Machine-Learning-based CTAO Telescope data processing

Financiado por la Unión Europea NextGenerationEU
Plan de Recuperación, Transformación y Resiliencia
AGENCIA ESTATAL DE INVESTIGACIÓN
Comunidad de Madrid
Astrofísica y Física de Altas Energías

III CONGRESO IPARCOS 2024

3rd IPARCOS Meeting – December 2024

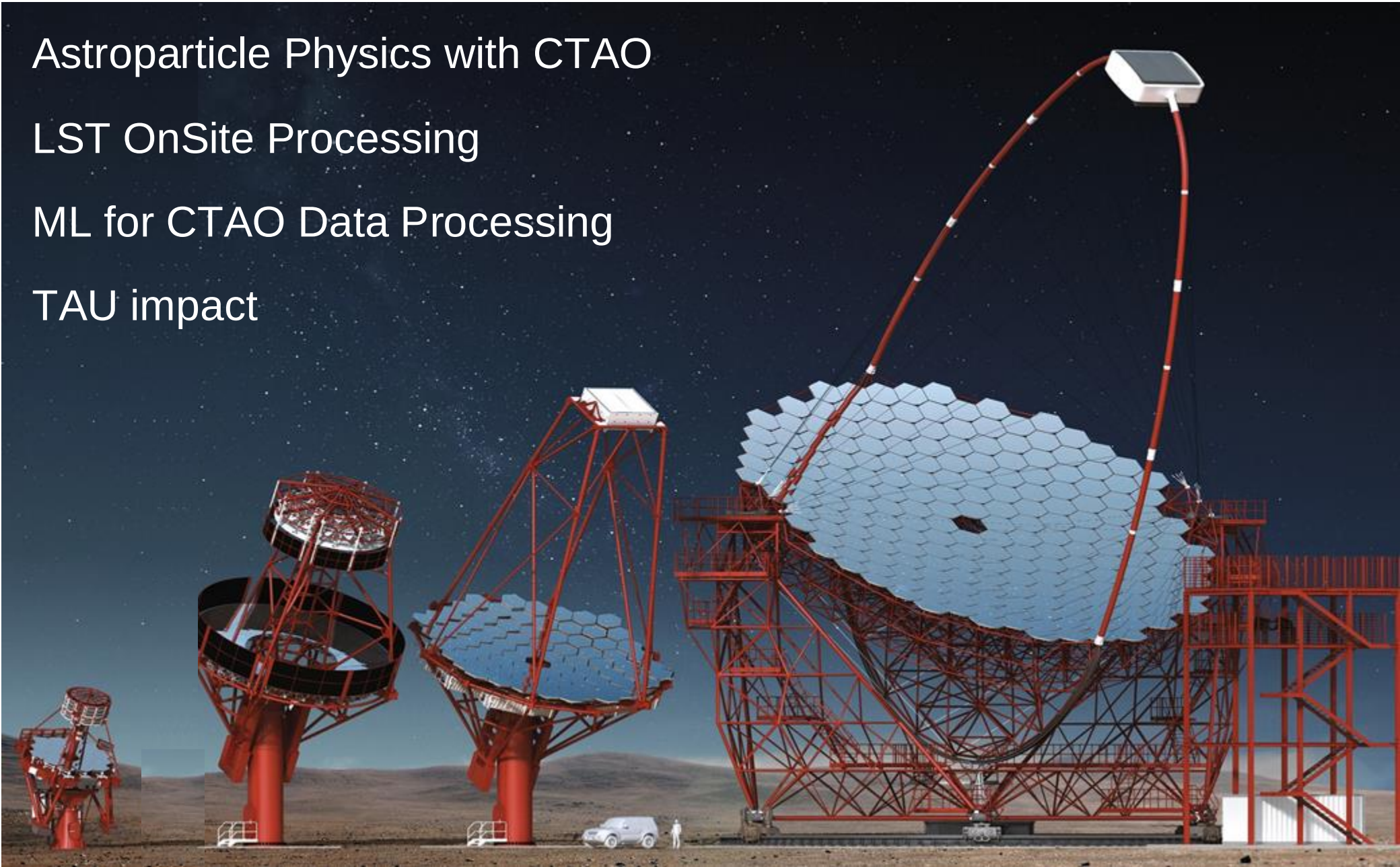
A. Pérez-Aguilera, A. Cerviño, M. Laínez, D. Martín, J. Buces, M. Molina, L.A. Tejedor, D. Nieto, J.L. Contreras, M. López, J.A. Barrio

IPARCOS

UNIVERSIDAD COMPLUTENSE MADRID

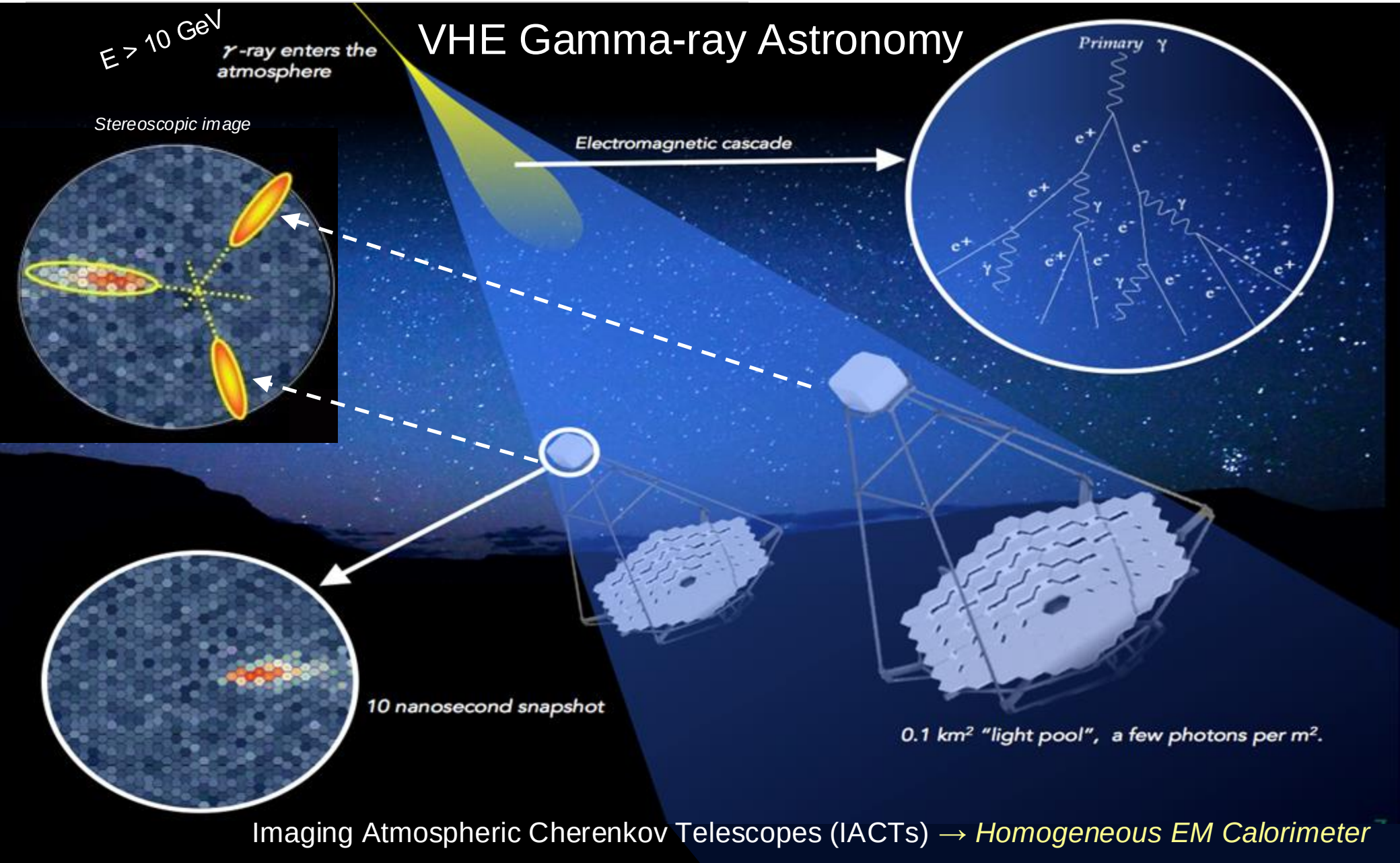
High Energy Physic Group – Institute of Particles and Cosmos Physics – University Complutense of Madrid

Contents



Astroparticle Physics

VHE Gamma-ray Astronomy



Imaging Atmospheric Cherenkov Telescopes (IACTs) → *Homogeneous EM Calorimeter*

CTA Observatory

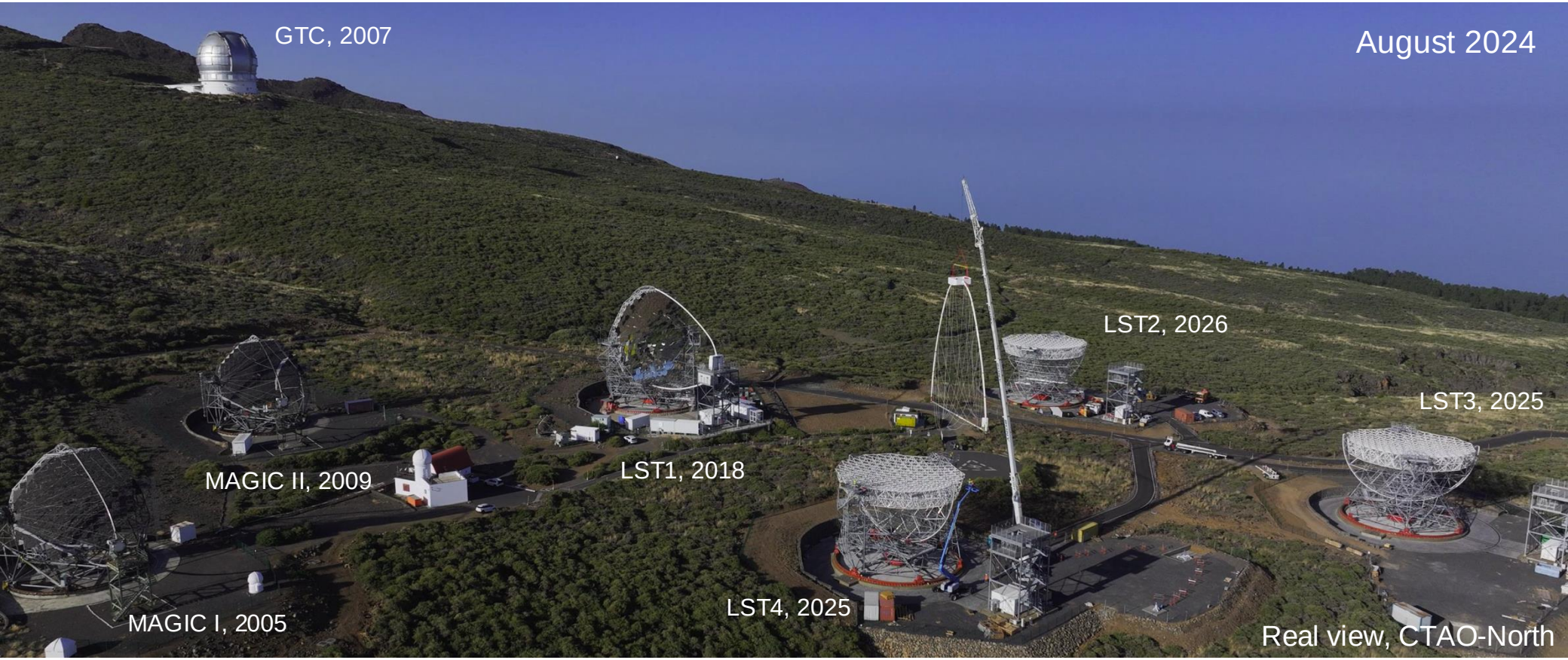
Cherenkov Telescope Array Observatory

Sensitivity improvement x10
Energy range extension x10
Angular resolution improvement

Two sites:
La Palma (*Canaries*) / Chile
~100 telescopes

CTA Observatory

Cherenkov Telescope Array Observatory



Sensitivity improvement x10
Energy range extension x10
Angular resolution improvement

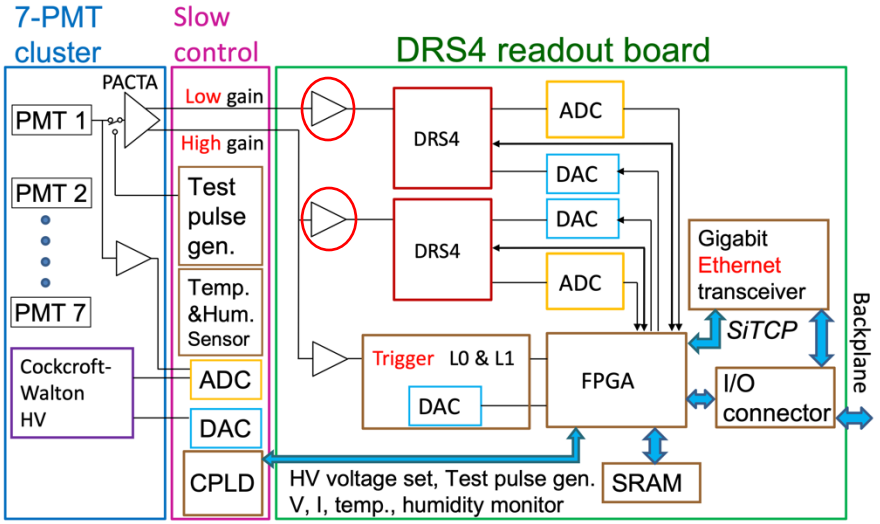
Two sites:
La Palma (*Canaries*) / Chile
~100 telescopes

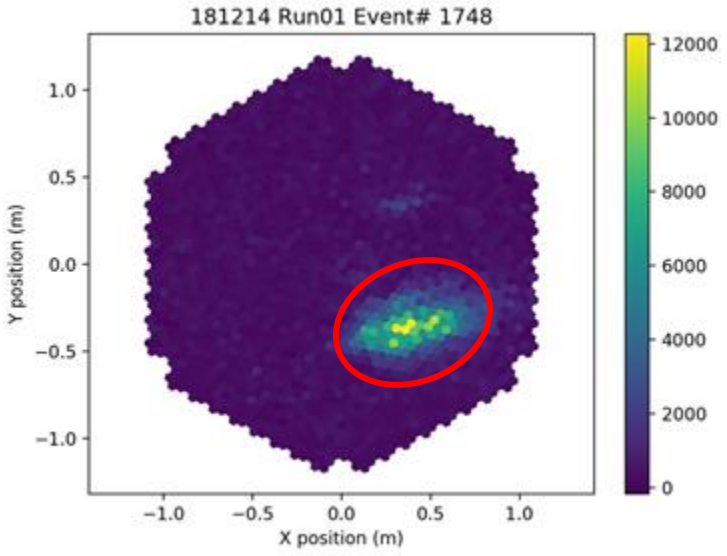
Data Processing @ CPUs

A. Cerviño
PhD student session

LST OnSite Processing for Data Volume Reduction

- OnSite Processing
  - LST1 alone produces ~10-20 TB/night raw data; 3 more LSTs to enter commissioning in ~2025
  - Data is processed on-site in a *temporary* data center at the telescope base, with ~1800 cores and 5.3 PB HD
  - Onsite pipeline process raw data every morning in a highly parallel way
  - Building on MAGIC OnSite experience → until recently saving all raw data



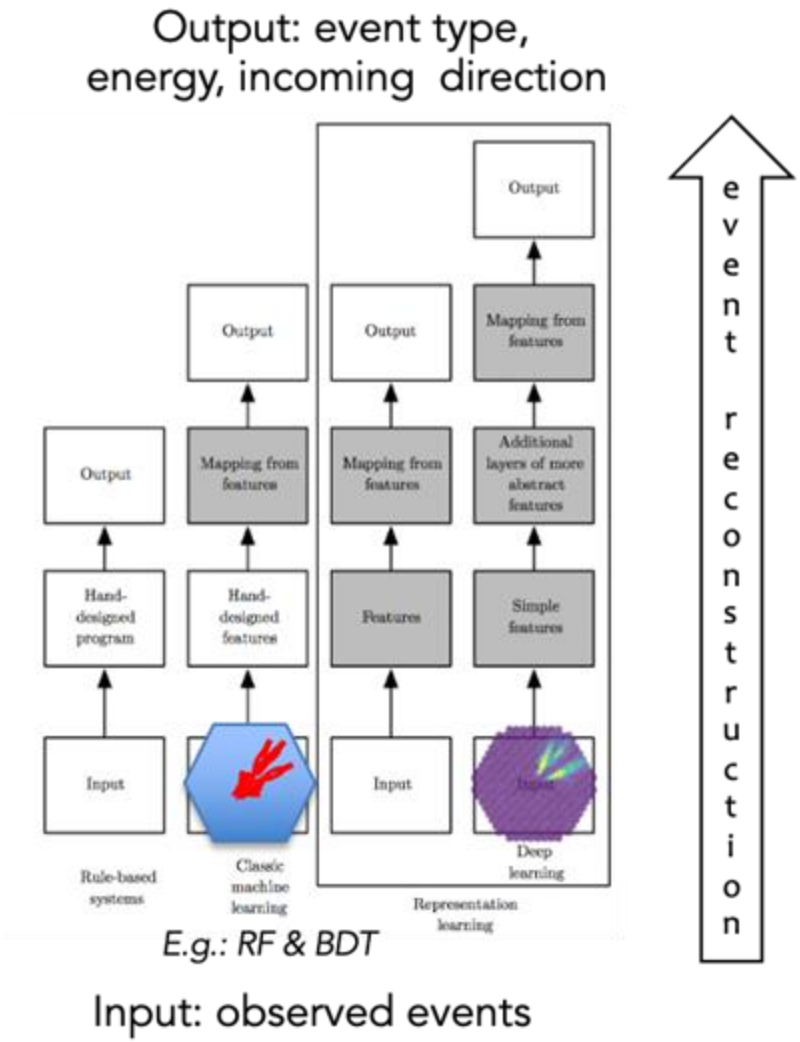


- Data Volume Reduction @ OnSite pipeline
  - 1st step: select only one of the 2 PMT-amplifier gains
  - 2nd step: Region of Interest selection, along with UAH
  - LST as test-bed for CTAO DVR

ML@GPUs for CTAO Reco

CTLearn

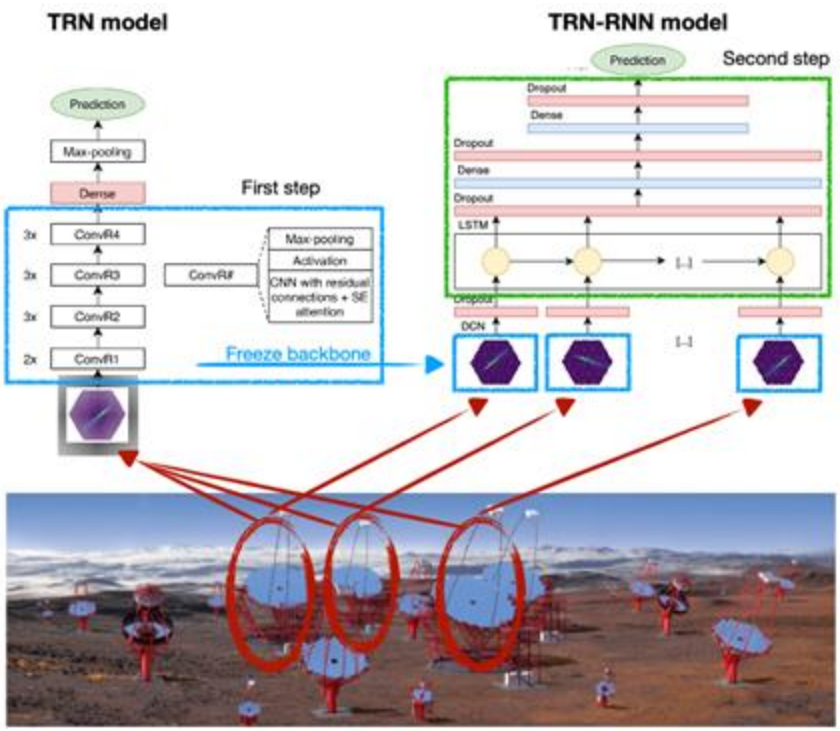
- High-level Python package for using deep learning at IACT:
  - gamma/hadron/Night-Sky-Background separation
  - Event Reconstruction
  - ➢ Based of pixel information
- Founded at GAE, led by GAE and UniGe
- Open source:
  https://github.com/ctlearn-project/ctlearn
  https://zenodo.org/records/11475531

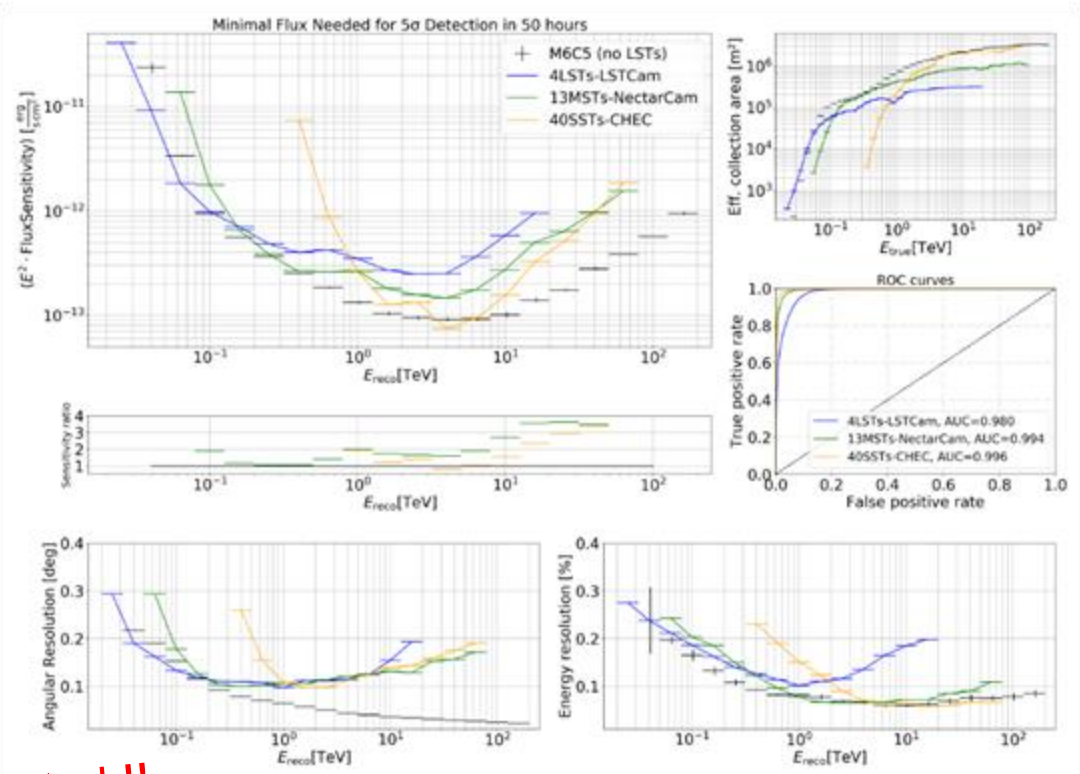


ML@GPUs for CTAO Reco

CTLearn *on simulation & real data*

A. Cerviño
PhD student session





Application to CTAO simulated data demonstrated !!

Application to MAGIC real data demonstrated !!

T. Miener et al., PoS(ICRC2021) 730

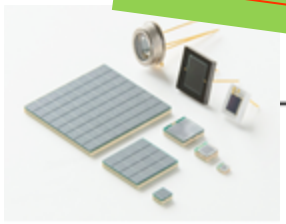
LST1 real data coming soon

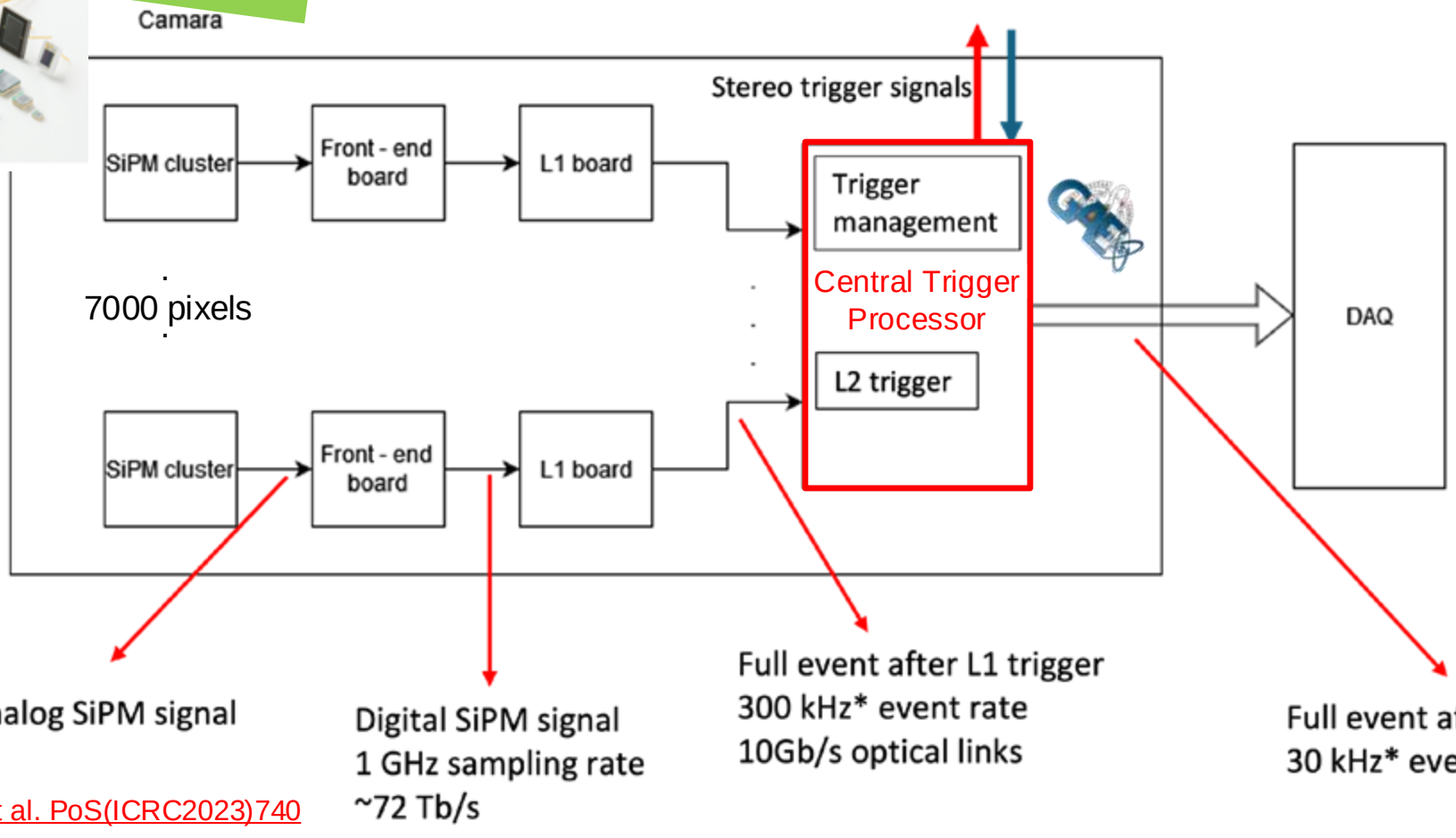
ML@FPGAs for CTAO Trigger

Advanced LST SiPM Camera*

V. Moya
PhD student session

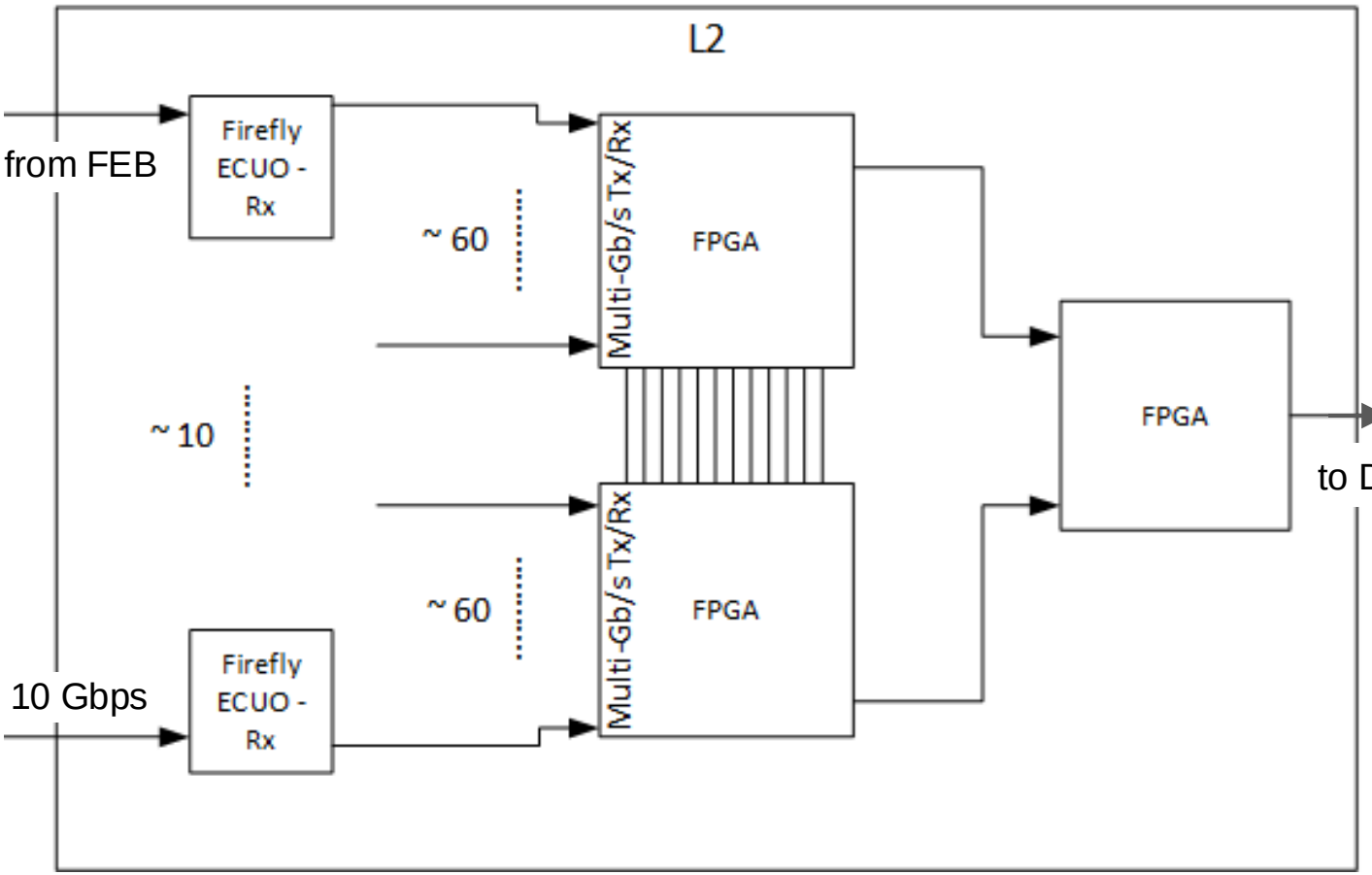
Candidate for mid-term upgrade of CTAO telescope cameras





*M. Heller et al. PoS(ICRC2023)740

Central Trigger Processor: ML-based L2 trigger on FPGAs



L2 Trigger O(1 μs) latency

FPGAs in layers to allow scalability

First layer of FPGAs for data reception, formatting and first layers of ML algorithms

Last FPGA will take care of the fuull-image ML layers, Event Building and sending the data to DAQ.

Tentative FPGA models: Kintex UltraScale KU085, KU095 or KU115, with moderate cost and a large number of high-speed resources.

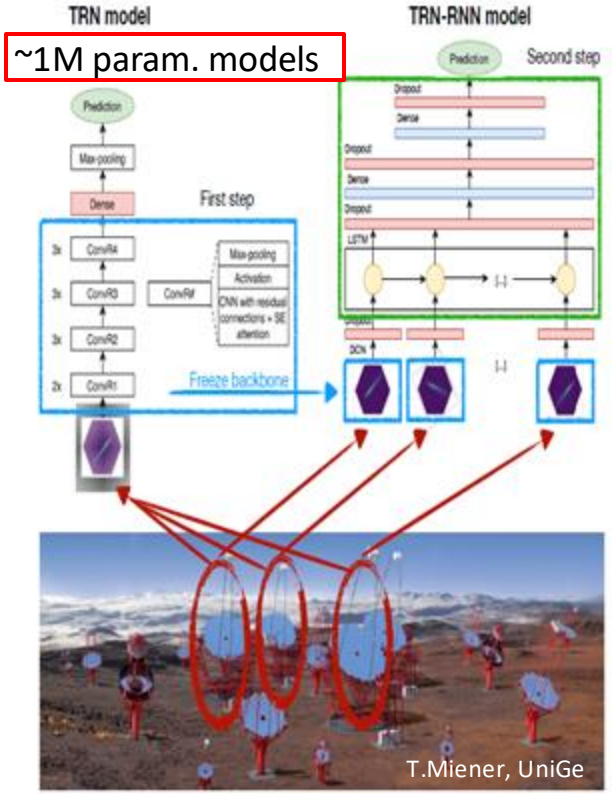
Prelim. number of high-speed channels, still under design

ML@FPGAs for CTAO Trigger

ML algorithms for CTAO camera trigger

A. Pérez-Aguilera
PhD student session

- Train small supervised CNNs & optimize large CNNs to fit in FPGAs → promising shower/NSB separation
- Train unsupervised DBSCAN models (along with UniGe)
- Translating ML models to firmware code



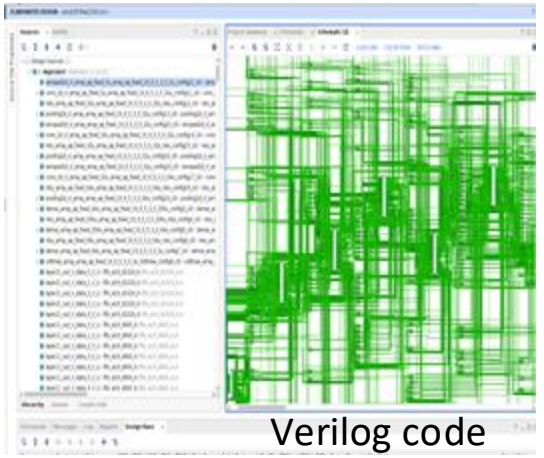
~2k param. models

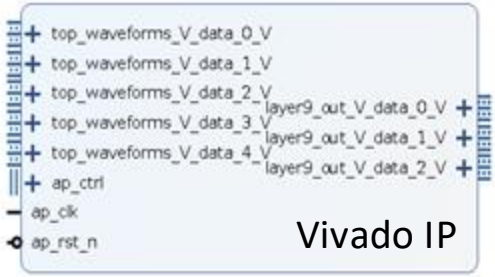
Model: "CTLearn_model"

Layer (type)	Output Shape	Param #
waveforms (InputLayer)	[(None, 30, 30, 5)]	0
SingleCNN_block (Functiona l)	(None, 16)	1536
fc_particletype_1 (Dense)	(None, 32)	544
particletype (Dense)	(None, 3)	99
type (Softmax)	(None, 3)	0

Total params: 2179 (8.51 KB)
Trainable params: 2179 (8.51 KB)
Non-trainable params: 0 (0.00 Byte)

hls4ml

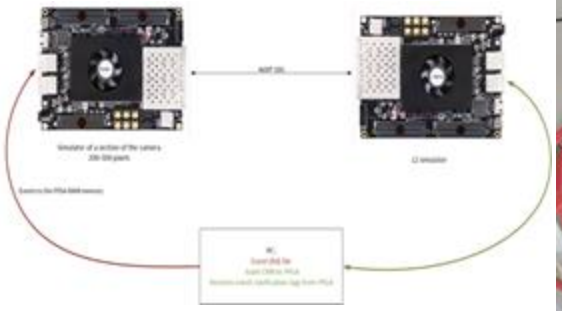

Verilog code

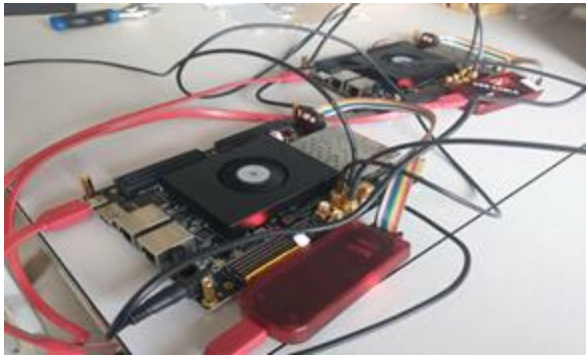


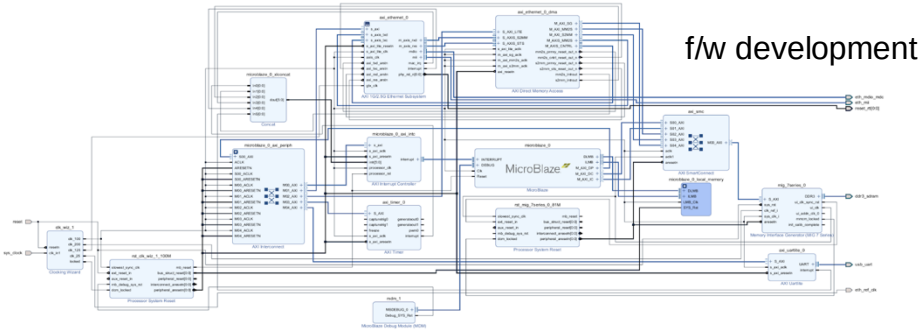
ML@FPGAs for CTAO Trigger

Demonstrators:

- CNN@FPGA trigger test-bench
  - 2 development boards with Kintex UltraScale-tier FPGAs
  - One board to simulate pixel data information; the other to implement firmware core and L2 algorithms

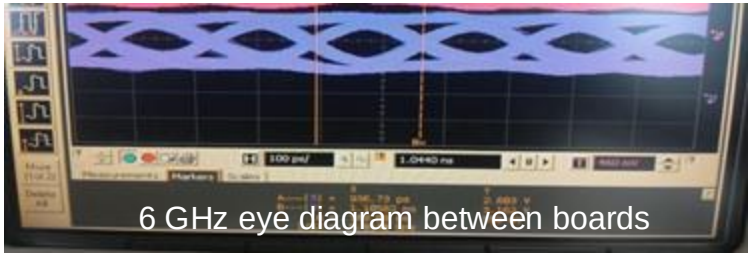




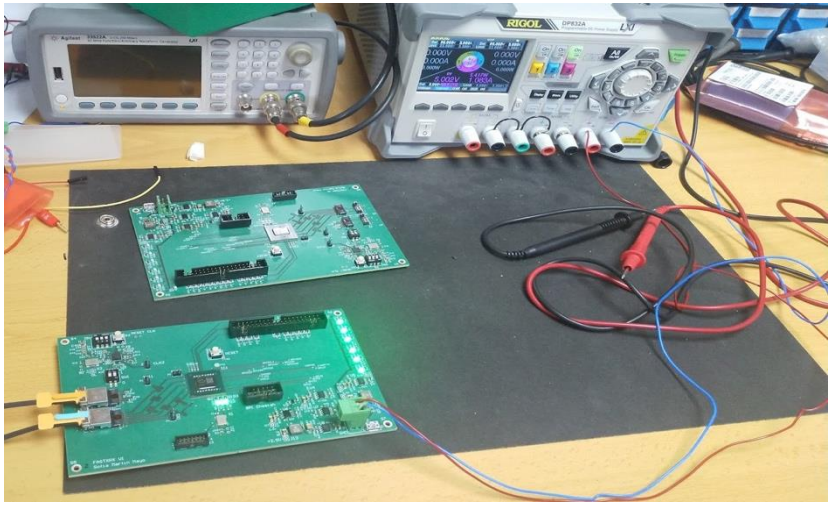


f/w development

- CTP high-speed test-bench
  - Multi-Gbit *Firefly* optical transceiver + 12-layer Microwave PCB + FPGA interface.
  - Test the capabilities and quality of the companies to produce the final board



6 GHz eye diagram between boards



TAU* Impact

- Scope & impact:
  - TAU-CTA: ~92 k€ (UCM ~500 k€, CM ~1.5 M€, Cataluña ~15 M€)
  - Engineer costs: A. Pérez-Aguilera, May 2023 – June 2025
  - Support of initial involvement in Advanced LST SiPM-CAM
  - Current design and prototyping of a SiPM-CAM key element
  - Allow to apply for other grants (PDC2023 ✓ , EU INFRA-Tech 🤞)
  - Transfer of knowledge started for fast ML-based image processing on FPGAs
- Network
  - ***Spain***: CNID/COMCHA (ML@xx, OnSite Proc.); Ciemat/IFIC (ML@FPGAs), UPM (HLS)
  - ***International***: SiPM-CAM (Ciemat, UB, UniGe, INFN); CERN DRD7 (ML@FPGAs)
- Plans (2026; 2028):
  - ***LST OnSite***: automatic processing to the 4 LSTs array; integrate Onsite Pipeline in CTAO scheme
  - ***CTLearn***: compliance with CTAO software requirements and pipeline integration, optimize DL models for real data; CTAO offline ML-based DVR
  - ***ML@FPGAs***: deploy & benchmark simple CNNs, optimization of CNNs for FPGAs & start knowledge transfer; build full-scale CTP prototype, benchmark optimized CNNs

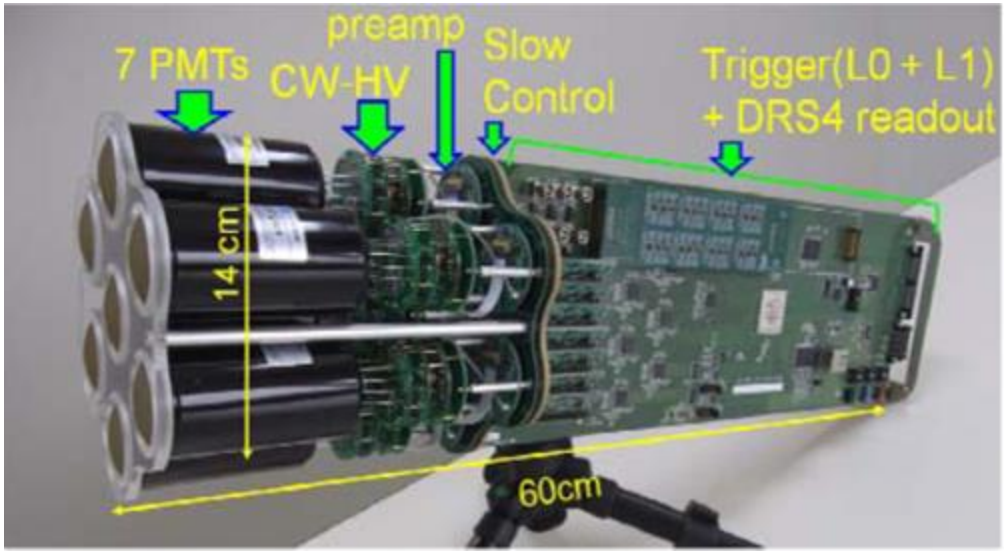
* Planes Complementarios - AstroHEP: Tecnologías Avanzadas para la Exploración del Universo y sus Componentes, Comunidad de Madrid

Summary

- CTAO ESFRI construction started
- Recent funds → UCM-GAE involvement in ML-based R&D for CTA
- Sinergies with Spanish/International teams pursued
- Transfer of knowledge pursued/expected from ML@FPGA activities

Backup

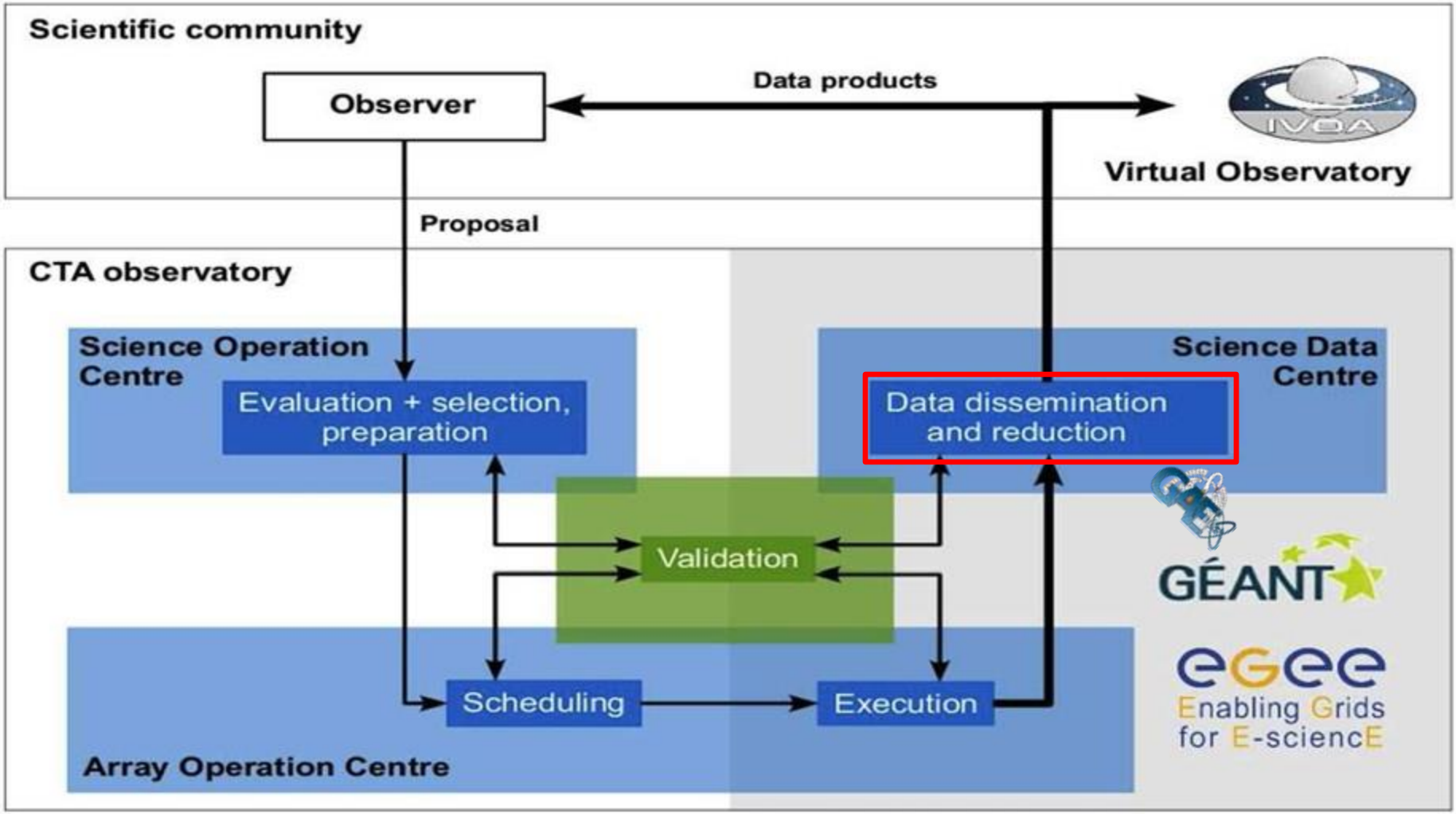
Cherenkov Telescope Array Observatory



~2000-pixel PMT-based camera

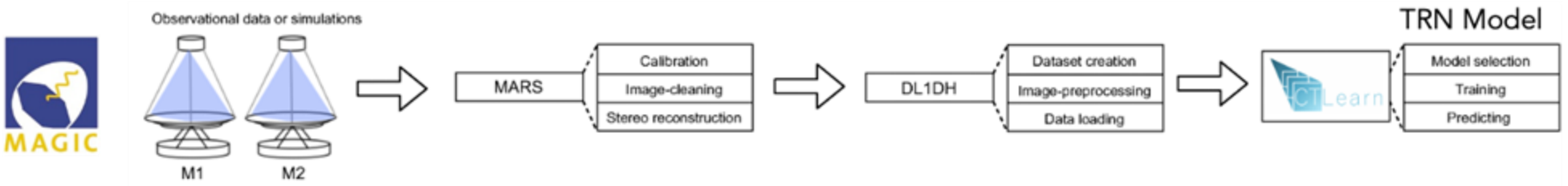
CTA Observatory

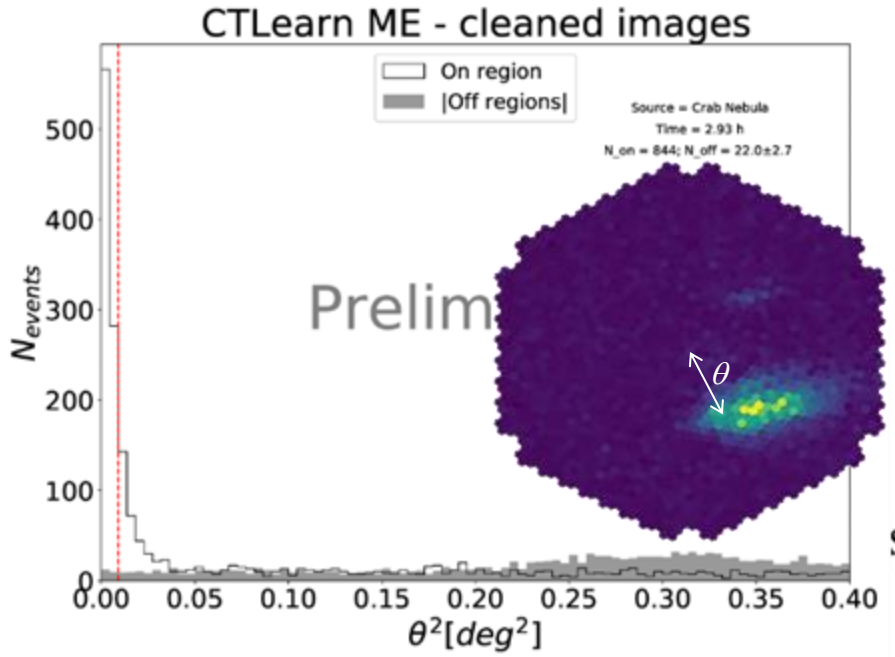
Cherenkov Telescope Array Observatory



ML@GPUs for CTAO Reco

CTLearn *on real data*





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)

Application to MAGIC real data demonstrated !!

LST1 real data coming soon

Milestones Towards The Intermediate Arrays

Can make significant achievements along the way

	Milestone	Criteria
#1	Operate LSTN-01	• Telescope accepted • CTAO Software package integrated • Illuminator, WS
#2	Open science data challenge	• 7 years of simulated observations provided as DL3 data • CTAO Software DPPS and SUSS
#3	Operate 1 SSTS + 1 MSTS	• CTAO-S Infrastructure • Computing s/w and h/w integrated • Illuminator, WS
#4	Operate 4 LSTN + 1 MSTN	
#5	Operate ≥ 5 SSTS + 2 MSTS + 2 LSTS (JAB)	

LSTs

S. McMuldroch, CTAO DG (*Nov. 2024)*

CTAO CENTRAL ORGANISATION

We Are Looking For 22 New People In 2025!

- Hiring across nearly all disciplines!
- Deputy project manager
- Project / quality assurance
- Scientists in multiple areas
- Systems engineers of all different types (general, requirements, CM, AI&V etc.)
- North and South Site staff
  (operator, engineers, safety, north site construction management)
- Software developers and coordinators (across all work packages)
- Telescope group – camera expert

S. McMuldroch, CTAO DG (*Nov. 2024)*