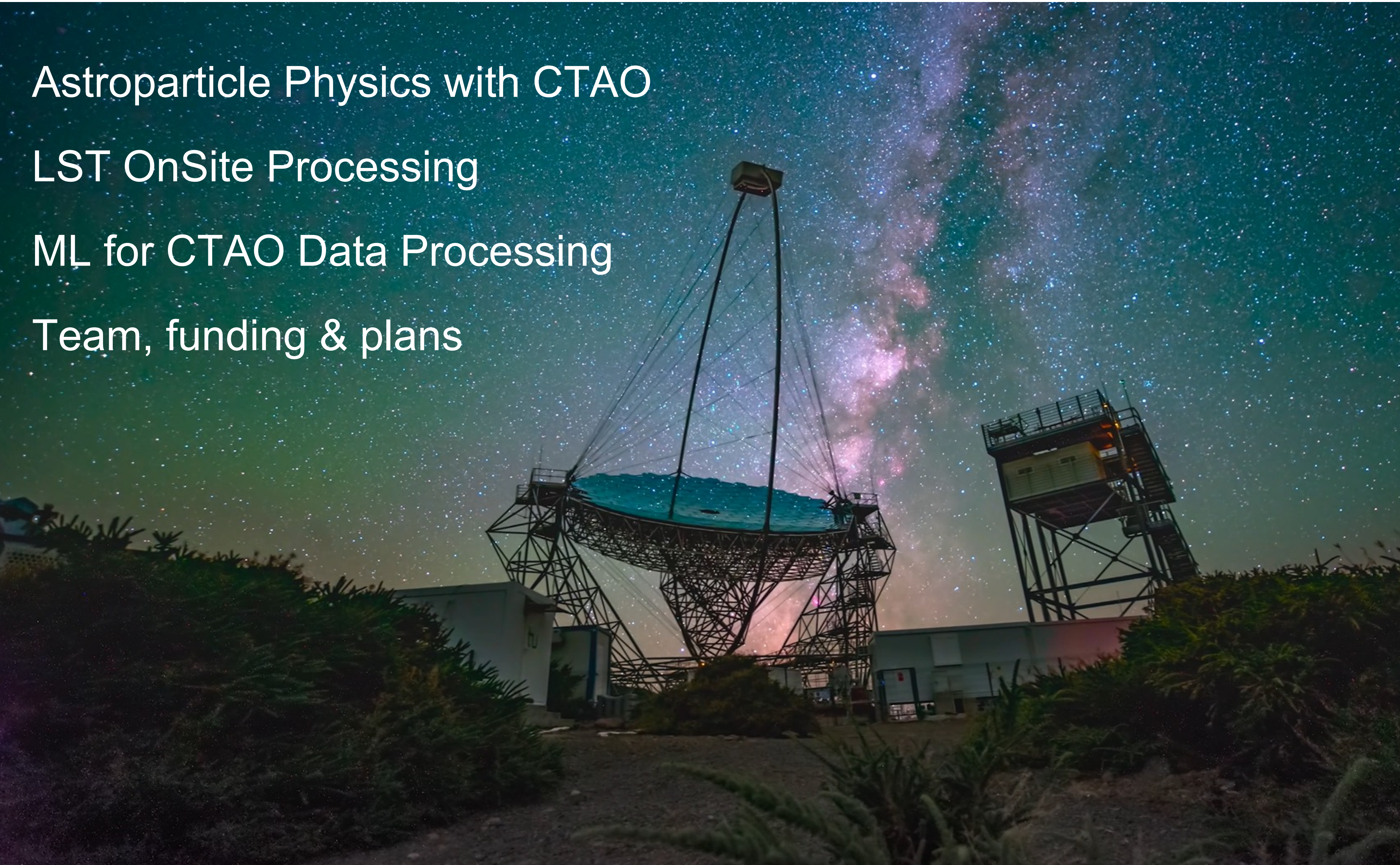
# Machine-Learning-based CTAO Telescope data processing

*XVII CPAN Days - COMCHA Session – November 2025*

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

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

# Contents



Astroparticle Physics with CTAO

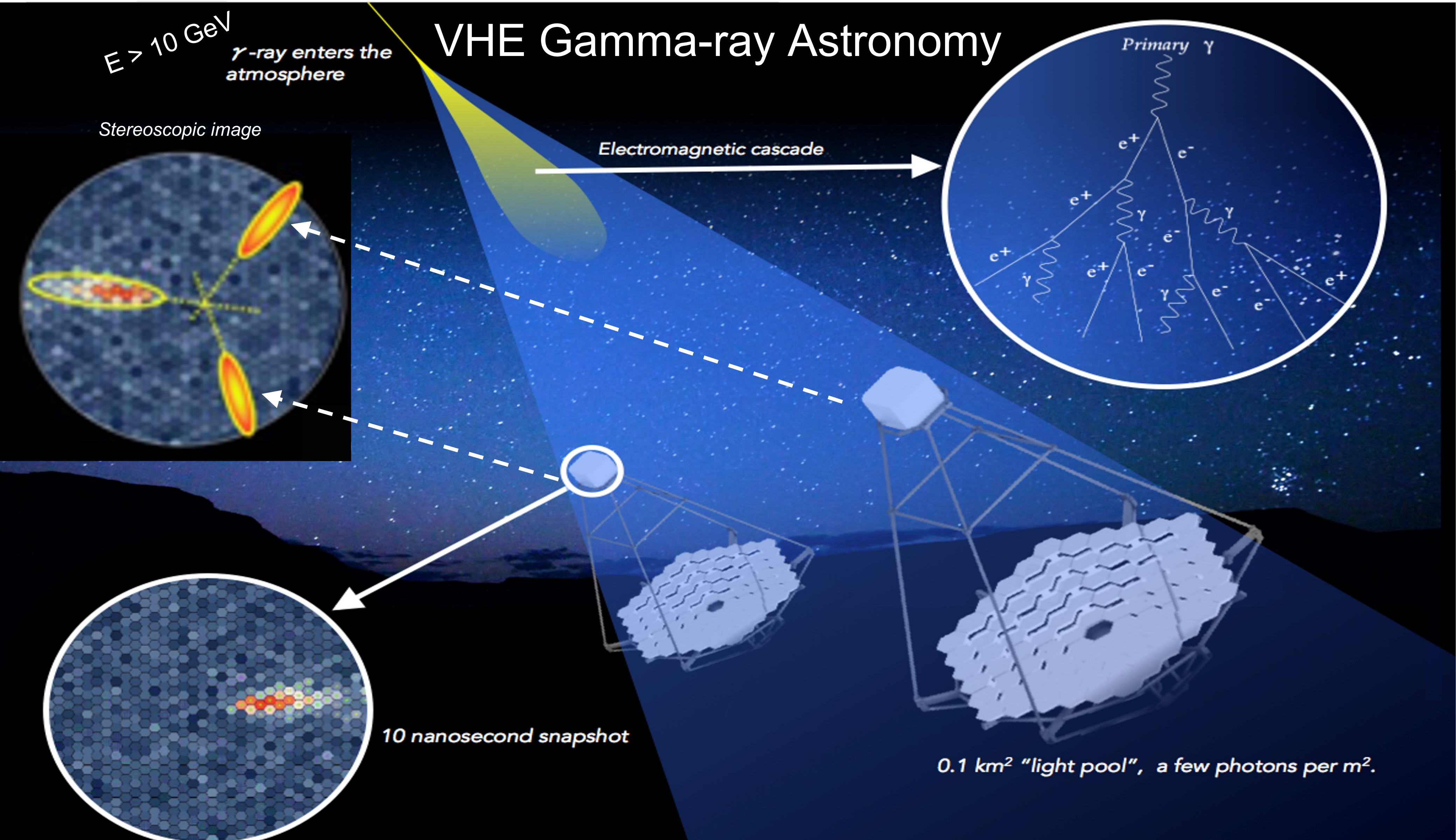
LST OnSite Processing

ML for CTAO Data Processing

Team, funding & plans

# Astroparticle Physics

## VHE Gamma-ray Astronomy




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

# CTA Observatory

## Cherenkov Telescope Array Observatory

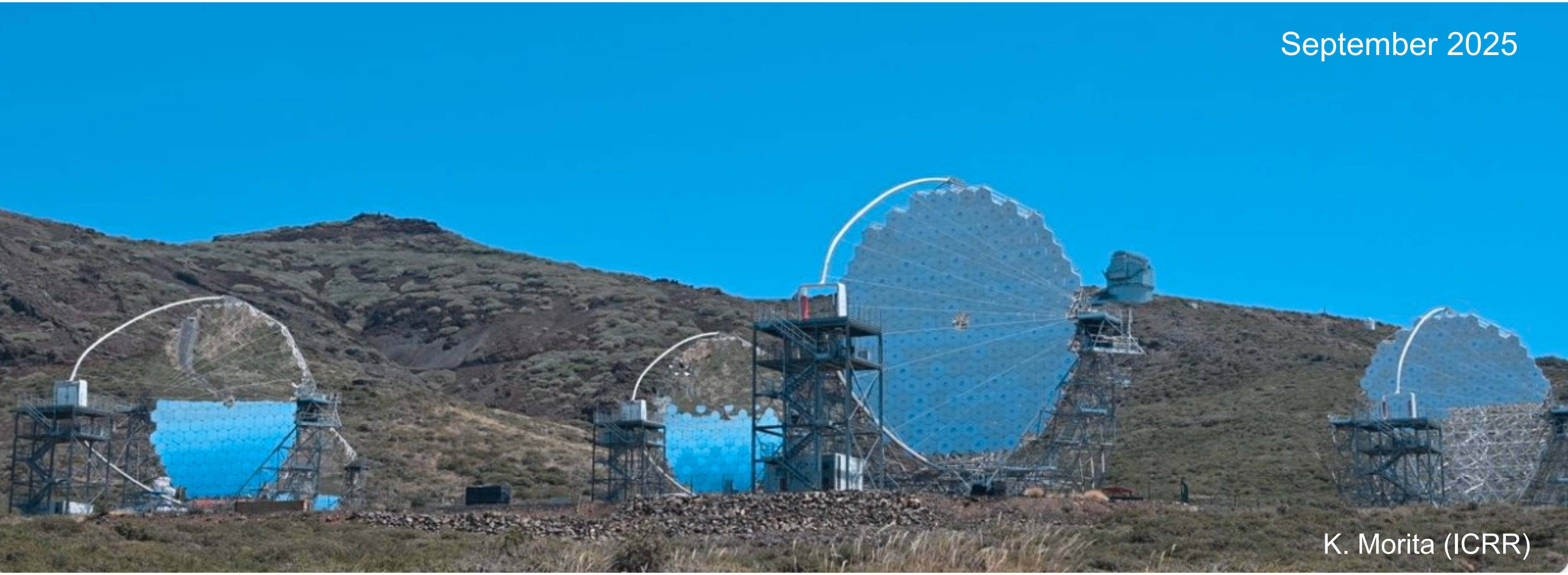

Sensitivity improvement x10
Energy range extension x10
Angular resolution improvement

→

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

# CTA Observatory

## CTAO-North

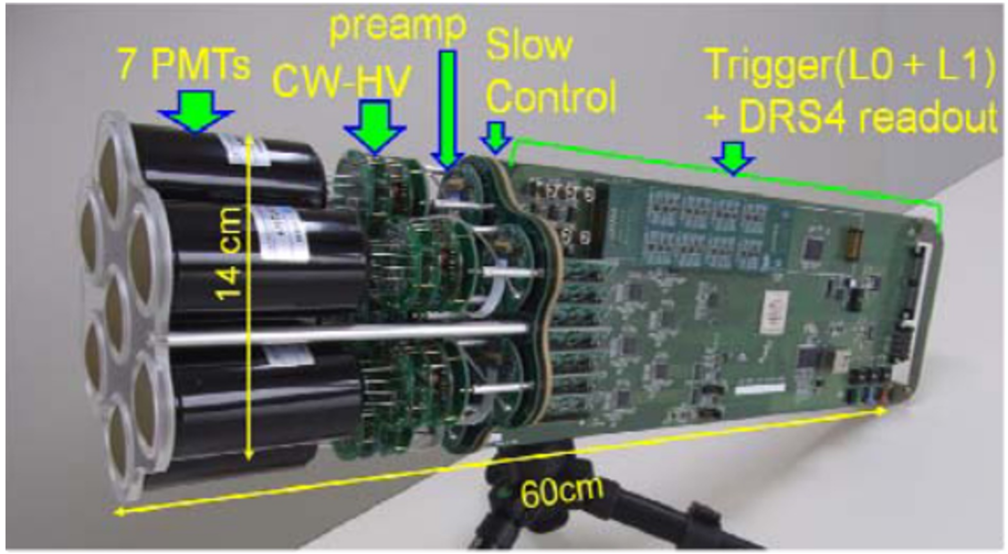



Sensitivity improvement x10
Energy range extension x10
Angular resolution improvement

→

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

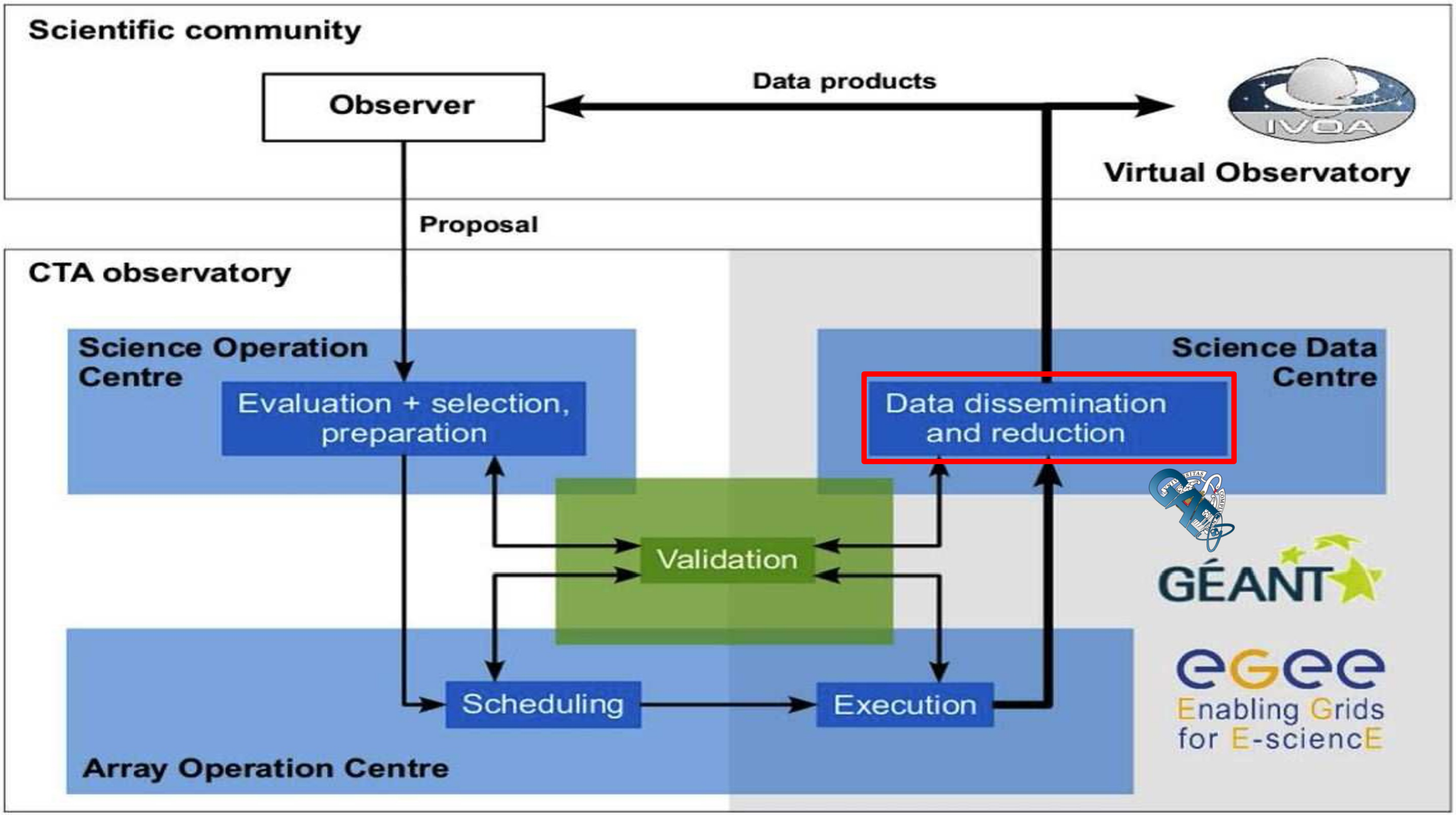
## Cherenkov Telescope Array Observatory




~2000-pixel PMT-based camera

# CTA Observatory

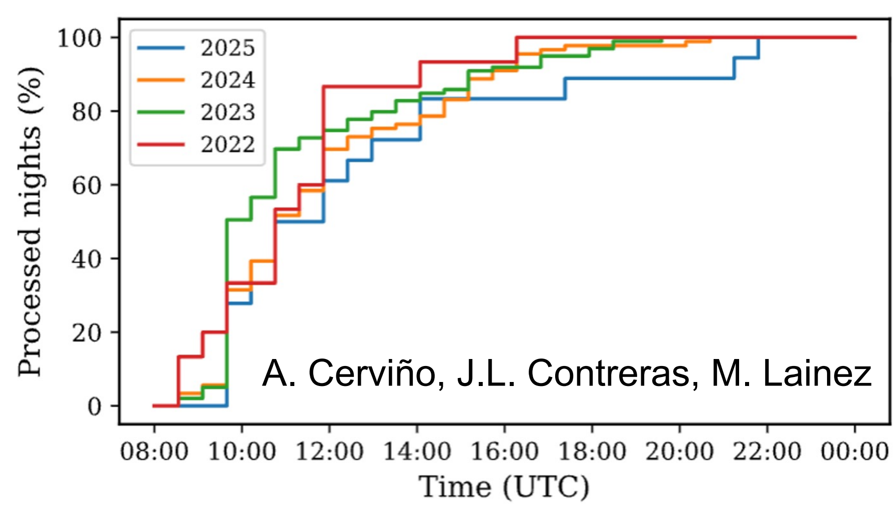
## Cherenkov Telescope Array Observatory



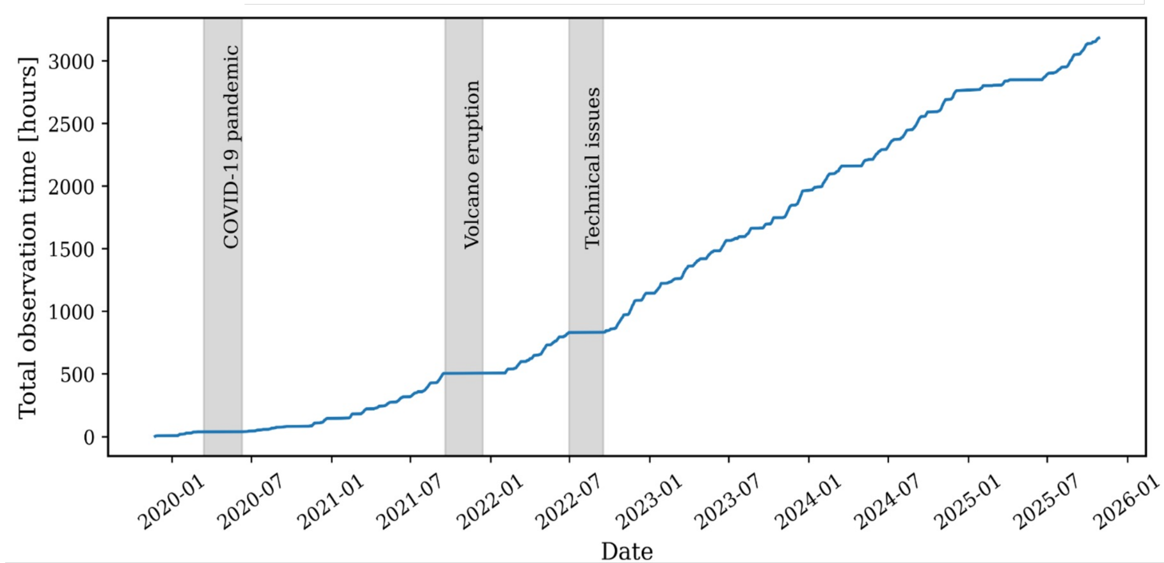
# Data Processing @CPUs

## 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 ~2026
  - Data is processed on-site in a temporary data center at the telescope base with ~1800 cores and 5.7 PB HD
  - Onsite pipeline, processing raw data every morning in a highly parallel way
  - Until recently saving all raw data




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



# ML@GPUs for CTAO Reco

## CTLearn

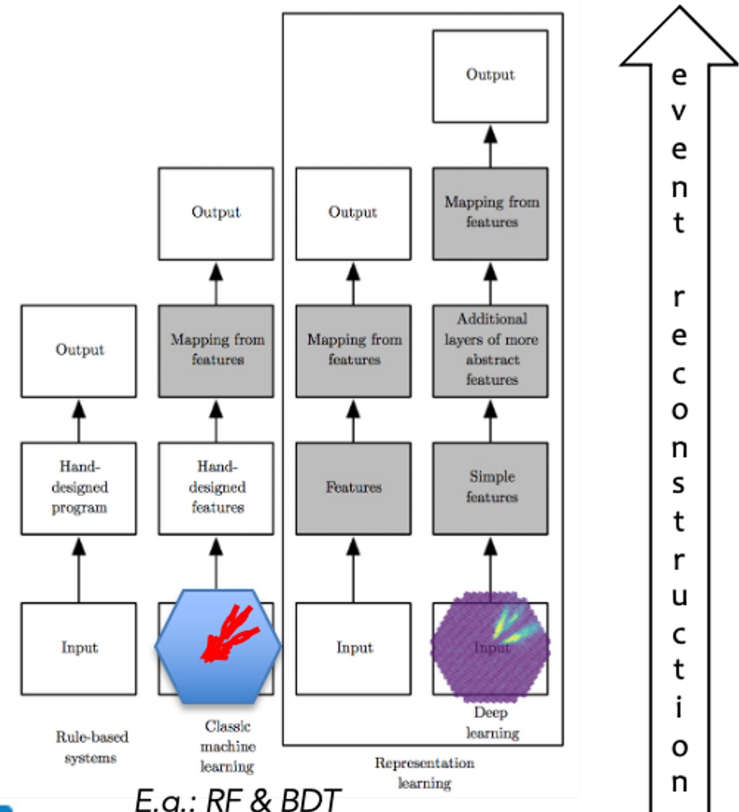
J. Buces, A. Cerviño, D. Martín, D. Nieto

- CTLearn is a high-level Python package for using **Deep Learning models** aiming for:
  - **IACT data analyses**
  - **CTAO Offline Data Volume Reduction**
- Core functionality:
  - Full-event reconstruction of various IACTs in monoscopic and stereoscopic mode
  - **CNN-based analysis on raw waveforms** possible through the efficiently data management package dl1-data-handler
  - Application of an **AI-based Trigger system**, where neural networks are ported on **FPGAs** for real time processing.
- Latest release: v0.10.2 (21/03/2025)
- Local computing resources + Artemisa

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

DOI 10.5281/zenodo.3342952

Output: event type, energy, incoming direction

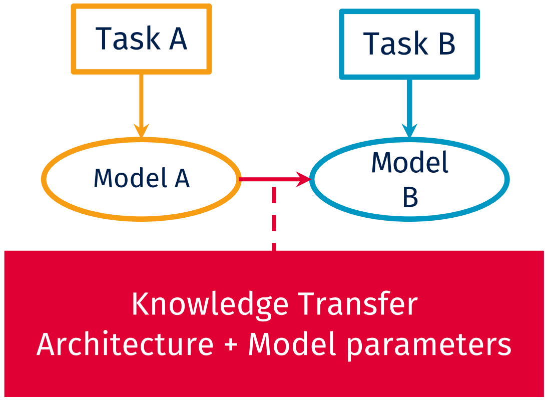



Input: observed events

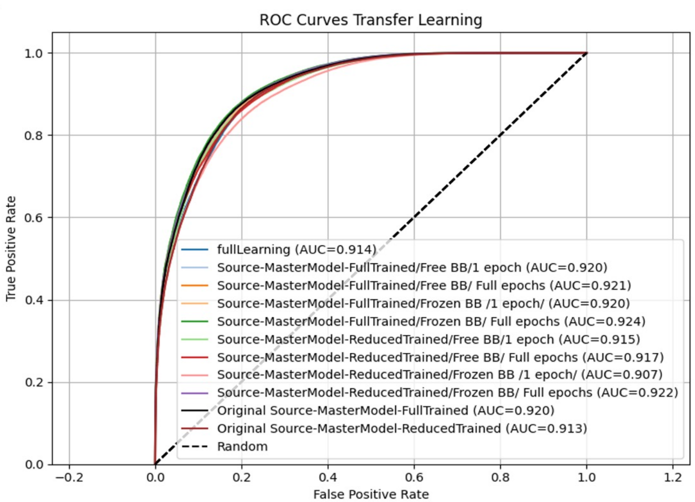
# ML@GPUs for CTAO Reco

## CTLearn - Optimization techniques - Preliminary results

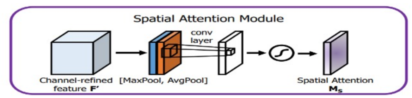
### Transfer Learning



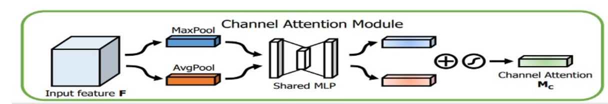

- Save up to **75% of training time**
- Good metrics with less resources



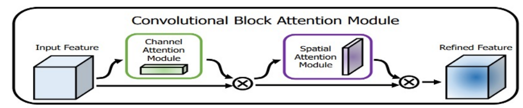

### Attention Experiments



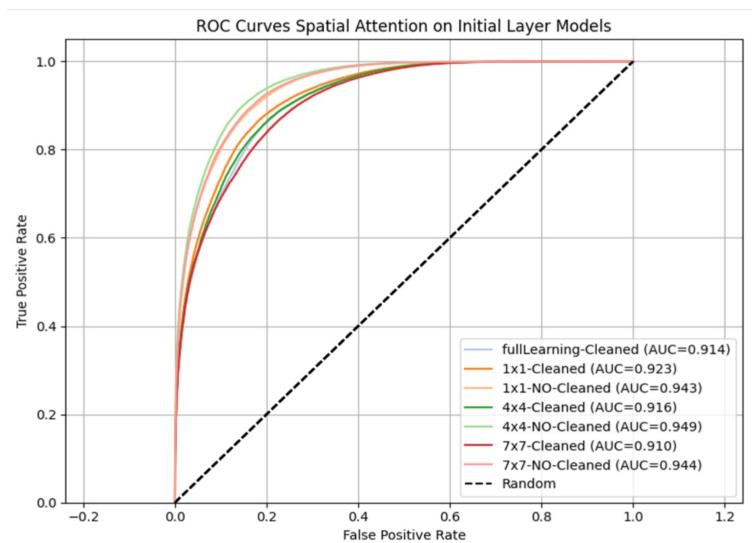

*Space*




*Time*




- Better understanding of the **CNN Explainability**
- **Cleaning step may be omitted**

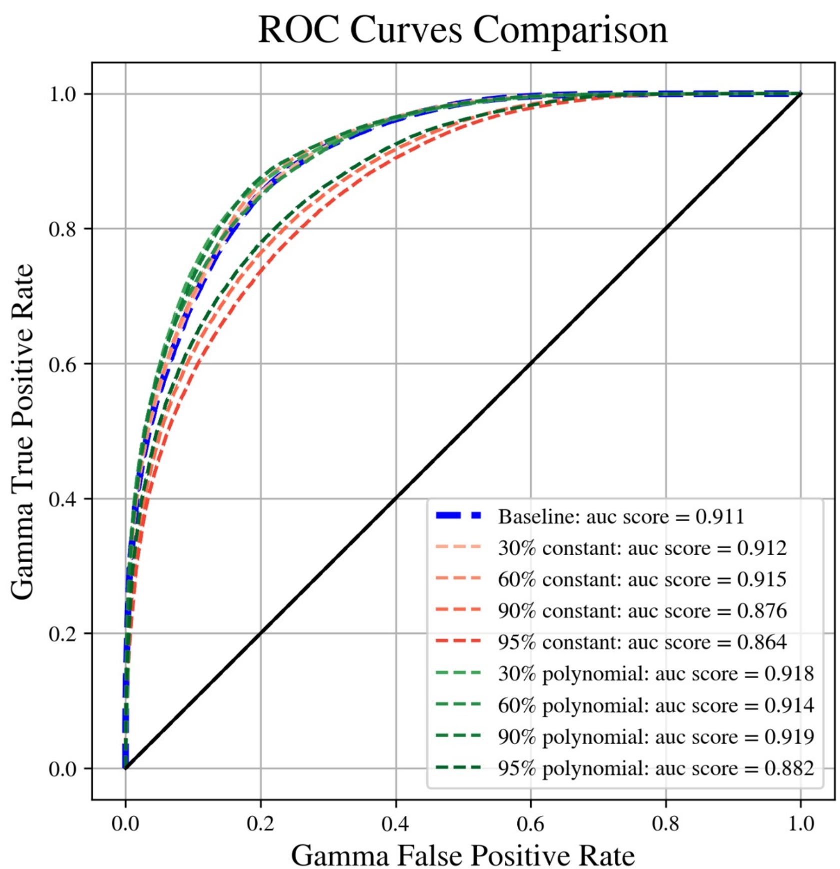


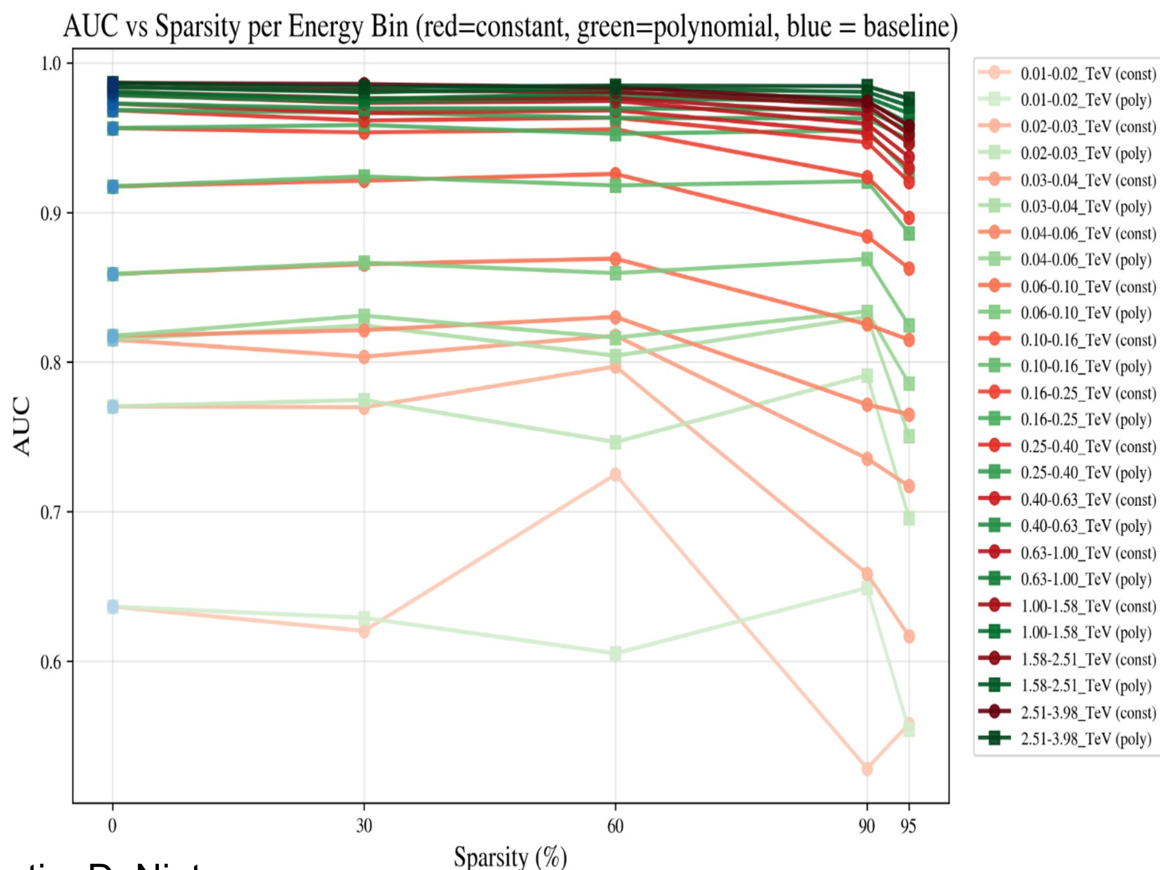

A. Cerviño, D. Nieto

# ML@GPUs for CTAO Reco

## ML algorithm compression

- Prunning algos on Reco data → later to be used for camera trigger
- Polynomial prunning reducing 90% parameters maintaining performance



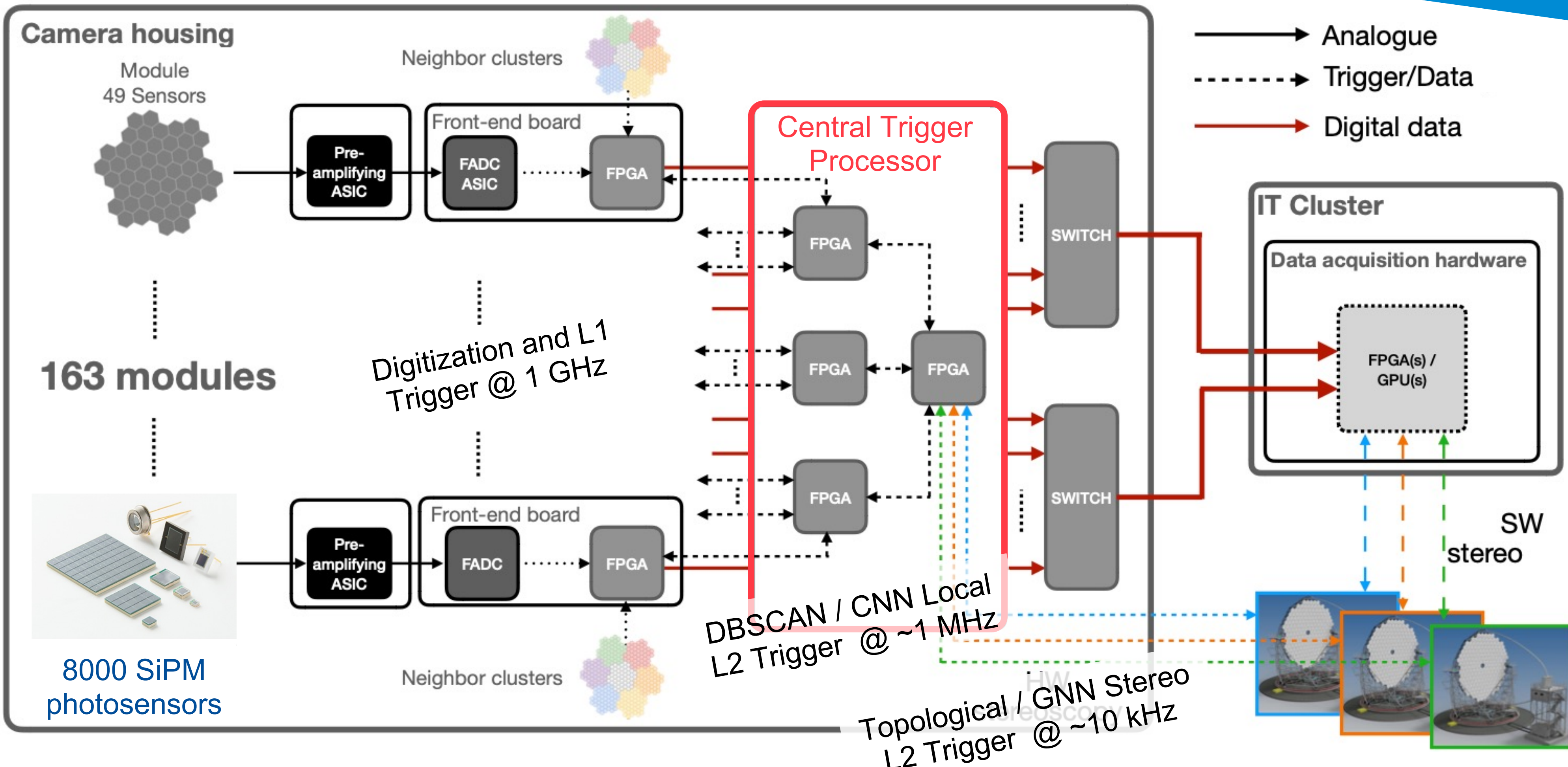




D. Martin, D. Nieto

# ML@FPGAs for CTAO Trigger

## CTAO-LST SiPM Advanced Camera*

Candidate for mid-term upgrade of CTAO telescope cameras

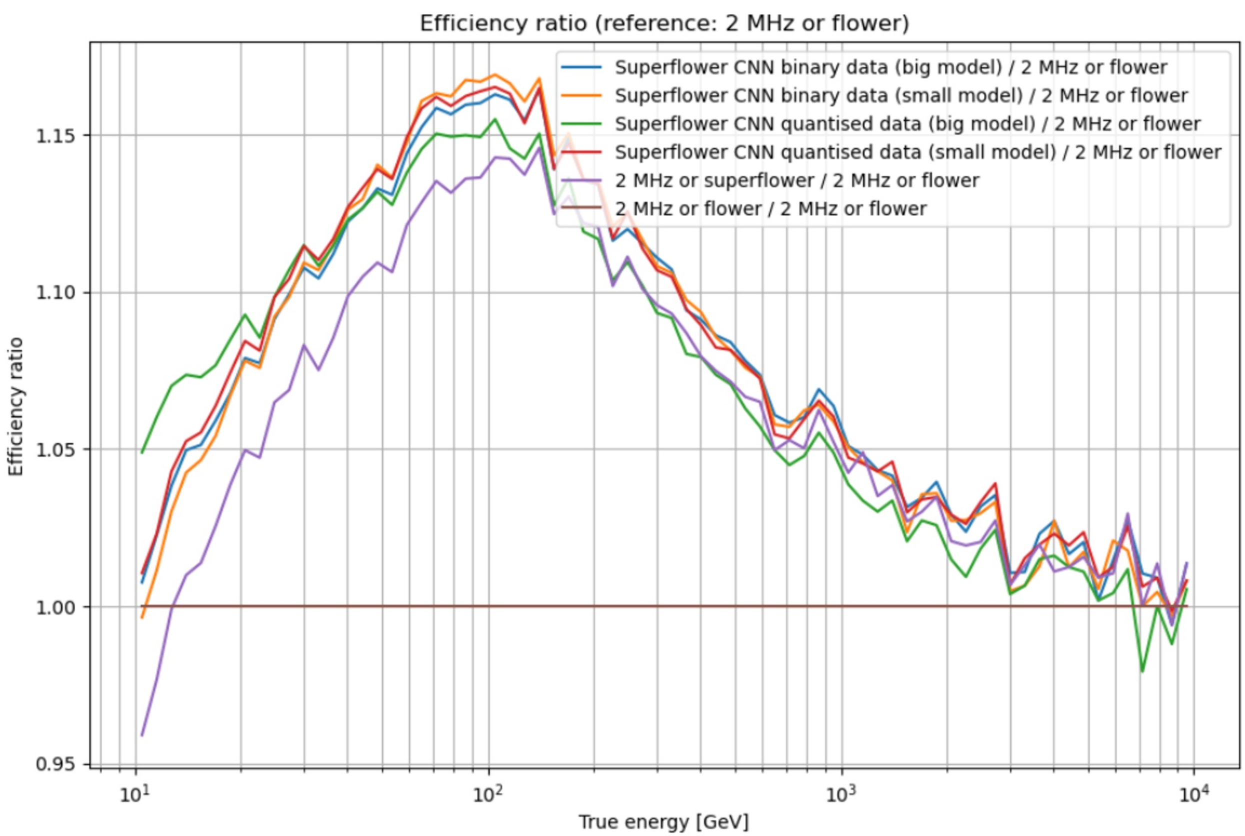


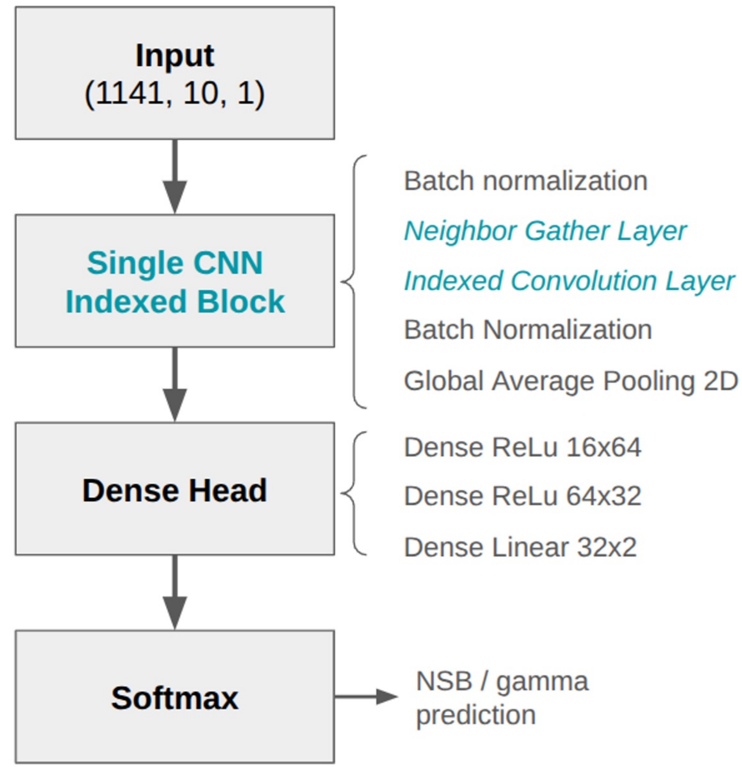

M. Heller, UniGe, 2025

*PoS(ICRC2025)673

# ML algorithms for CTAO camera LL2 trigger

- Very light models with custom layers with ~ 3k parameters
- Moderate gain in gamma-ray efficiency

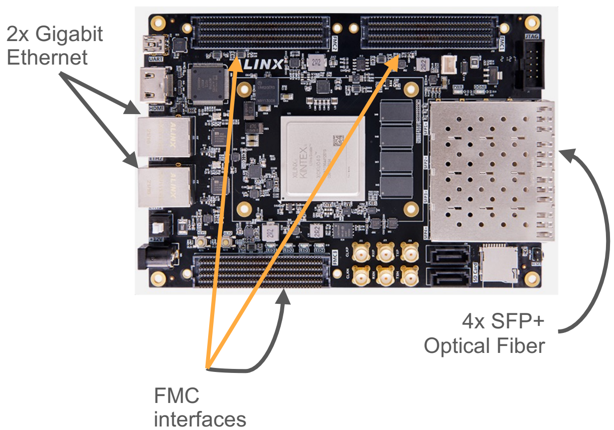



J. Buces, D. Nieto, J.A. Barrio

## CTP test benches

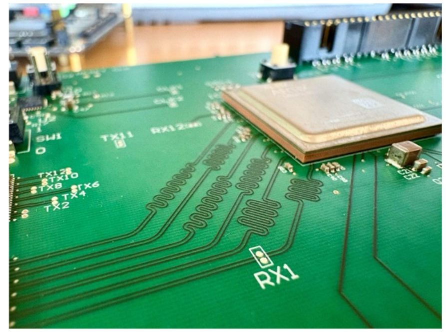
### #1 Machine Learning @ FPGAs

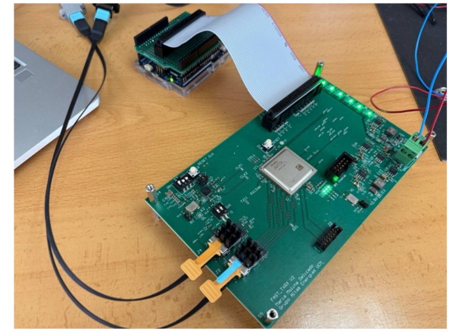



- 2x ALINX AMD Xilinx Kintex UltraScale XCKU040
  - 20 gigabit transceivers @ 16.3 Gbps
  - 4GB high-speed DDR4 RAM
- Data transfer between PC and FPGA using IPBus protocol

M. Molina, A. Pérez-Aguilera, L.A. Tejedor, J.A. Barrio

### #2 High-speed lines



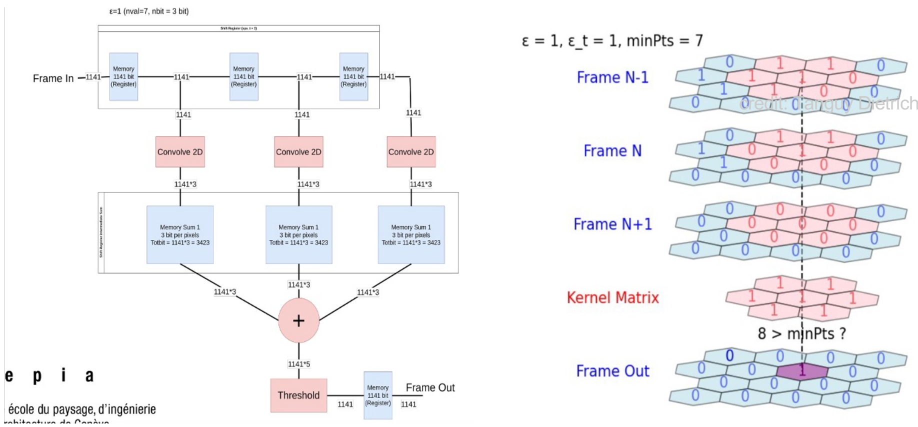



- Main components
  - Xilinx UltraScale+ with 12 gigabit transceivers
  - Two 12-channel Samtec FireFly 14 Gbps optical connectors (TX and RX)
- Design and manufacturing
  - 12 layer PCB design and high-speed differential pair routing
  - Validation of the substrate and the PCB manufacturer
- Latest results
  - Slow control firmware for configuration and monitoring of FireFly modules on real time
  - Successful signal integrity testing
  - Excellent performance up to 10 Gbps

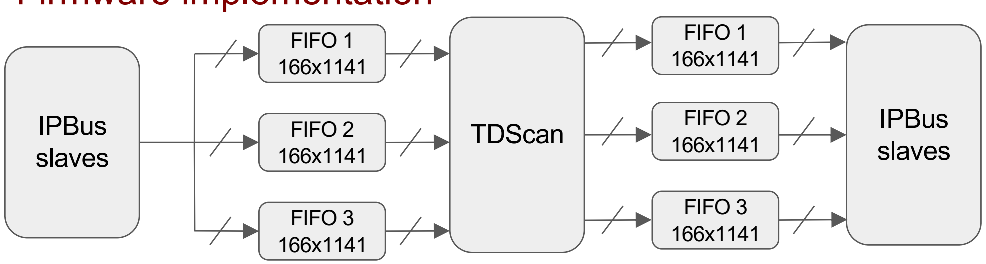
# ML@FPGAs for CTAO Trigger

## CTP test bench #1: Machine Learning algorithms

### **TDSCAN:** DBSCAN-like parallel 2+1D conv. over the whole camera (HESGE-HEPIA)



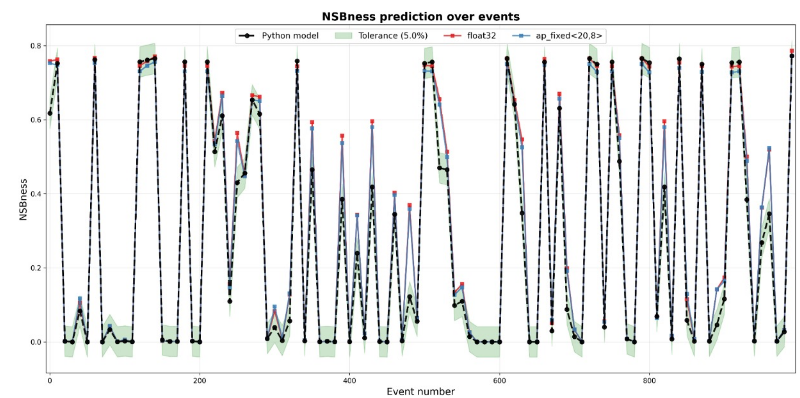

### Firmware implementation




- Latency tests demonstrated proper operation @ 1 GHz (processing 1141 bit frame per ns)
- Data check tests confirmed same results as the Python script

### **Custom CNNs @ FPGAs using Vitis**

- Reconstruction of the original Python model on C++
- Data quantization (PTQ) from float32 to fixed point




- HLS optimization
- Post-synthesis results: ~7μs latency/interval on XCKU040 and ~3.5μs on XCKU115 (bottleneck on 1141 cycles)
- Next steps: overcome the bottleneck with different models with smaller input dimensions

# CTP test bench #2: Signal integrity (BER) testing

No failed bits detected during the tests

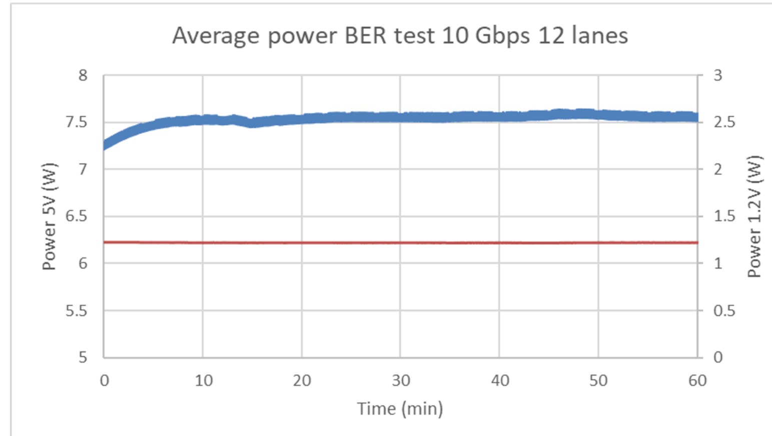
- BER @ 6 Gbps < $4.394 \cdot 10^{-14}$
- BER @ 10 Gbps < $2.63 \cdot 10^{-14}$

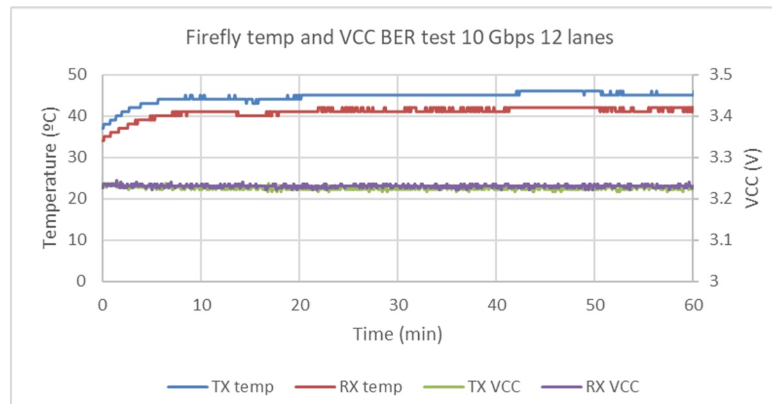
Testing conditions: raw signal (no encoding) and approx one hour duration with continuous monitoring of FireFly's temperature and power supplying

Unexpected results @ 2 Gbps

- Channels #6 and #11 BER ~$4.4 \cdot 10^{-11}$
- Remaining channels BER < $1.35 \cdot 10^{-13}$ (0 errors)
- Resonance or impedance matching issues
- Further investigation is required

Next steps → Repeat tests with candidate protocols for the final CTPB (e.g., Aurora, JESD204C) and characterize their latency, throughput and other relevant metrics



# Team, funding & plans

- Team
  - ML@FPGAs: 2 faculty (phys + h/w eng), 4 predoc (2 phys + 2 h/w eng)
  - ML@GPUs: 1 faculty (phys), 1 predoc (phys & s/w eng)
  - OnSite Processing: 1 faculty (phys), 1 predoc (phys)
- Network
  - Spain: CNID/COMCHA (ML@xx, OnSite Proc.); Ciemat/IFIC (ML@FPGAs)
  - International: AdvCam (Ciemat, UniGe, INFN-Padova); CERN DRD7 (ML@FPGAs)
- Dedicated grants
  - Running: Spanish (PDC2023+PPCC) 2-year for predocs & h/w
  - Requested: Spanish 2-year for predocs (PDC2025 *call*), Spanish 4-year for predocs & h/w (*CDTI call for Fire Detection on-board Satellites*, incl. CIEMAT*)*, *EU-InfraTECH-2026 4-year for predocs & h/w*
- Plans for AdvCam Trigger
  - ***2026***: complete 2-testbench demonstrator, deploy & benchmark simple CNNs, compressed CNNs & DBSCAN-like for LL2, test GNNs for SL2
  - ***2028***: build 1/4-scale CTP prototype, deploy & benchmark optimized AI-based trigger

# Summary

- CTAO ESFRI construction started
- PDC2023 + PPCC → involvement in ML-based R&D for CTA
- Sinergies with COMCHA teams to be pursued
- Transfer of knowledge pursued/expected from ML@FPGA activities

# Acknowledgements

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

Financiado por la Unión Europea NextGenerationEU

Plan de Recuperación, Transformación y Resiliencia

AGENCIA ESTATAL DE INVESTIGACIÓN

The research here presented has been partially supported by the MICIU/AEI/10.13039/501100011033 and by the EU-NextGenEU/PRTR under grant PDC2023-145839-I00, and ERDF/EU under grant PID2022-138172NB-C42

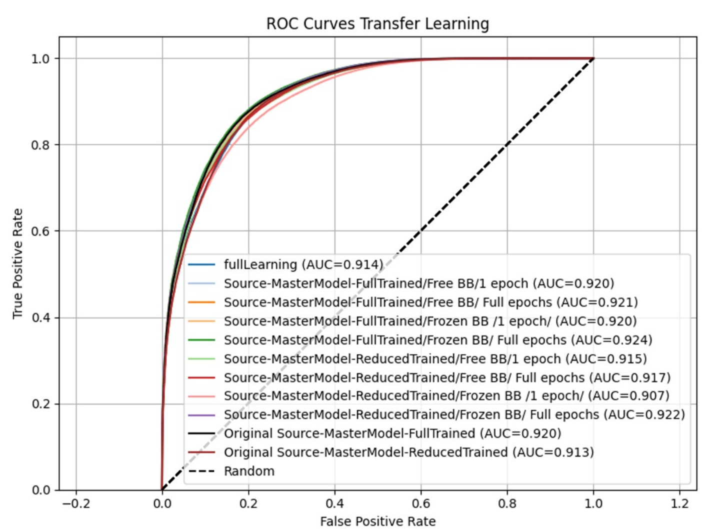
IPARCOS

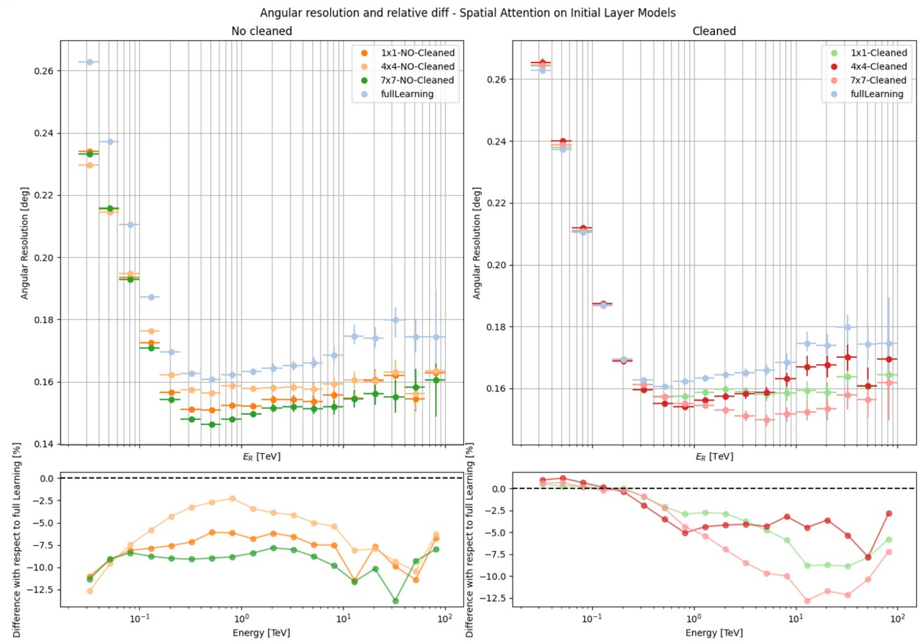
UNIVERSIDAD COMPLUTENSE MADRID

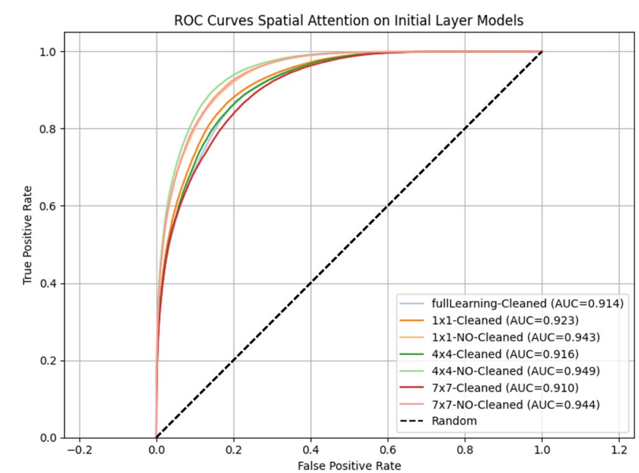
# Backup

# CTLearn - Optimization techniques - Preliminary results

| Time (relative to full learning) | Energy | Direction | Type |
|---|---|---|---|
| Full Learning | 1 | 1 | 1 |
| Full Trained - Free - 1epoch | 1.46 | 0.55 | 1.35 |
| Full Trained - Free - full epochs | 1.45 | 0.54 | 1.27 |
| Full Trained - Frozen - 1epoch | **0.45** | **0.31** | **0.44** |
| Full Trained - Frozen - full epochs | **0.49** | - | **0.45** |
| Reduced Trained - Free - 1epoch | 1.87 | 1.72 | 1.55 |
| Reduced Trained - Free - full epochs | 1.85 | 1.81 | 1.16 |
| Reduced Trained - Frozen - 1epoch | **0.21** | **0.31** | **0.27** |
| Reduced Trained - Frozen - full epochs | **0.30** | **0.5** | **0.19** |

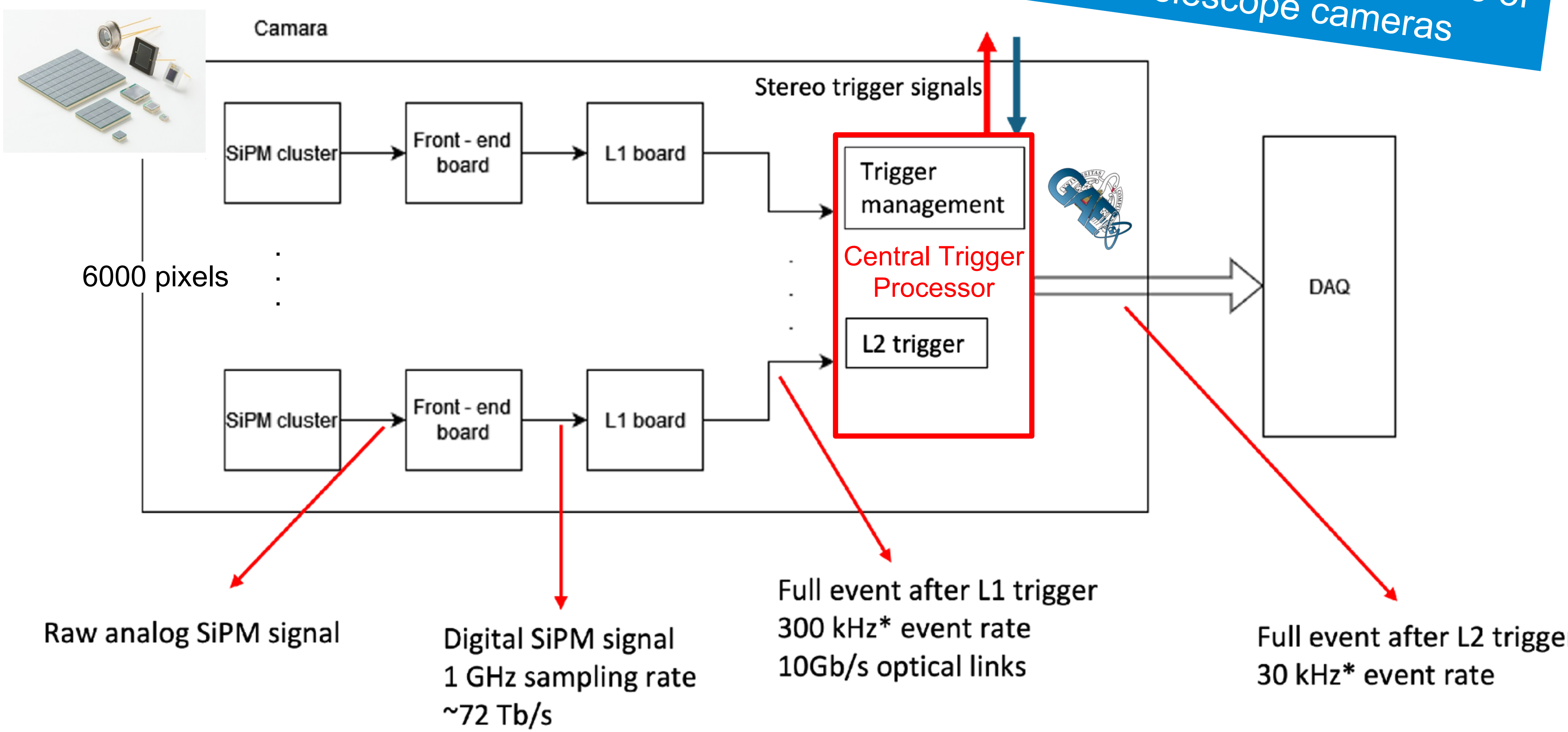


## Advanced LST SiPM Camera*

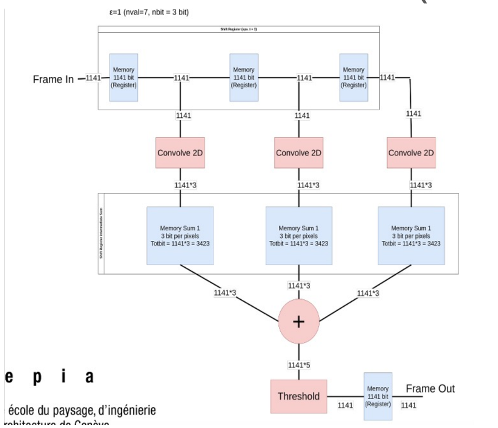
Candidate for mid-term upgrade of CTAO telescope cameras

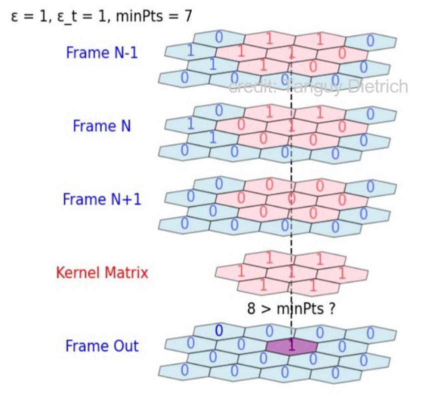



*M. Heller et al. PoS(ICRC2023)740

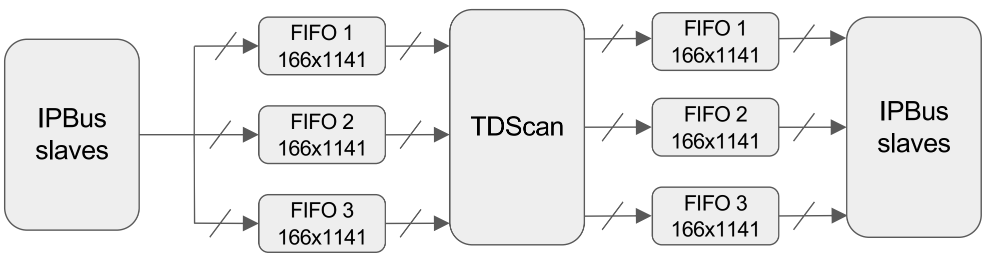
# CTP test bench #1: Machine Learning algorithms

**TDSCAN** → parallel 2+1D convolution over the whole camera (HESGE-HEPIA)




## Firmware implementation

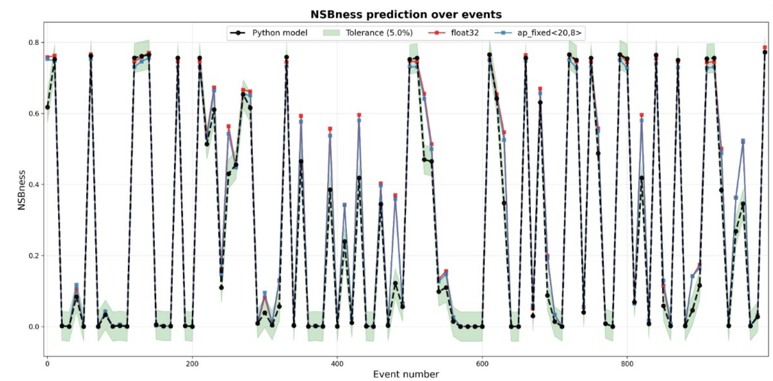



- Latency tests demonstrated proper operation @ 1 GHz (processing 1141 bit frame per ns)
- Data check tests confirmed same results as the Python script

## Custom CNNs @ FPGAs using Vitis

Reconstruction of the original Python model on C++
- Weights and biases extraction into .h files
- Pre-computation of batch normalization layers
- Development of core functions and wrappers on C++

Data quantization (PTQ) from float32 to fixed point
- Profiling of activations and accumulators range
- Comparative analysis of precision between models




HLS optimization: Loop and function pipelining → Loop unrolling and parallelization → Dataflow and array partition→ Interface optimization and data packing

Post-synthesis results: ~7μs latency/interval on XCKU040 and ~3.5μs on XCKU115 (bottleneck on 1141 cycles)

Next steps: overcome the bottleneck with different models with smaller input dimensions