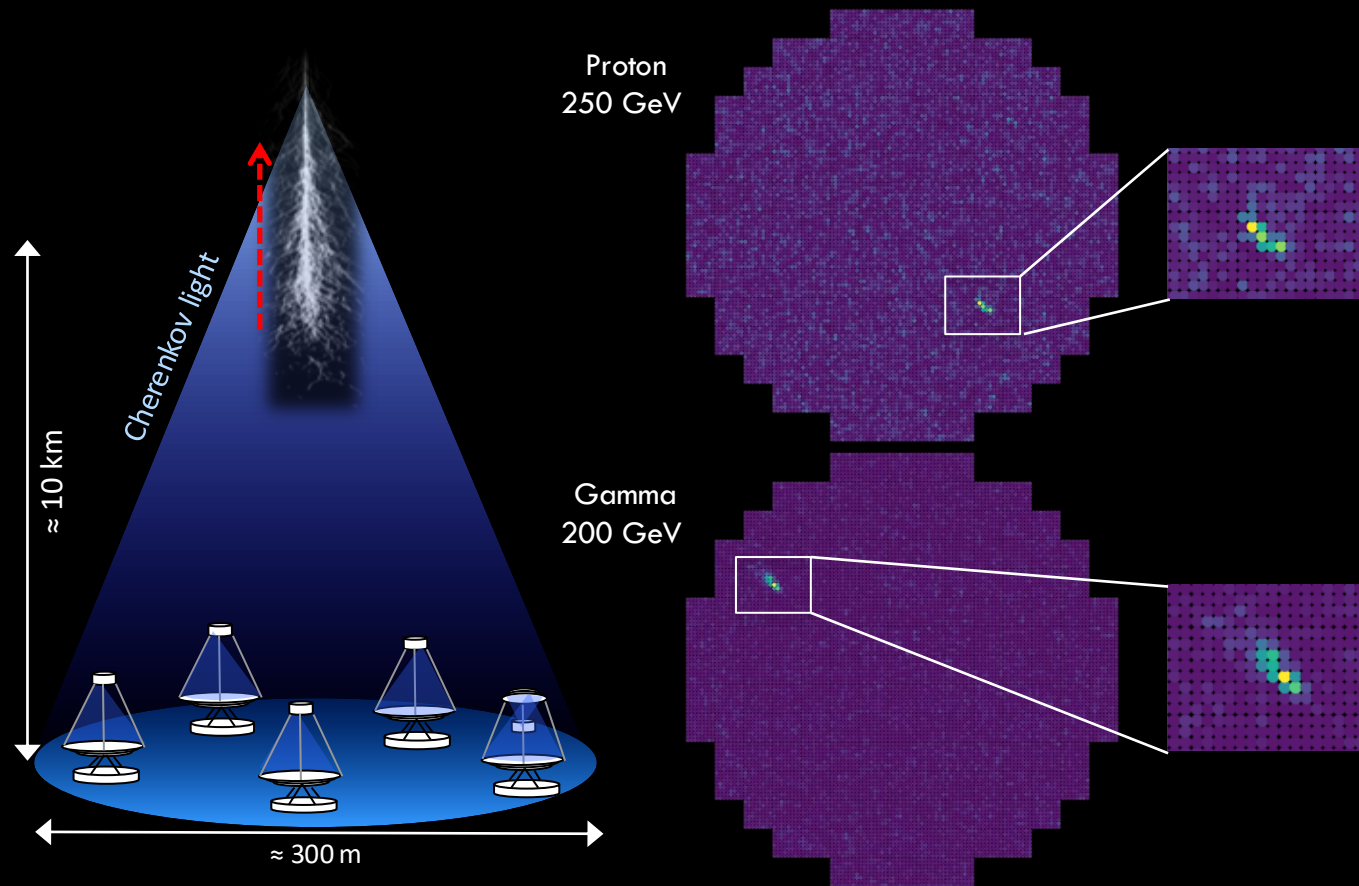


- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ -ray collection area ($\sim 10^5 \text{ m}^2$)
- Large background from charged CR
 - Partly irreducible (e^-/e^+ , single-EM, with current methods)
- Energy window: tens GeV - tens TeV
- Event reconstruction from image:
 - Type of primary event
 - Primary energy estimation
 - Primary arrival direction



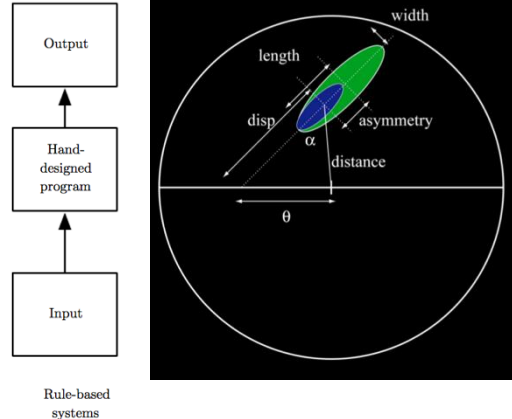






Output: event type,
energy, arrival direction

- Event type: box cuts
- Event energy: parametrization
- Event direction: parametrization

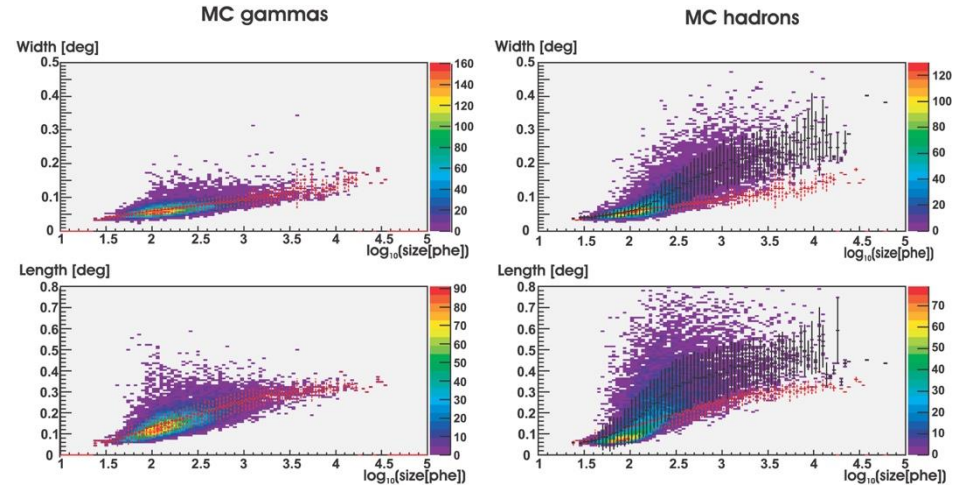


e
v
e
n
t

r
e
c
o
n
s
t
r
u
c
t
i
o
n

Input: observed events

- Based on image parametrization (Hillas parameters)

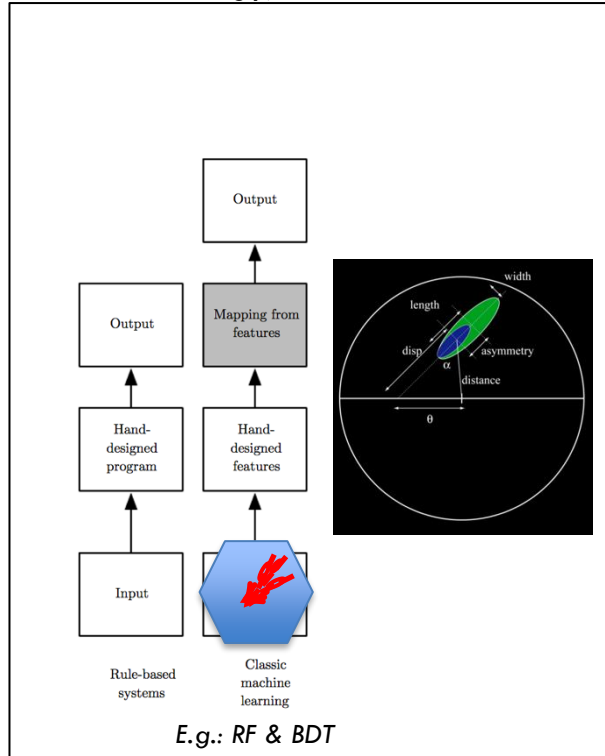


$$E = E(\text{size}, \text{distance}, h_{max})$$

$$DISP = A(SIZE) + B(SIZE) \cdot \frac{WIDTH}{LENGTH + \eta(SIZE) \cdot LEAKAGE2}$$

- Instrument calibration with real data not possible
- Strong dependency on Montecarlo simulations

Output: event type,
energy, arrival direction



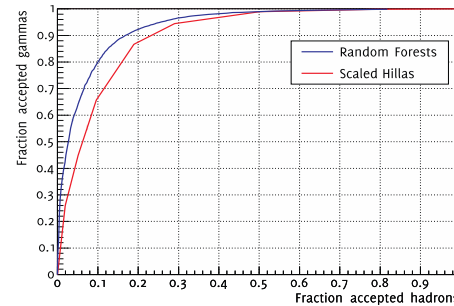
event reconstruction

Input: observed events

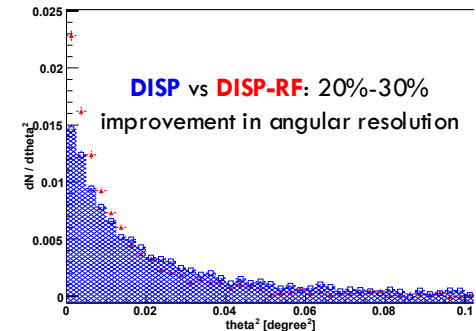
○ Current generation of IACTs: classic ML



- ML method:
 - Random Forest (RF)
- Applied to:
 - Background rejection
 - Arrival direction

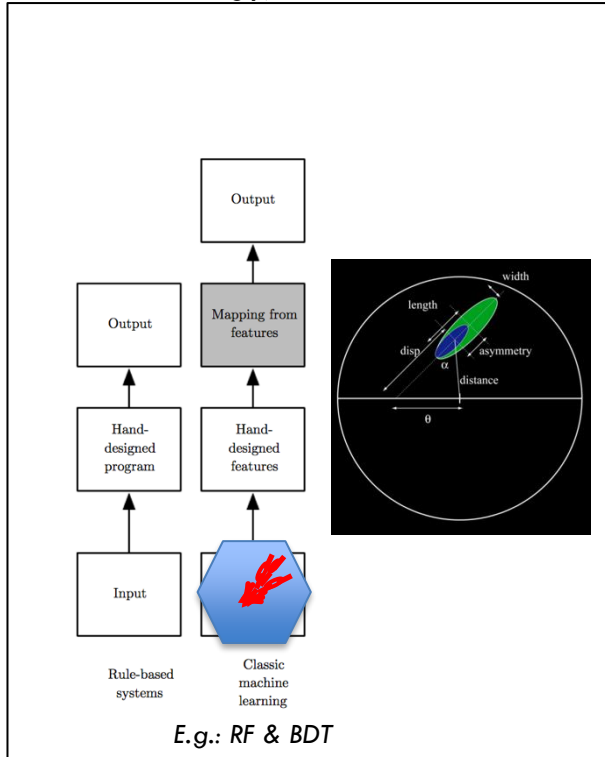


Albert et al., NIM-A 588:424-432 (2008)



Aleksic et al., A&A 524 A77 (2010)

Output: event type,
energy, arrival direction



e
v
e
n
t

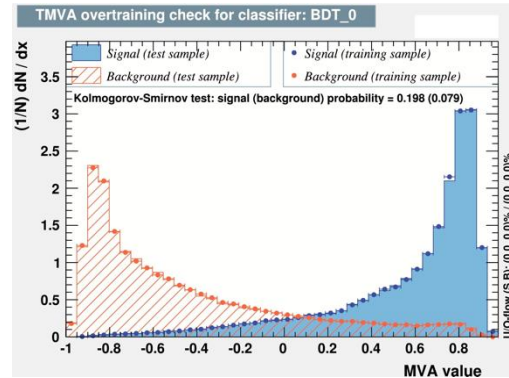
r
e
c
o
n
s
t
r
u
c
t
i
o
n

Input: observed events

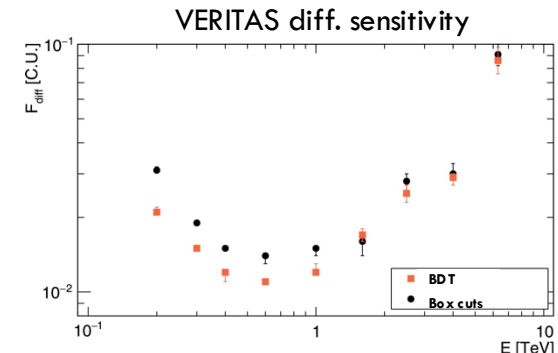
○ Current generation of IACTs: classic ML



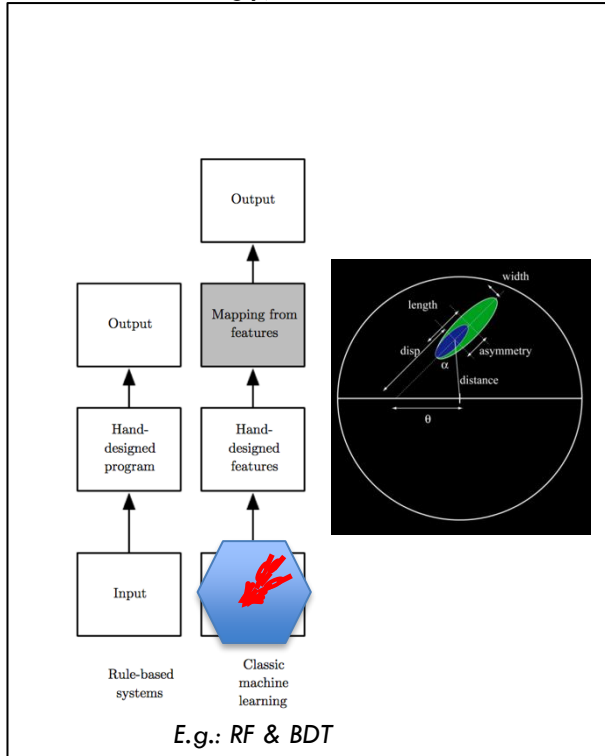
- ML method:
 - Boosted Decision Trees (BDT)
- Applied to:
 - Background rejection



Krause et al., APP V89 P1-9 (2017)



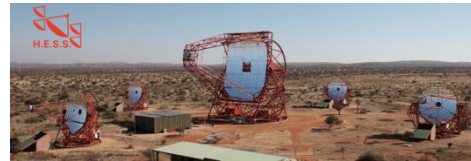
Output: event type,
energy, arrival direction



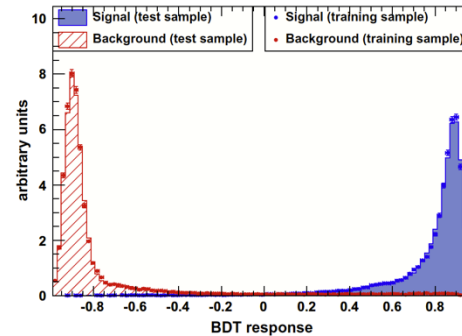
Input: observed events

event reconstruction

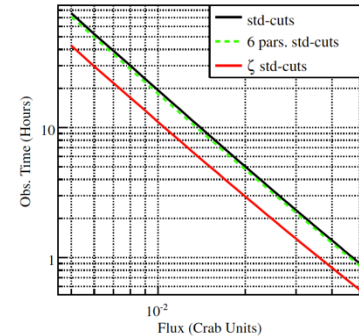
○ Current generation of IACTs: classic ML



- ML method:
 - Boosted Decision Trees (BDT)
- Applied to:
 - Background rejection



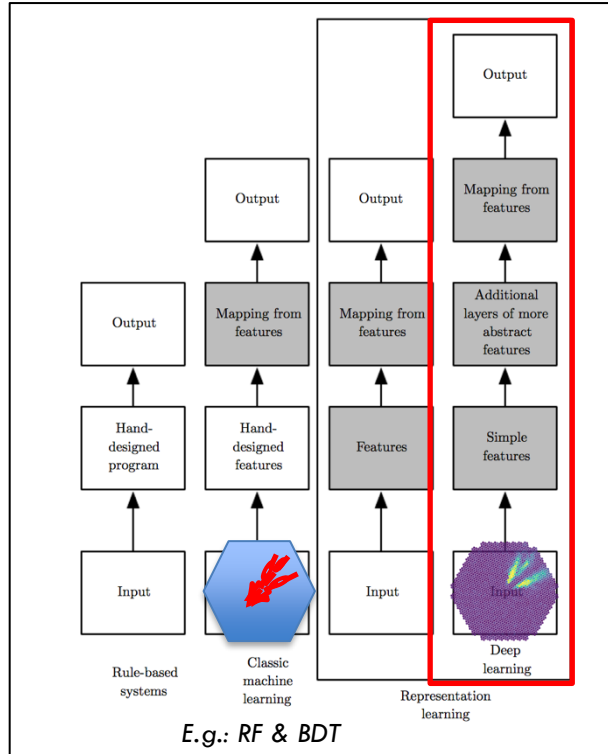
Becherini et al., APP V34-12 P858-870 (2011)



Ohm et al., APP V31-5 P383-391 (2009)

(Results for H.E.S.S. I only)

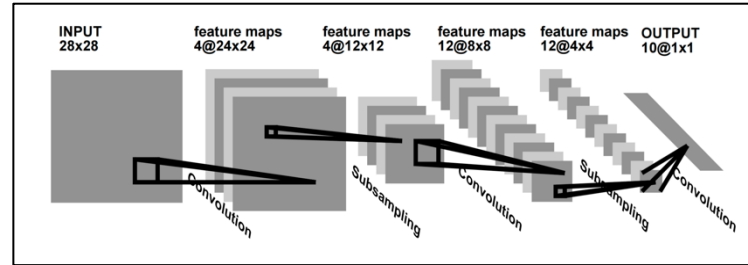
Output: event type,
energy, incoming direction



event reconstruction

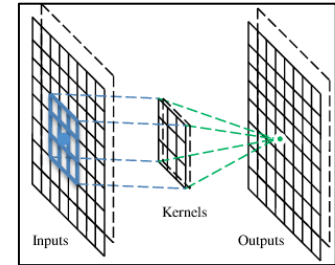
Input: observed events

Convolutional Neural Network (CNN)



LeCun *et al.*

Convolution



Guo *et al.*

- DL capable of **extracting** and mapping image features automatically with unprecedented classification accuracy. Hyper-active CS research field constantly improving
- Many HEP/Astro experiments already exploring/utilizing the technique (LIGO, LHC, MicroBooNe, NOVA, etc...)

Method:

- Use deep learning to reconstruct IACT events from non-parameterized images
 - Performance enhancement -> better sensitivity

But there are risk...

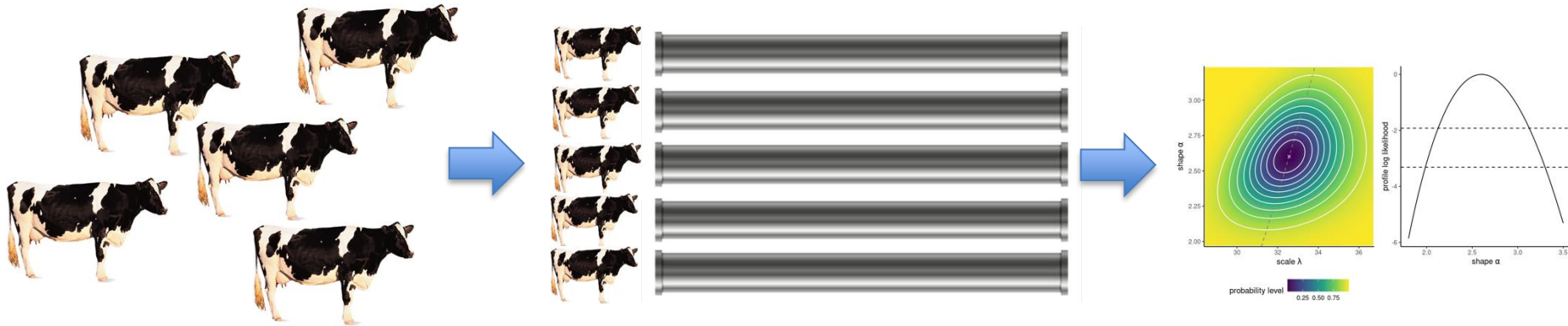
- MC reliability (e.g. network selecting some features from your MC not present in real data)

Issues with traditional MC + likelihood maximization in the era of big data in astroparticle physics



- **Information Loss:** Traditional parametric methods often fail to exploit all the valuable information embedded in raw data, relying instead on restrictive summary features
- **Scalability Bottleneck:** Traditional spatial and temporal likelihood maximization using Monte Carlo simulations struggles with current astronomical data volumes and detector topologies

Machine learning as a potential solution

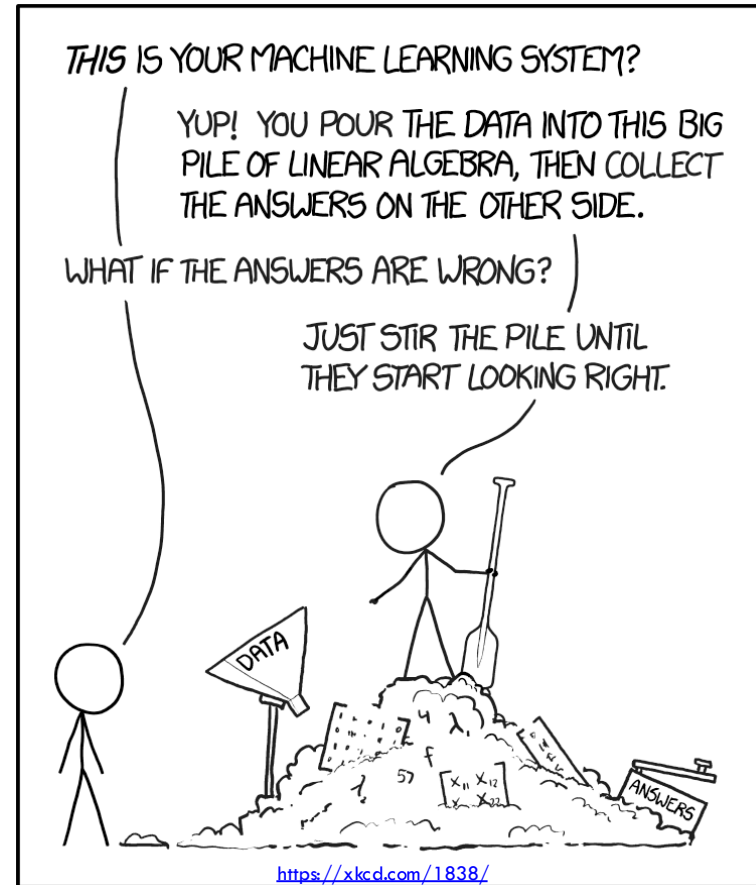


- **Neural networks** directly process raw information, leveraging GPU and FPGA parallelization to efficiently optimize analysis in high-dimensional spaces

Black box?

The Universal Approximation Theorem ([Hornik \(1991\)](#))

- UAT proves neural networks can approximate any continuous function
- Network designs possess the capacity to model highly complex, non-linear relationships to any desired accuracy
- Models act as mathematically sound interpolators mapping inputs to outputs, rather than simply memorizing data (if properly trained...)



- 5-20 fold better sensitivity w.r.t. current IACTs
- 4 decades of energy coverage: 20 GeV to 300 TeV
- Improved angular and energy resolution
- Two arrays (North: La Palma / South: Cerro Paranal)

CTAO

Low-energy range:

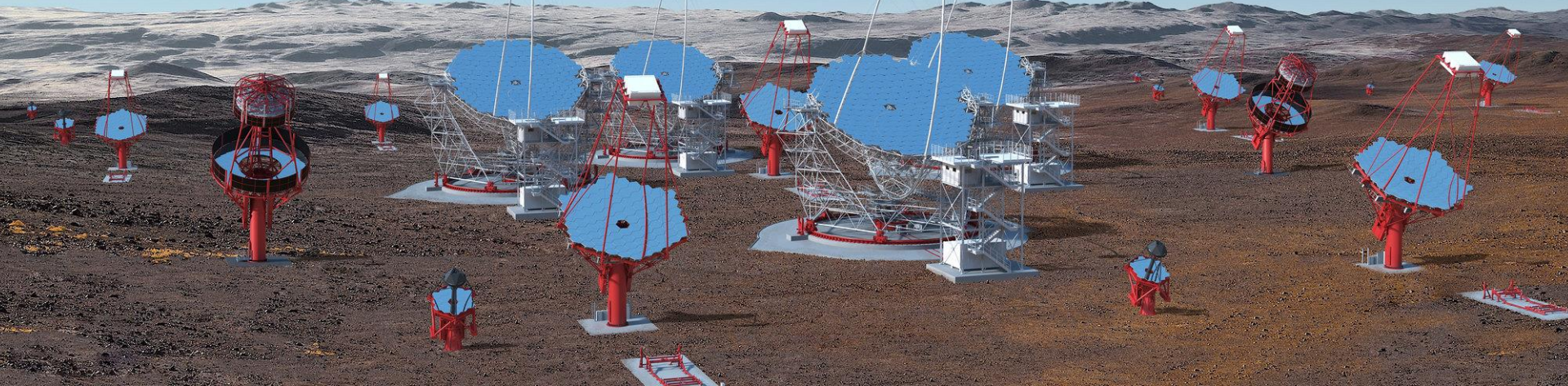
23 m \varnothing
Parabolic reflector
4.3° FoV
Energy threshold 20 GeV

Mid energy-range:

12 m \varnothing modified Davies-Cotton reflector
7.5° FoV
**Full system sensitivity in the
150 GeV – 5 TeV range**

High-energy range:

4 m \varnothing Schwarzschild-Couder reflector
10° FoV
**Several km² area at
multi-TeV energies**



- Stereoscopy:

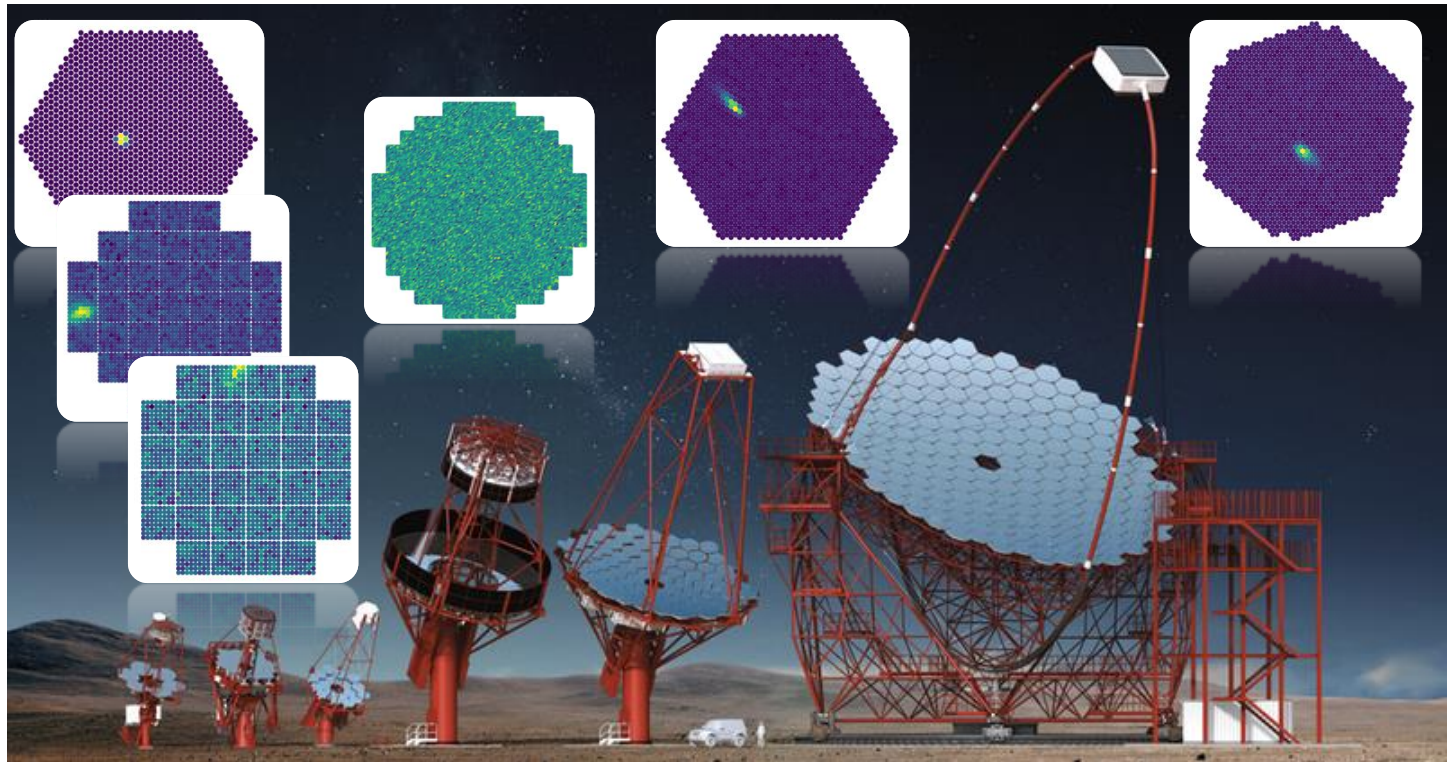
- Stereoscopic view of the extended air showers
- Compact “videos” rather than single snapshots
- Events effectively recorded in 4D!



CREDIT: DESY/Milde Science Communication

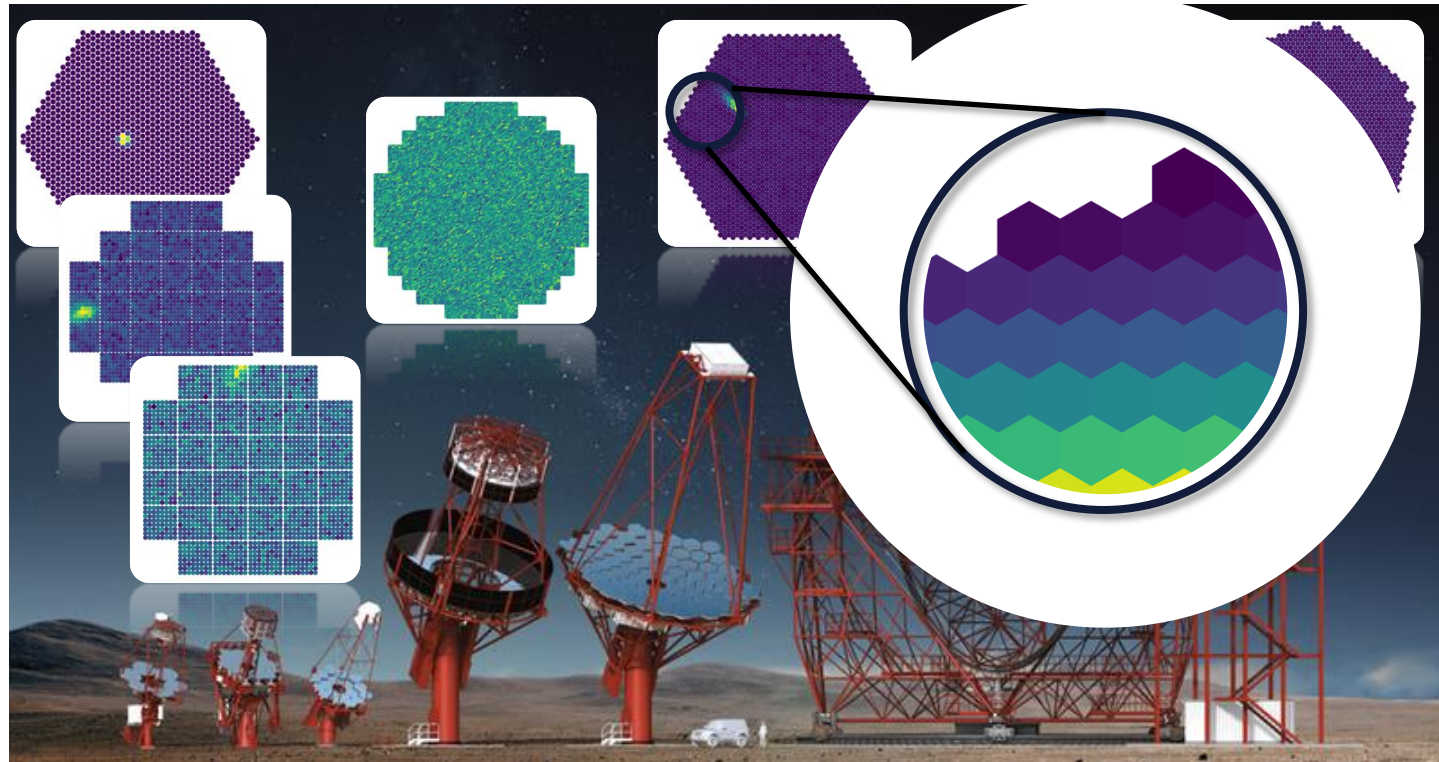
- Heterogeneity of instruments:

Camera images courtesy of T. Vuillaume



- Heterogeneity of instruments:

Camera images courtesy of T. Vuillaume

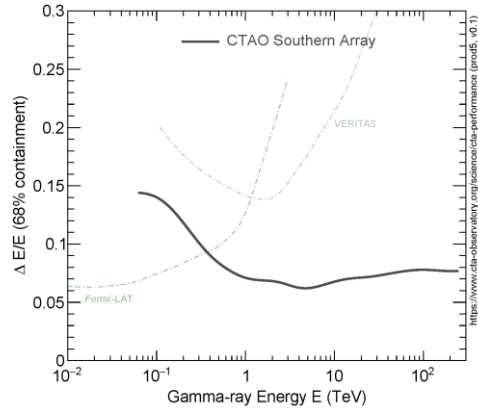


Hexagonal pixels

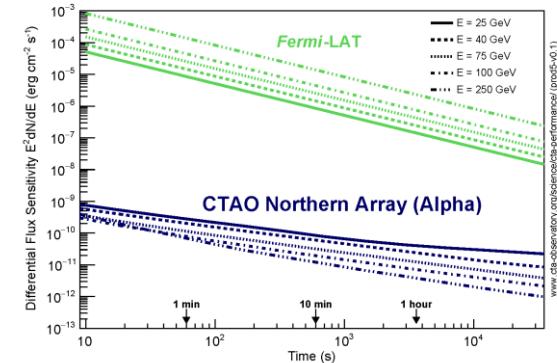
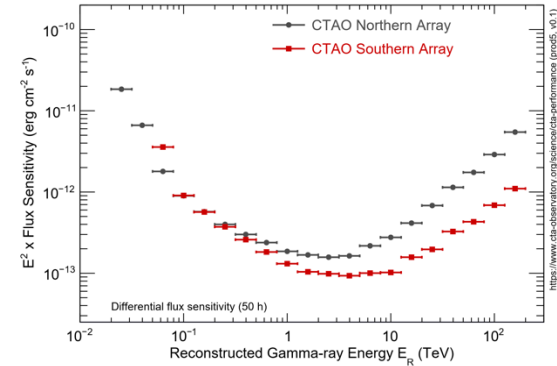
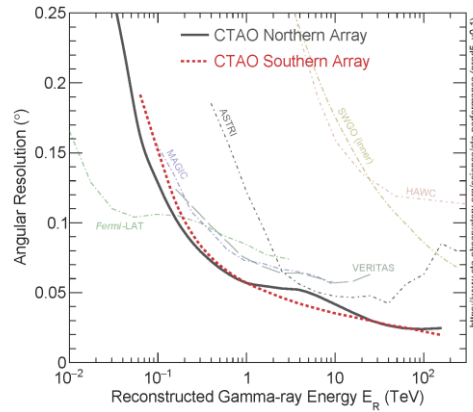
- Final metrics are far from trivial and entangled

Flux sensitivity

Energy resolution



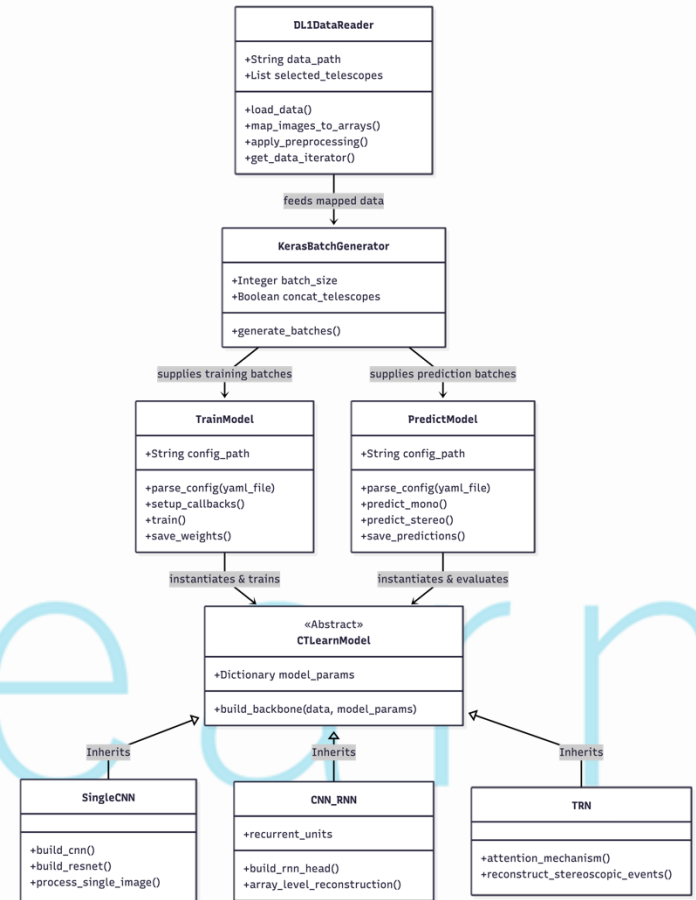
Angular resolution



- High-level Python package for using **deep learning for IACT event reconstruction**
- Configuration-file-based workflow and installation with conda drive reproducible training and prediction
- Supported ML backends:



- Open source:
 - github.com/ctlearn-project/ctlearn
 - ctlearn.readthedocs.io
 - zenodo.org/records/15021424



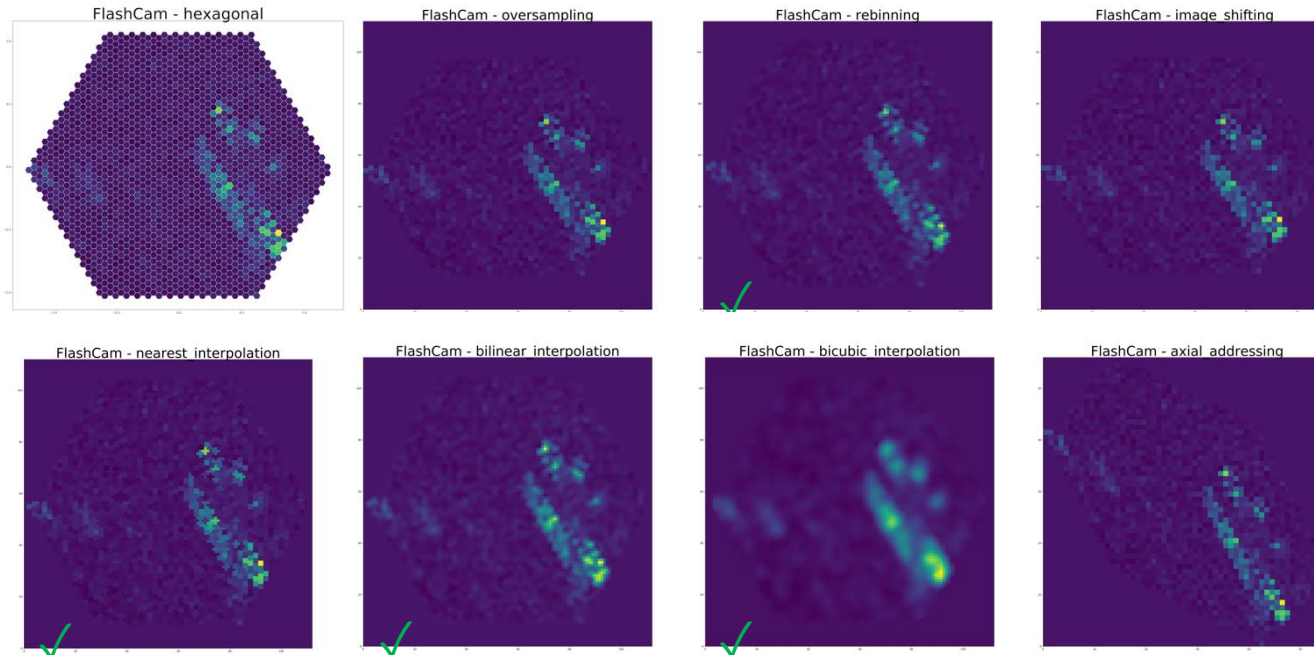
- Image mapping (preprocessing)



A. Brill, B. Kim, Q. Feng,
D. Nieto, T. Miener,
et al.



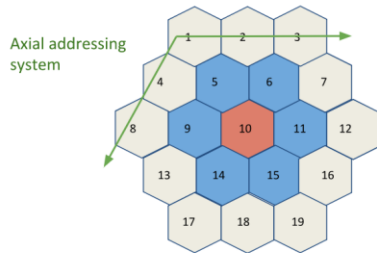
<https://github.com/ctlearn-project/>



✓ Angles and distances preserved

- Hexagonal convolution

- Convolution



Convolution kernel

Index matrix

1	2	3		
4	5	6	7	
8	9	10	11	12
	13	14	15	16
		17	18	19

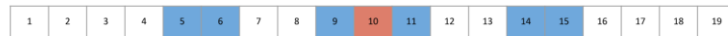


Image stored as a vector



T. Vuillaume,
M. Jacquemont, et al.

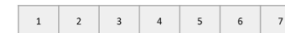
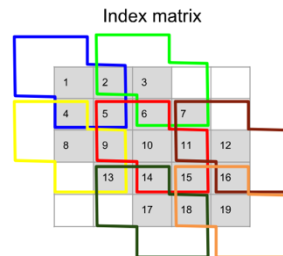


<https://github.com/IndexedConv>

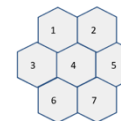
$W \times$

5		
6		
9		
10		
11		
14		
15		

- Pooling

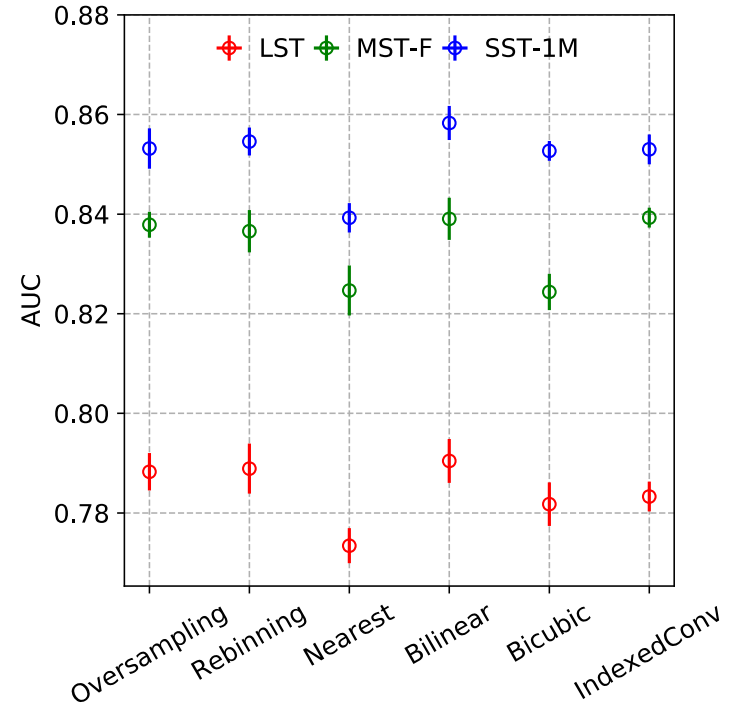
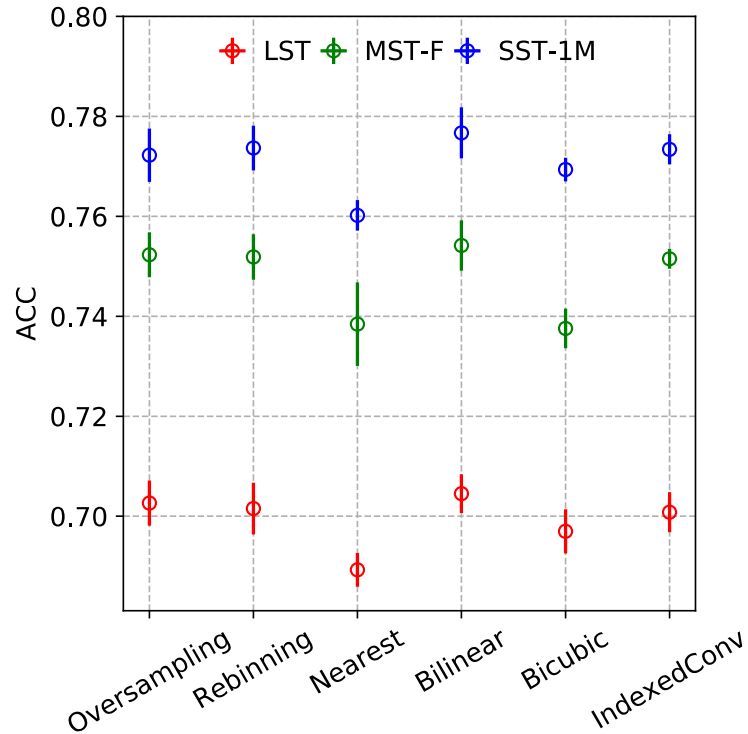


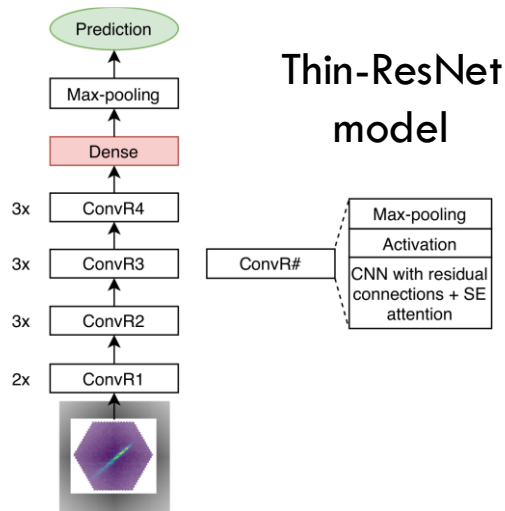
Rebuild index matrix



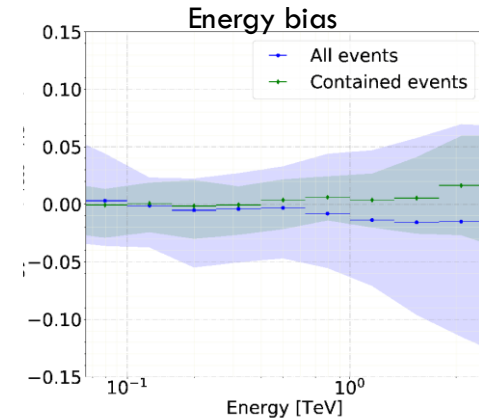
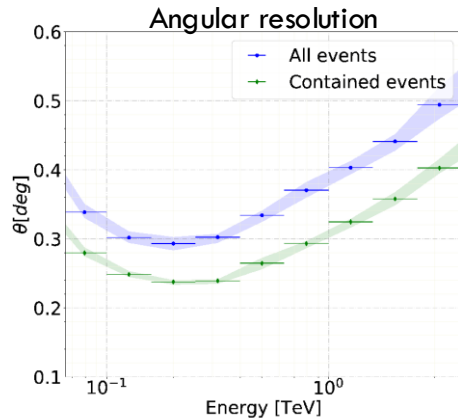
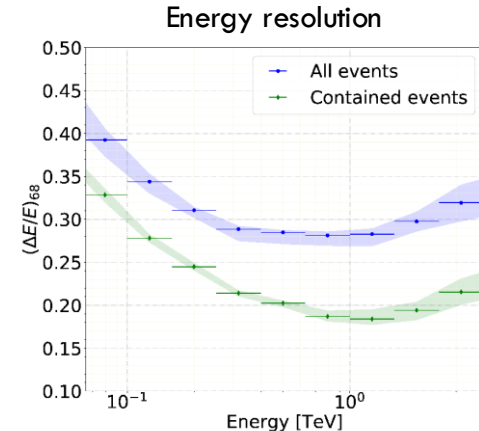
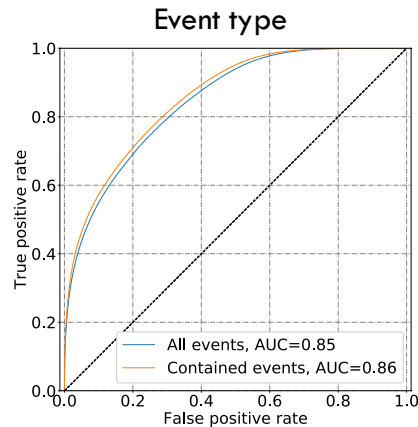
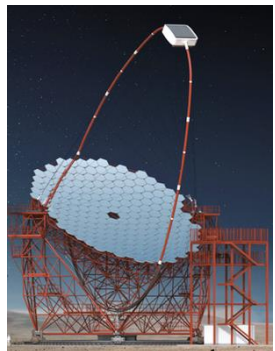
(M. Jacquemont et al. 2019)

- Comparison of methods for classification task





Full-event reconstruction for single-telescope data achieved!

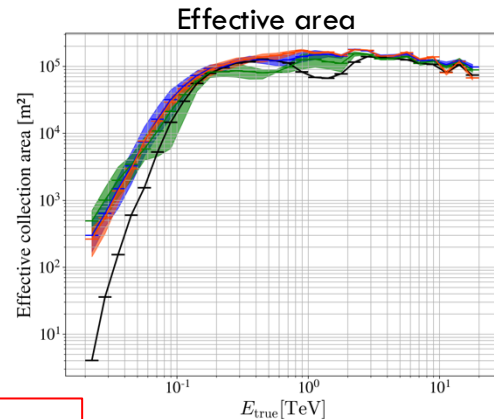
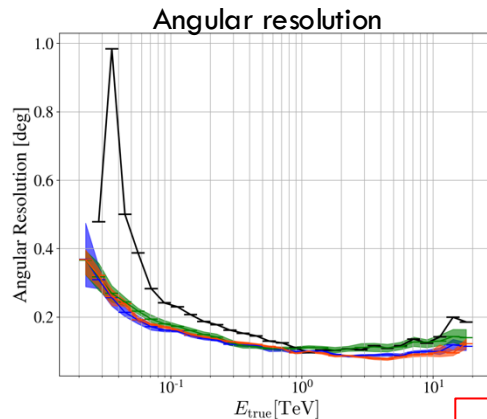


[D. Nieto et al. ADASS XXX 2020](#)

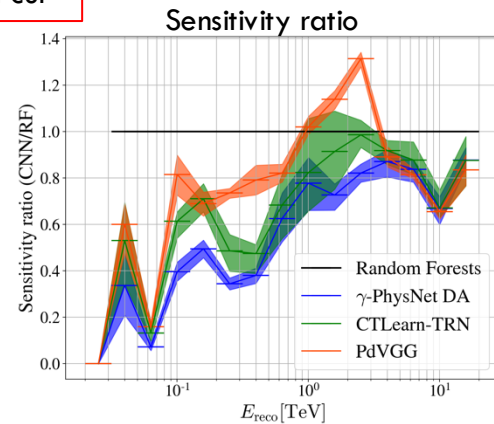
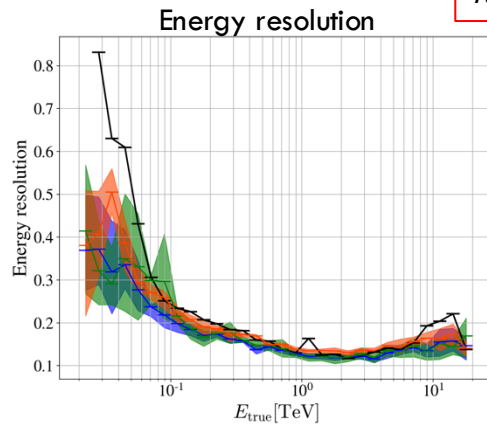


UNIVERSITÀ
DEGLI STUDI
DI PADOVA

- Crosschecking three different implementations
- Same datasets, same cuts
- Different models
- Comparison against standard analysis (RF)

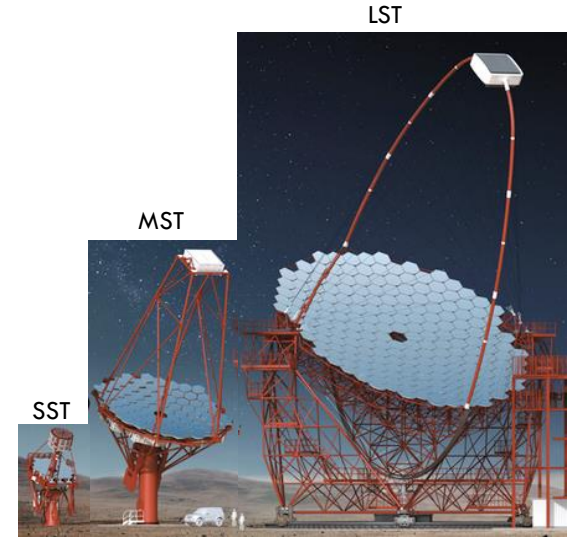
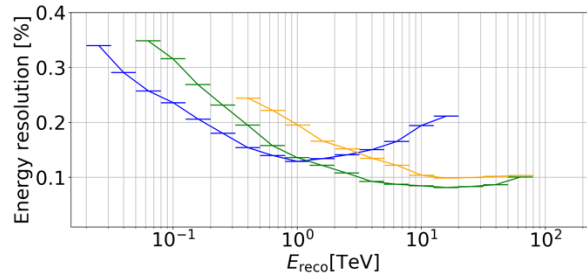
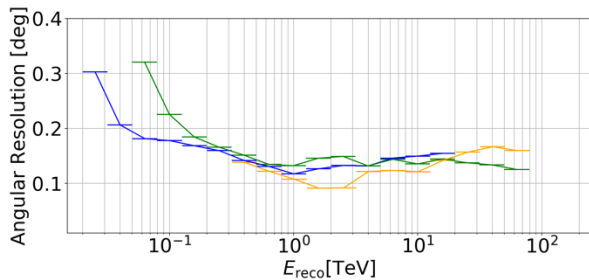
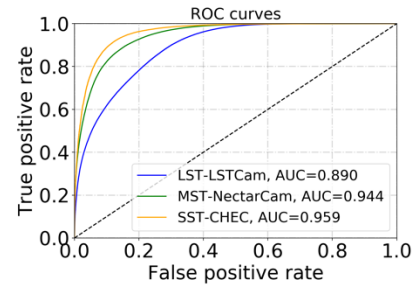
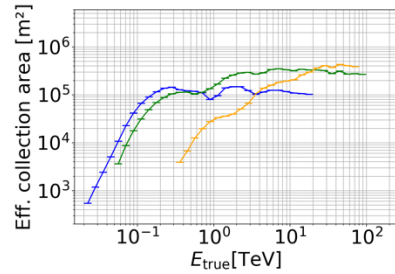
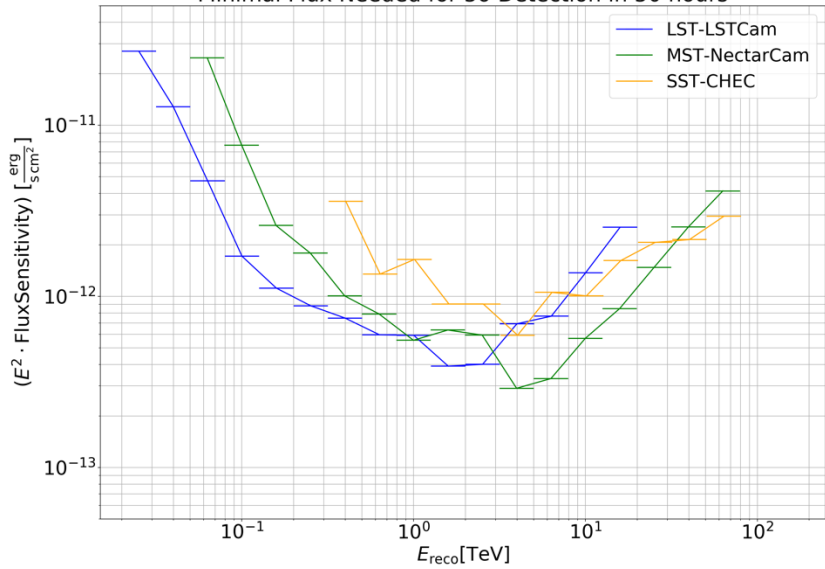


Mid cut

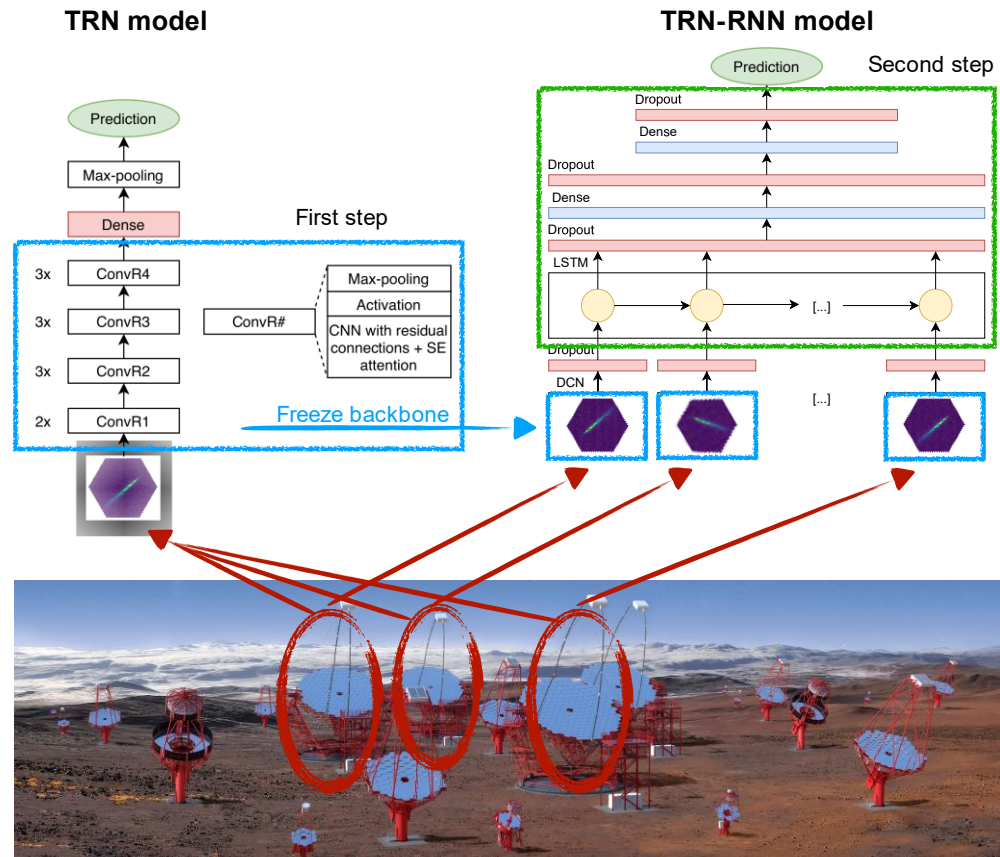


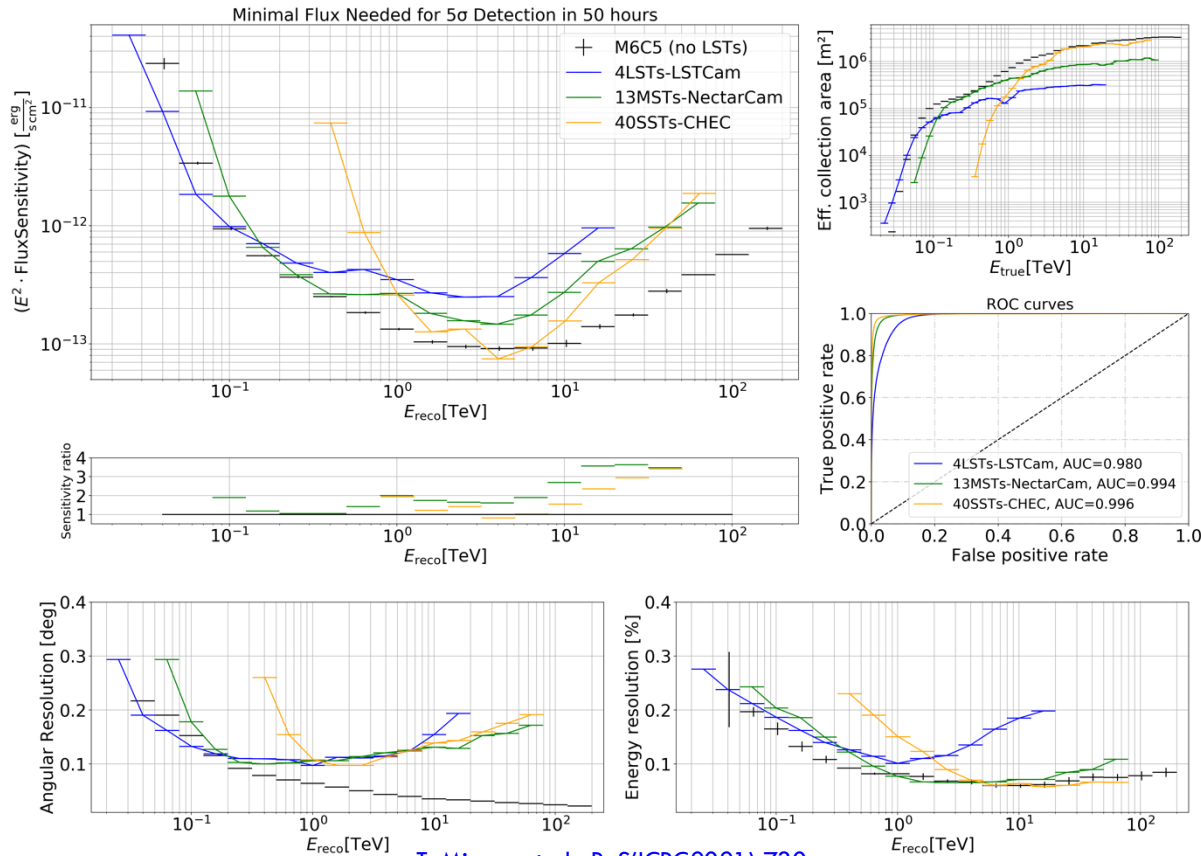
[P. Grespan et al. PoS\(ICRC2021\) 771](#)

Minimal Flux Needed for 5 σ Detection in 50 hours

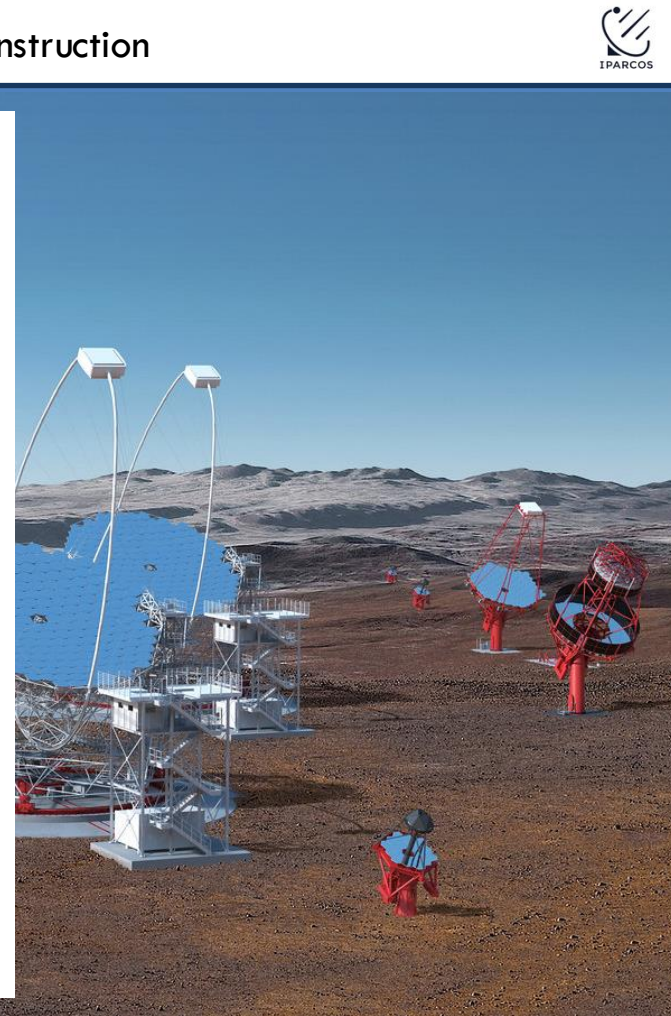


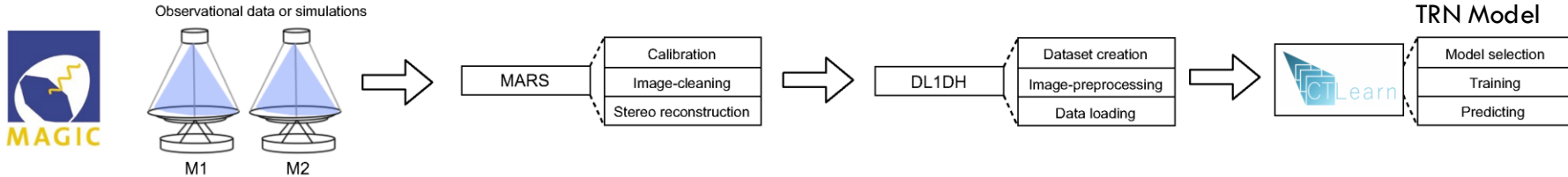
[T. Miener et al., PoS\(ICRC2021\) 730](#)



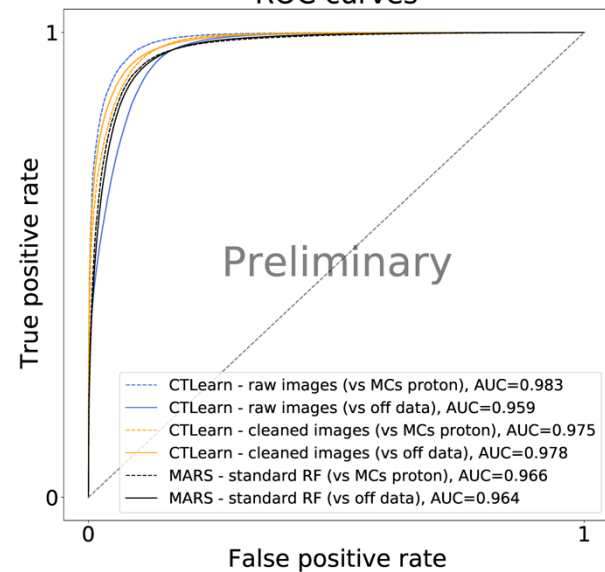


[I. Miener et al., PoS\(ICRC2021\) 730](#)

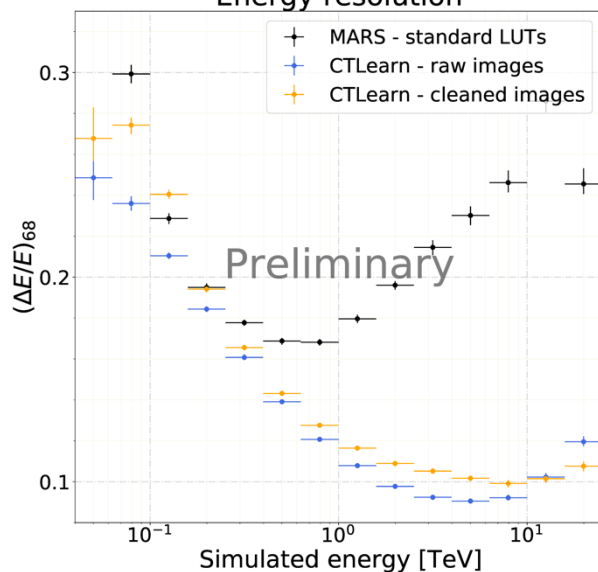




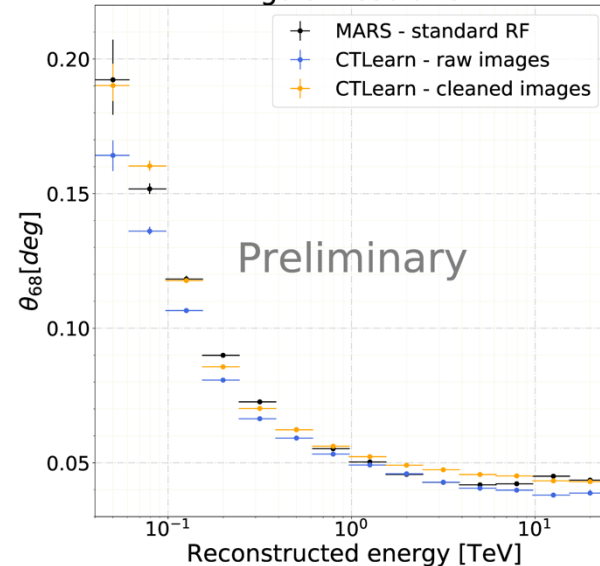
ROC curves



Energy resolution



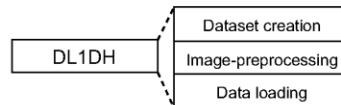
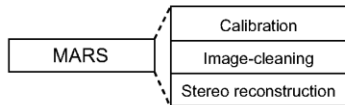
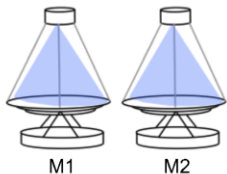
Angular resolution



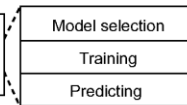
[T. Miener et al. 2021 \(ADASS XXXI\)](#)



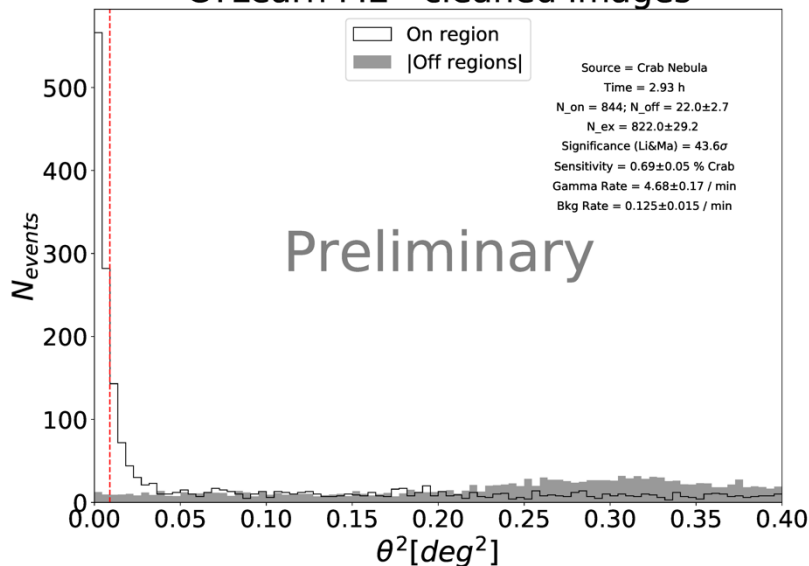
Observational data or simulations



TRN Model



CTLearn ME - cleaned images



Analysis	γ rate [/min]	bkg rate [/min]	Sen. [% Crab]	Sig. (Li&Ma)
MARS – ME	4.54 ± 0.16	0.119 ± 0.015	0.70 ± 0.05	43.0σ
CTLearn – ME (raw)	3.45 ± 0.14	0.133 ± 0.018	0.97 ± 0.08	36.5σ
CTLearn – ME (cleaned)	4.68 ± 0.17	0.125 ± 0.015	0.69 ± 0.05	43.6σ
MARS – LE	16.49 ± 0.35	3.861 ± 0.086	1.09 ± 0.03	61.1σ
CTLearn – LE (raw)	11.70 ± 0.32	3.832 ± 0.114	1.53 ± 0.05	47.5σ
CTLearn – LE (cleaned)	16.24 ± 0.35	3.872 ± 0.086	1.11 ± 0.03	60.4σ

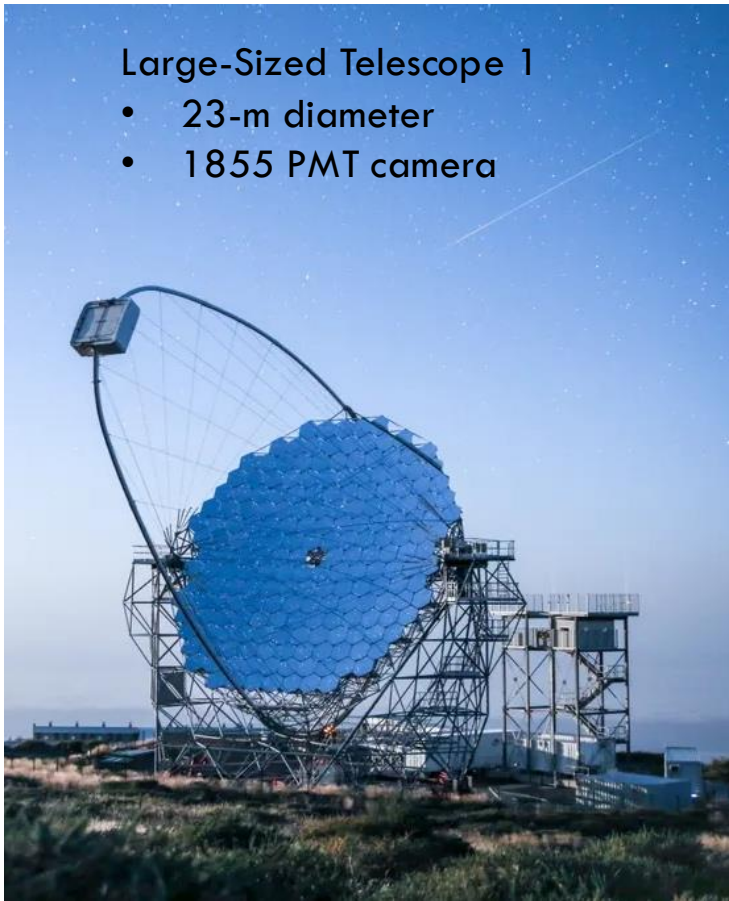
Analysis	N_{on}	N_{off}	N_{ex}
MARS – ME	819	21.0 ± 2.6	798.0 ± 28.7
CTLearn – ME (raw)	629	23.3 ± 3.1	605.7 ± 25.3
CTLearn – ME (cleaned)	844	22.0 ± 2.7	822.0 ± 29.2
MARS – LE	3579	679.0 ± 15.0	2900.0 ± 61.7
CTLearn – LE (raw)	2730	673.7 ± 20.0	2056.3 ± 56.0
CTLearn – LE (cleaned)	3536	680.7 ± 15.1	2855.3 ± 61.3

Summary of all performed analyses of the same Crab Nebula sample

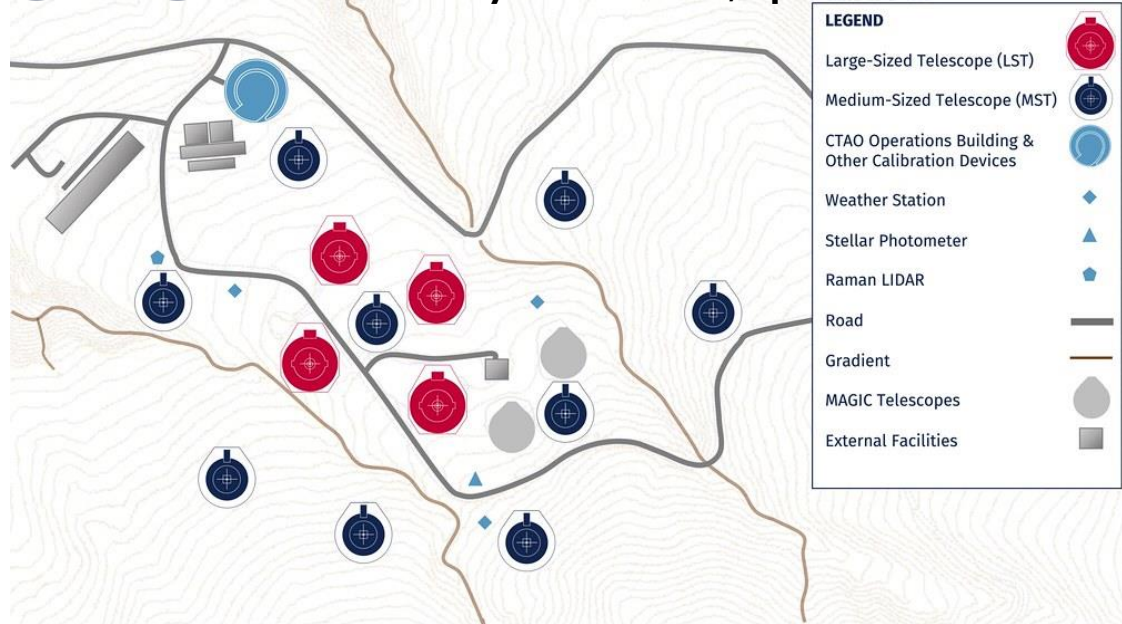
[T. Miener et al. 2021 \(ADASS XXXI\)](#)

Large-Sized Telescope 1

- 23-m diameter
- 1855 PMT camera

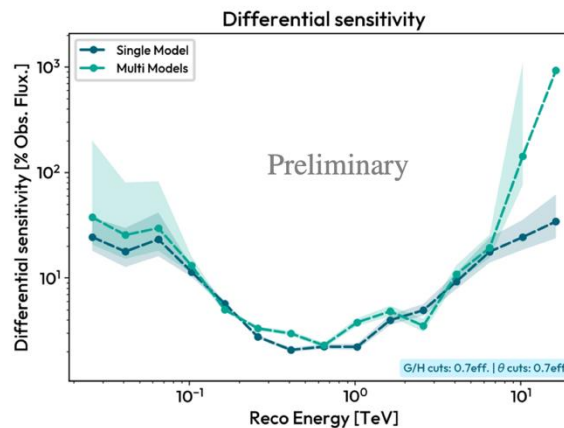
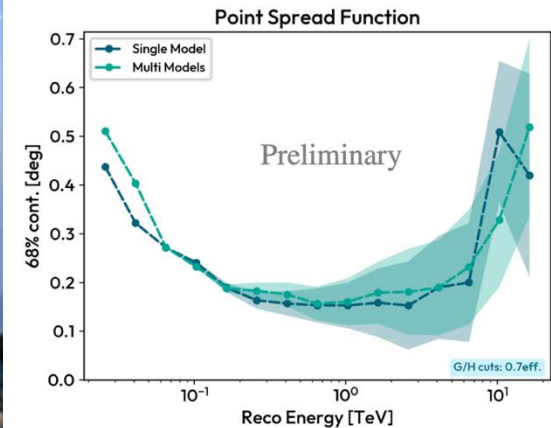
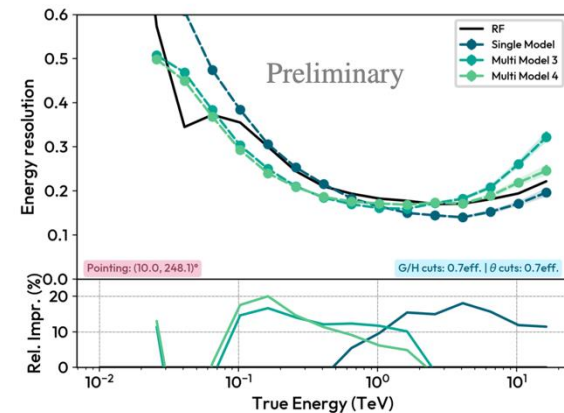
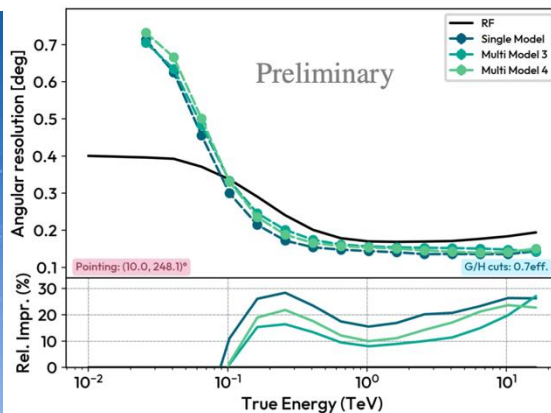


CTAO Northern Array in La Palma, Spain



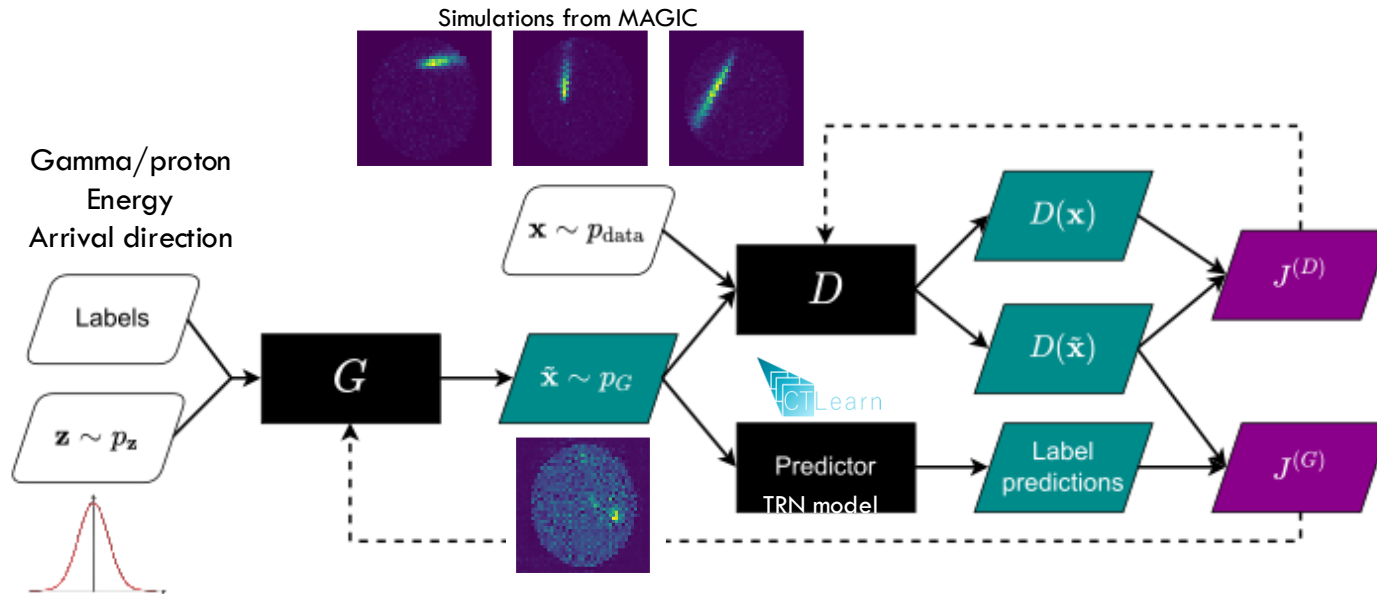
Large-Sized Telescope 1

- 23-m diameter
- 1855 PMT camera



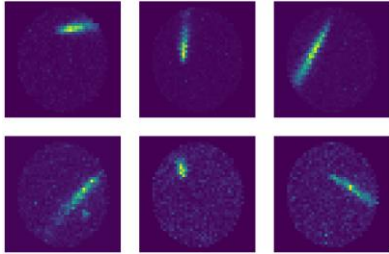
B. Lacave et al. (ICRC 2025)

- Auxiliary conditional generative adversarial networks (AC-GANs)

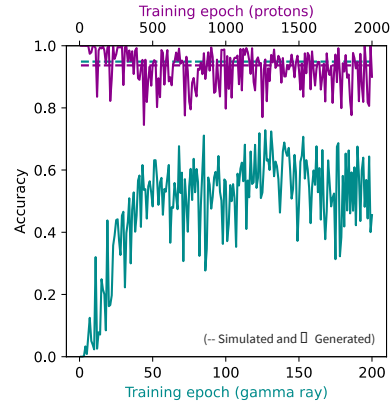
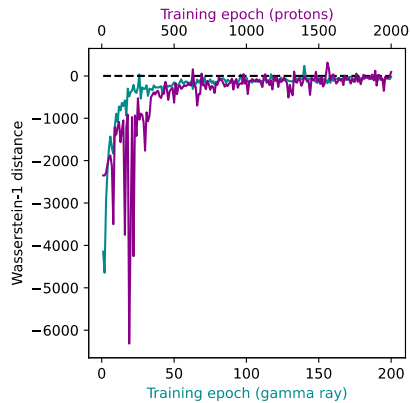
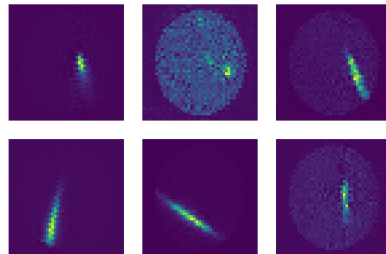


GAMMA RAYS

Simulated

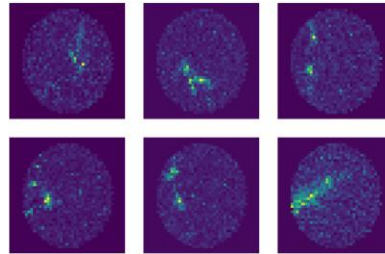


Generated

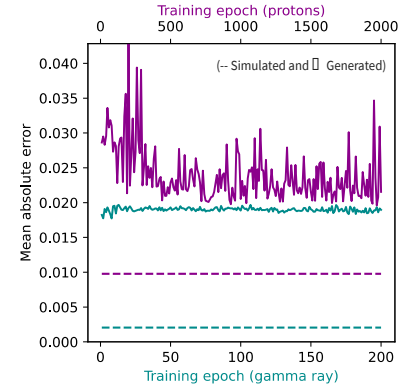
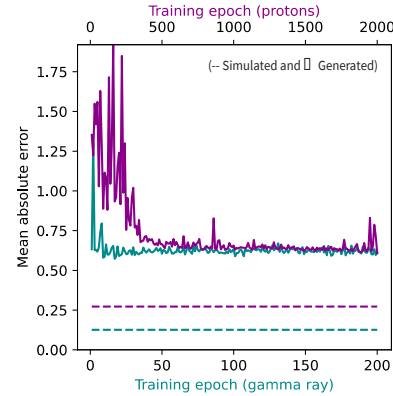
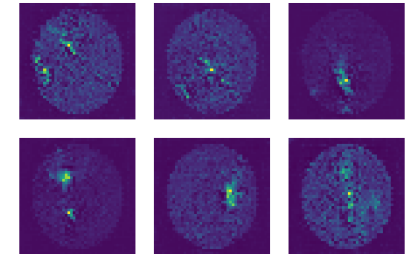


PROTONS

Simulated

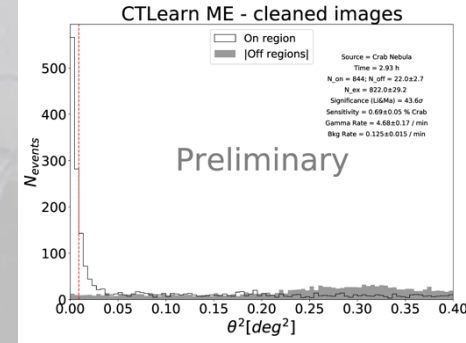
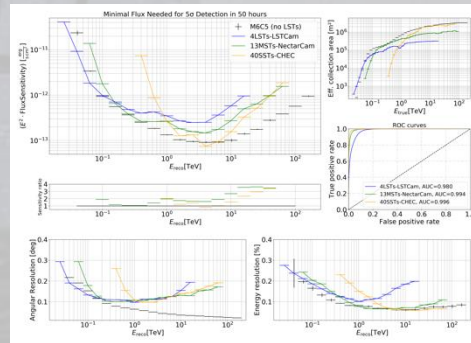
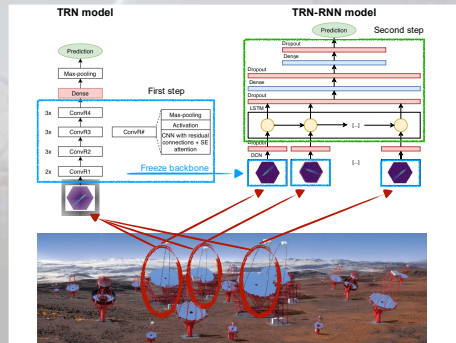
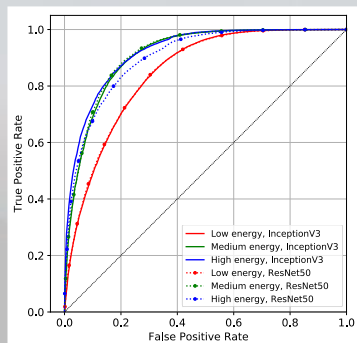


Generated



S. García-Heredia et al.

- Current-generation IACTs have enhanced their performances through ML
- **Next-gen** (even current-gen!) IACT may profit from **latest developments in ML**
- Ongoing efforts to exploit **deep learning** as an event reconstruction method for **IACTs**
 - Full-event reconstruction over simulated IACT events demonstrated
 - Application to real observations works!
 - Working on optimizing architectures & multi-task learning
 - Using AC-GANs as pseudosimulators
 - Tackling the real-data problem



We acknowledge the Spanish Ministry of Science and Innovation and the Spanish Research State Agency (AEI) through the government budget lines PGE2022/28.06.000X.711.04, 28.06.000X.411.01, and 28.06.000X.711.04 of PGE 2023, 2024, and 2025 funding grant PID2022-138172NB-C42 and grant PDC2023-145839-I00 funded by the Spanish MCIN/AEI/10.13039/501100011033 and by ERDF/EU and NextGenerationEU PRTR. We also acknowledge the grant “Tecnologías avanzadas para la exploración del universo y sus componentes” (PR47/21 TAU), funded by Comunidad de Madrid, by the Recovery, Transformation and Resilience Plan from the Spanish State, and by NextGenerationEU from the European Union through the Recovery and Resilience Facility. The authors gratefully acknowledge the computer resources at Artemisa and the technical support provided by the Instituto de Física Corpuscular, IFIC (CSIC-UV). Artemisa is co-funded by the European Union through the 2014-2020 ERDF Operative Programme of Comunitat Valenciana, project IDIFEDER/2018/048.

