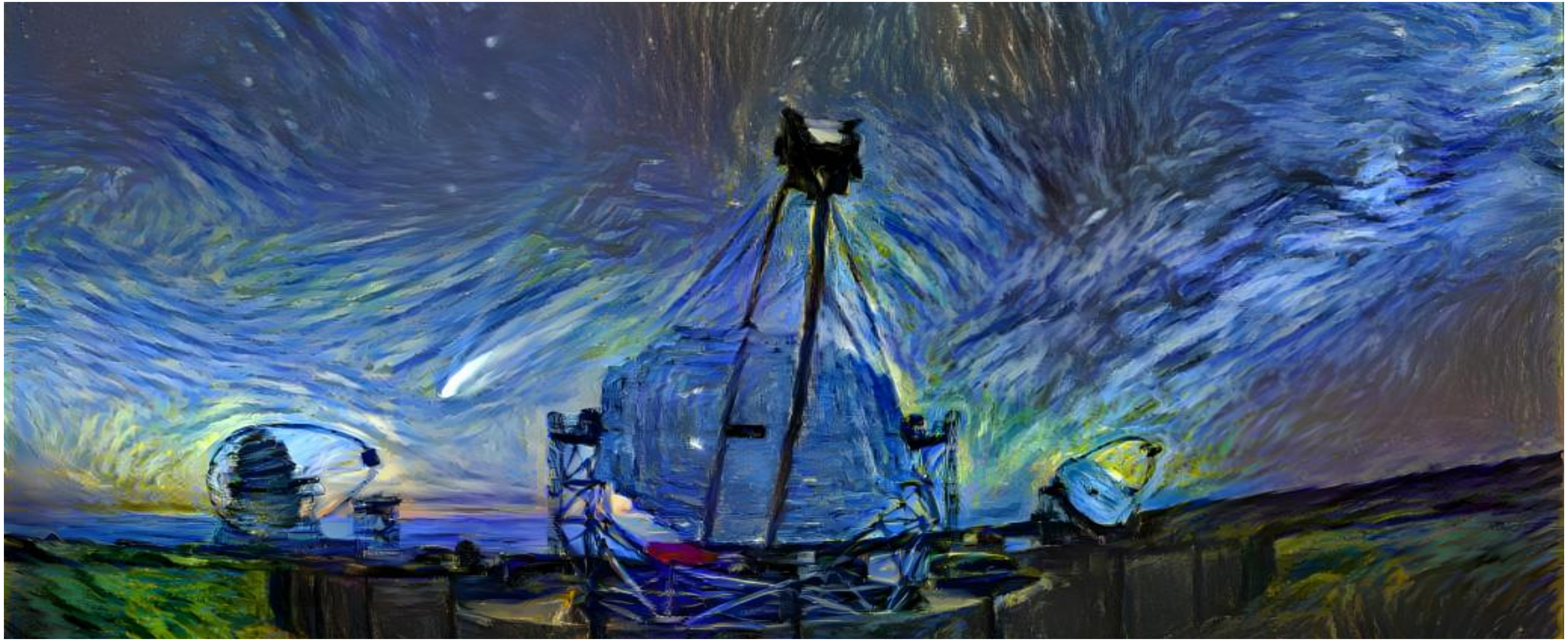
MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Astrofísica y física de altas energías
GAE
IPARCOS

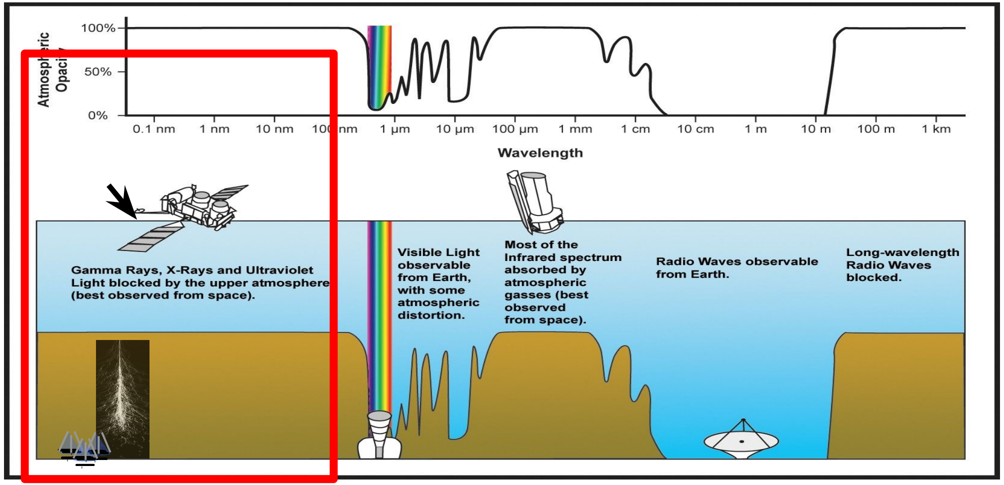
Deep-learning techniques in ground-based imaging gamma-ray observatories and the CTLearn package

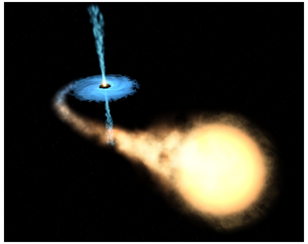
High Energies Group

Alexander Cerviño
rcervino@ucm.es
Supervisors: J.L. Contreras & D. Nieto



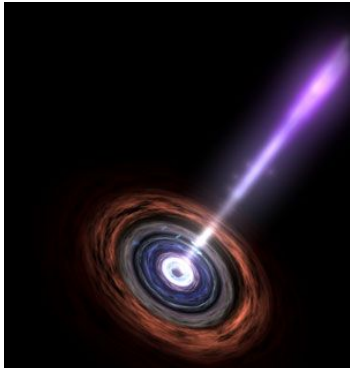
INTRODUCTION



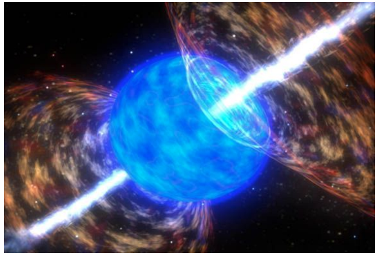


Gamma-ray Binaries

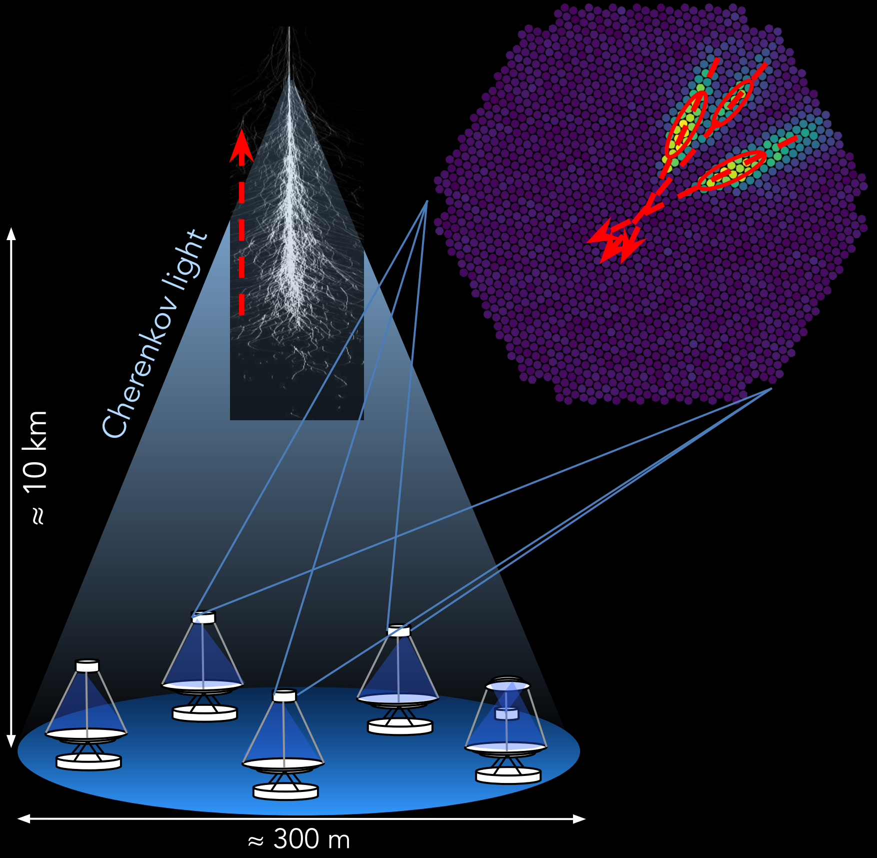
Supernova



Active Galactic Nuclei



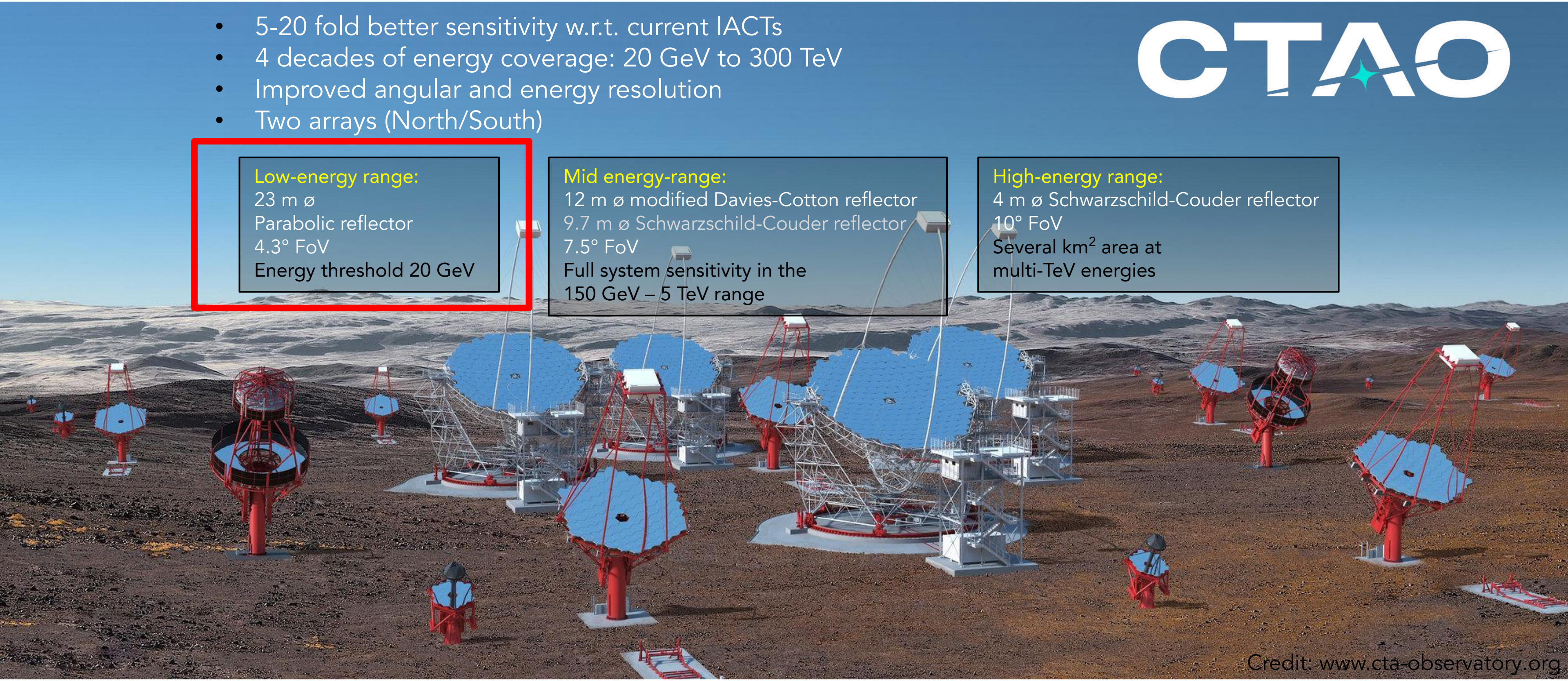
Gamma-ray Bursts



- Imaging Atmospheric Cherenkov Telescope (IACT)
- Detection of extended air showers using the atmosphere as a calorimeter
- Huge γ–ray collection area $(\sim 10^5\ m^2)$
- 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

Next-generation IACT: The Cherenkov Telescope Array

- 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/South)



Low-energy range:
23 m ø
Parabolic reflector
4.3° FoV
Energy threshold 20 GeV

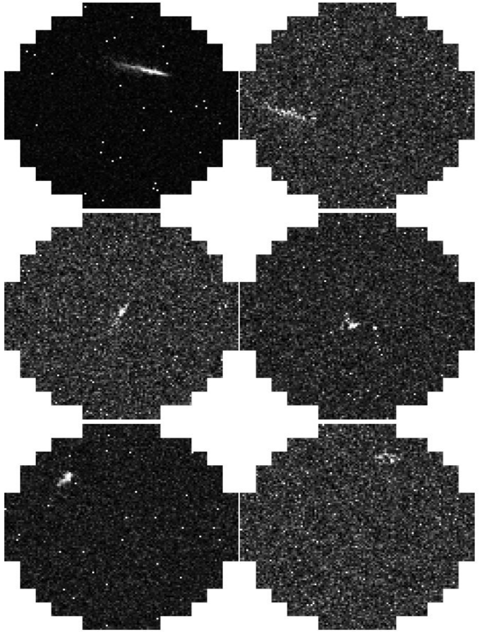
Mid energy-range:
12 m ø modified Davies-Cotton reflector
9.7 m ø Schwarzschild-Couder reflector
7.5° FoV
Full system sensitivity in the 150 GeV – 5 TeV range

High-energy range:
4 m ø Schwarzschild-Couder reflector
10° FoV
Several km^2 area at multi-TeV energies

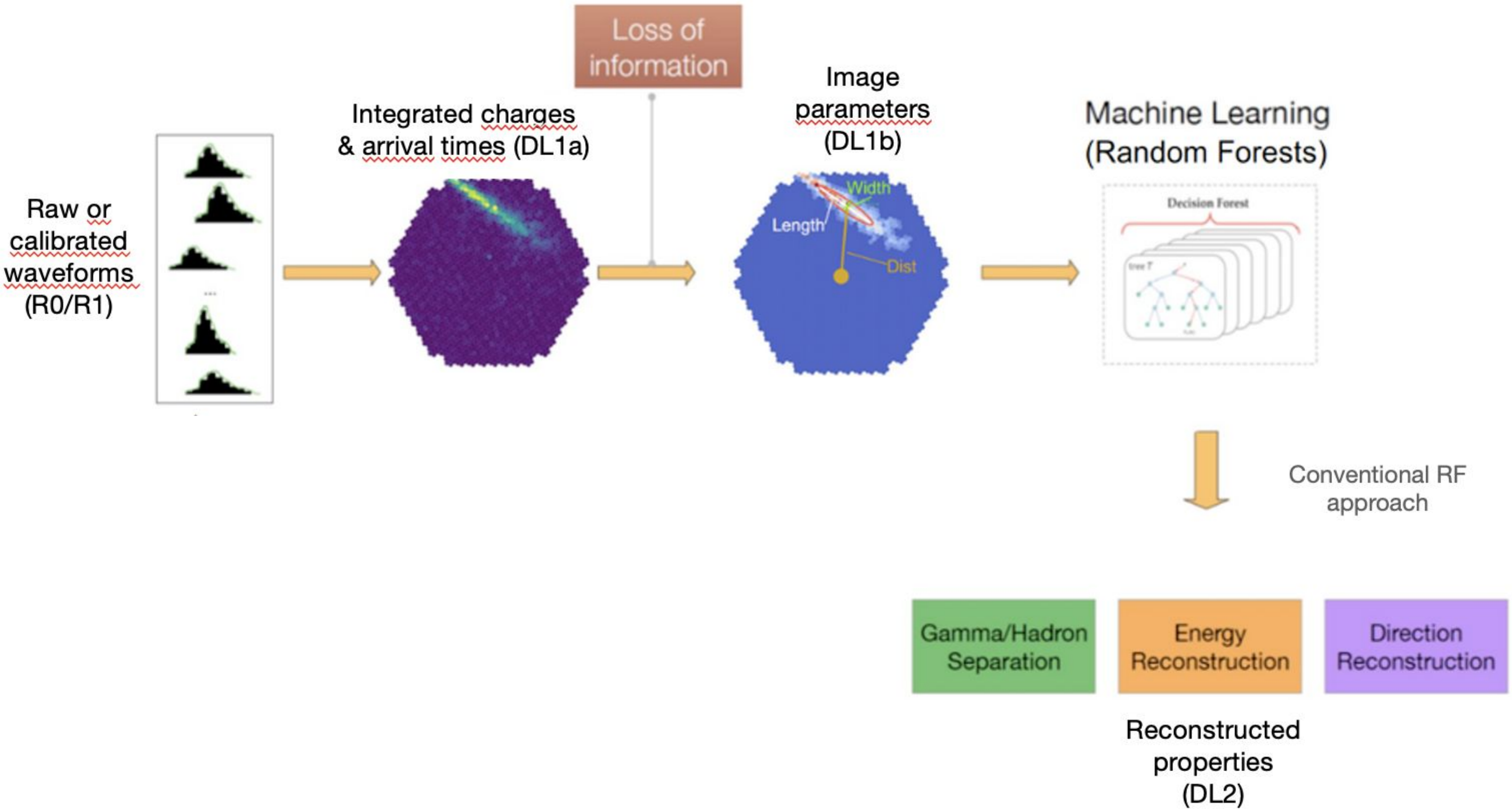
www.ctao.org

Science with CTA, arXiv:1709.07997

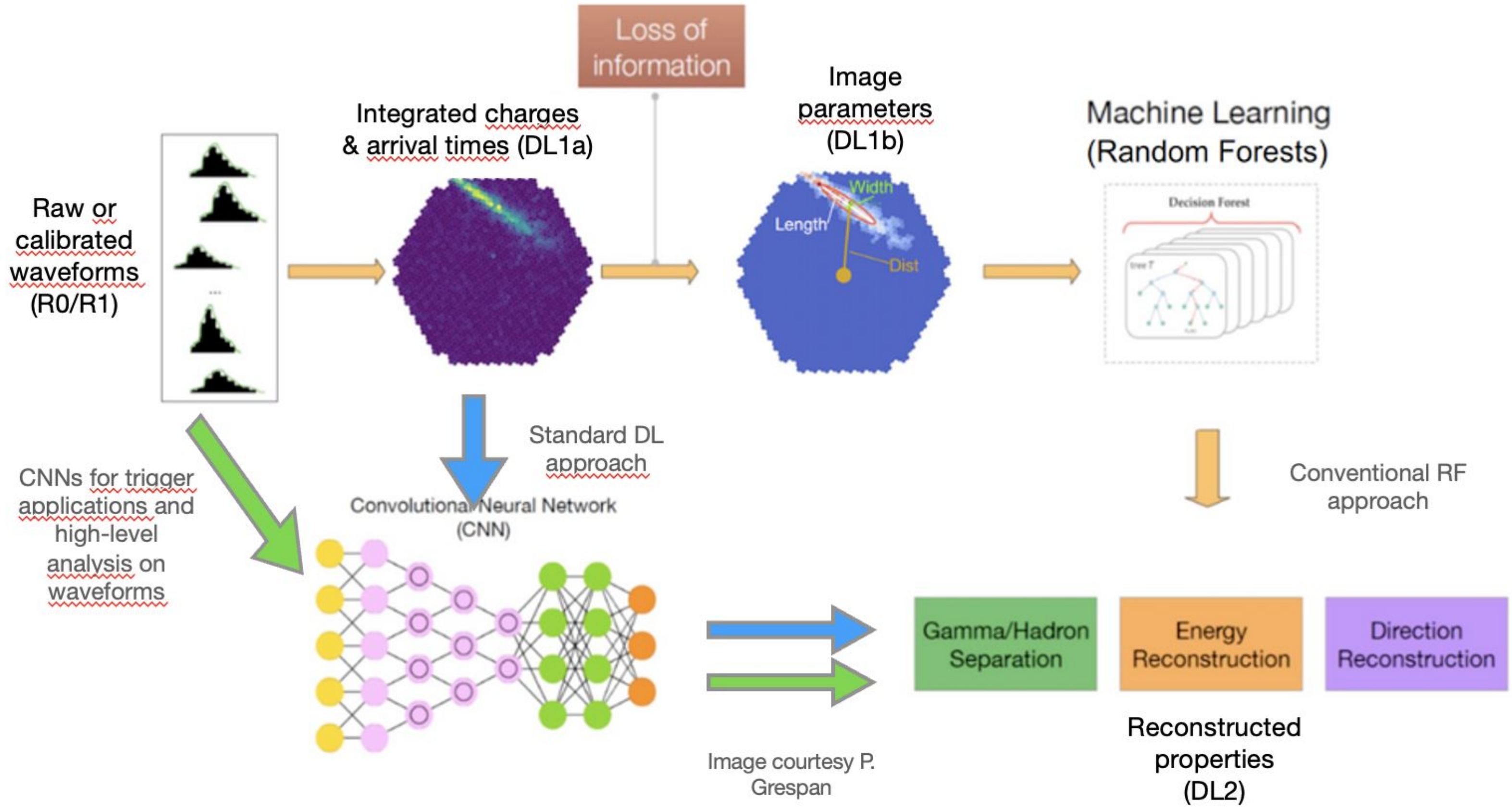
DEEP LEARNING APPLIED TO IACTs



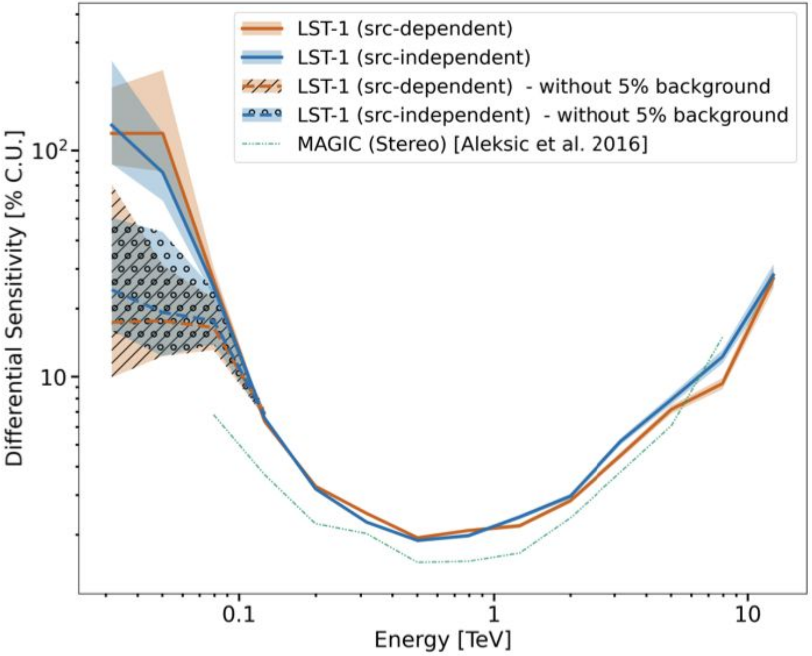
Event reconstruction in IACTs – Current state

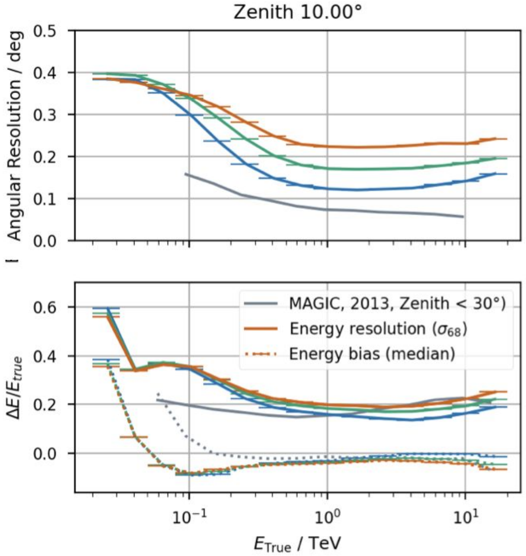


Event reconstruction in IACTs – DL Scenario



- Improvement of current reconstruction and performance metrics with Random Forest





LST Collaboration (2024)

- CTLearn is a high-level Python package for using **Deep Learning models for ground-based gamma-ray data analyses.**
- Core functionality:
  - **Full-event reconstruction** of various imaging atmospheric Cherenkov telescopes in **monoscopic and stereoscopic mode**
  - **CNN-based analysis on 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)

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

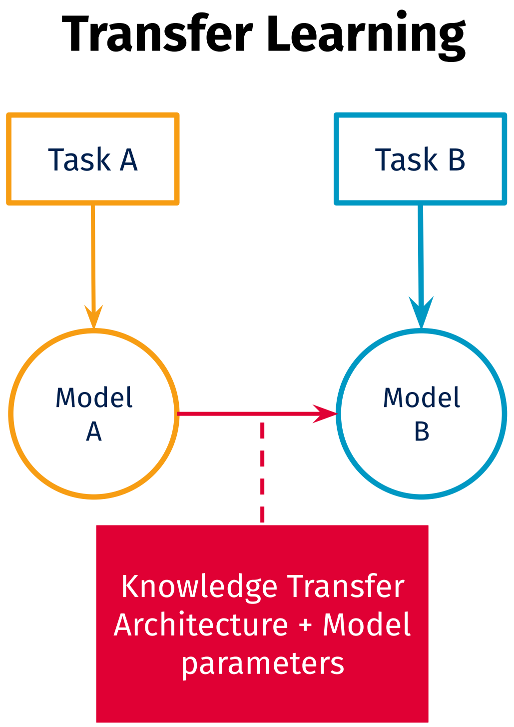
DOI 10.5281/zenodo.3342952

CTLearn

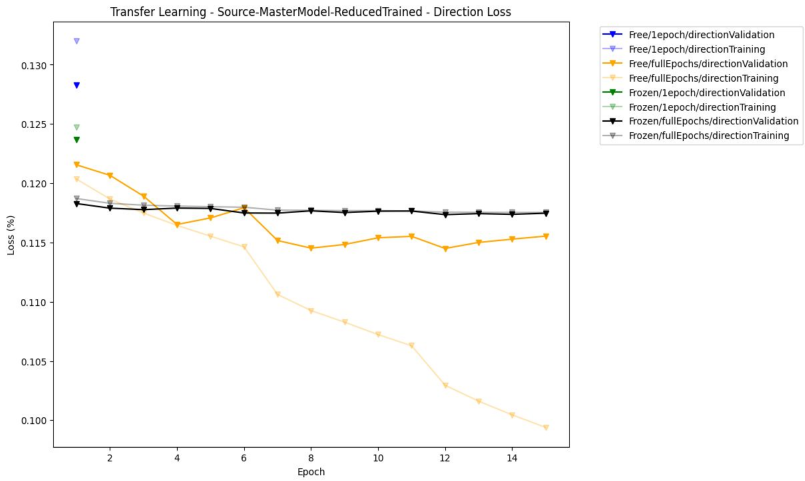
Share knowledge between models

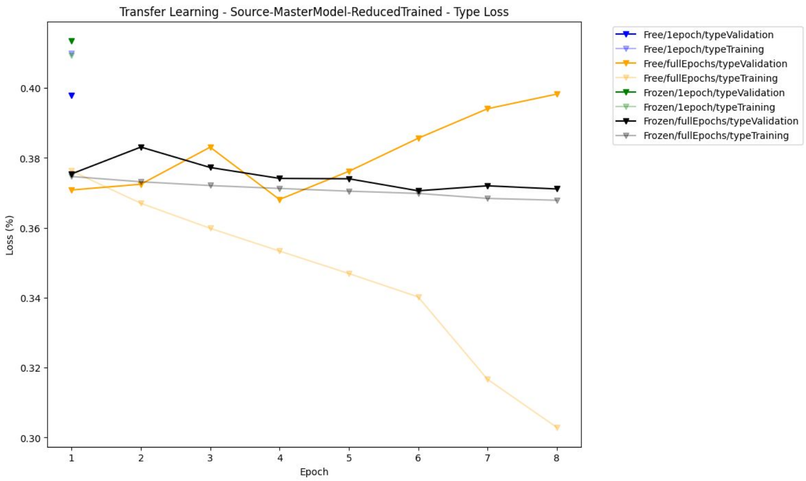
- Variants
  - Differente source models
  - Backbone (Free or Frozen)
  - Train the second model for different number of epochs

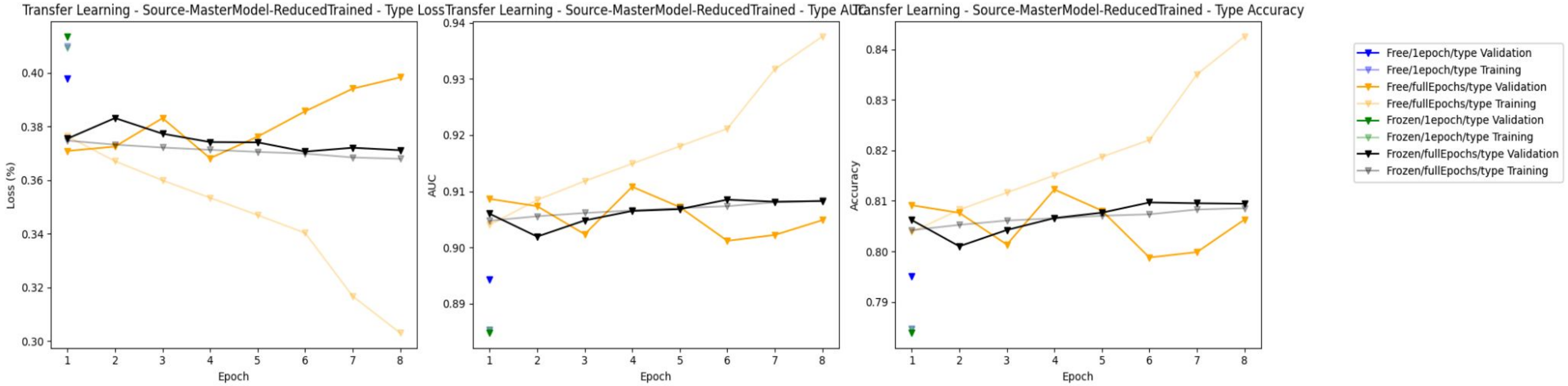
- Goals
  - Save training time
  - Have a main source model



CTLearn - Transfer Learning - Train



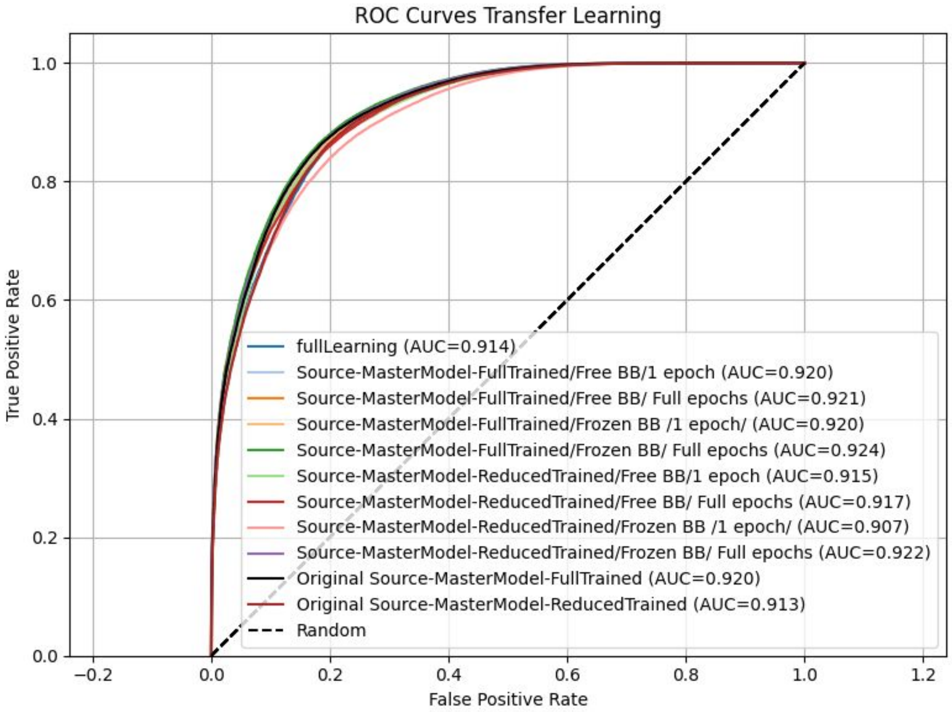




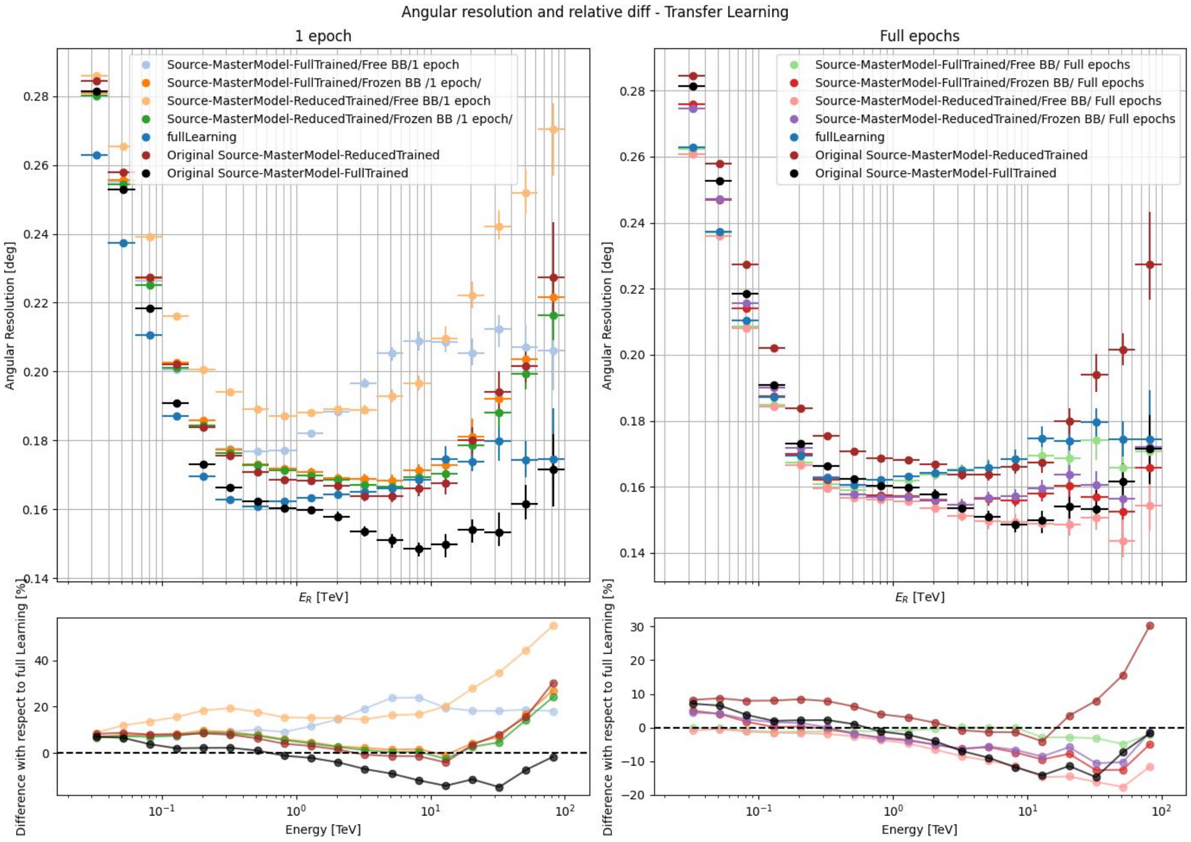
CTLearn - Transfer Learning - Train

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**

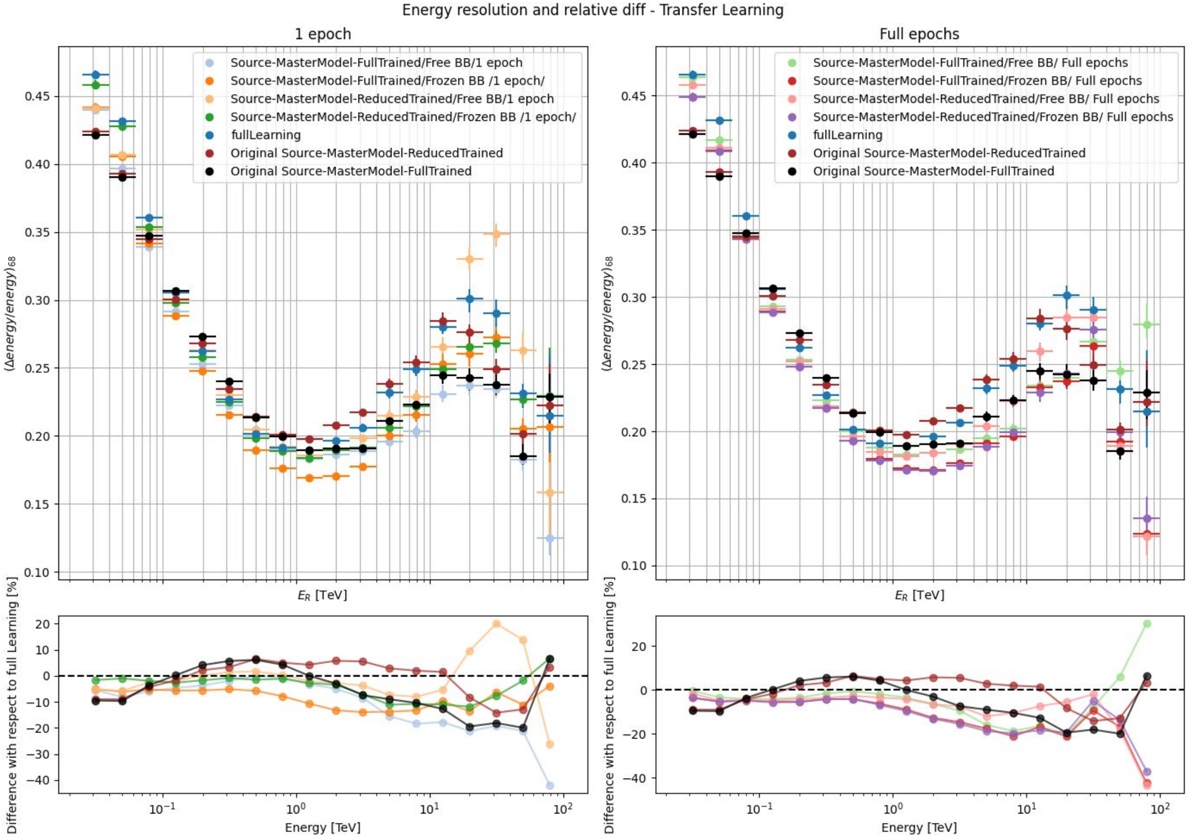
CTLearn - Transfer Learning - Test



CTLearn - Transfer Learning - Test

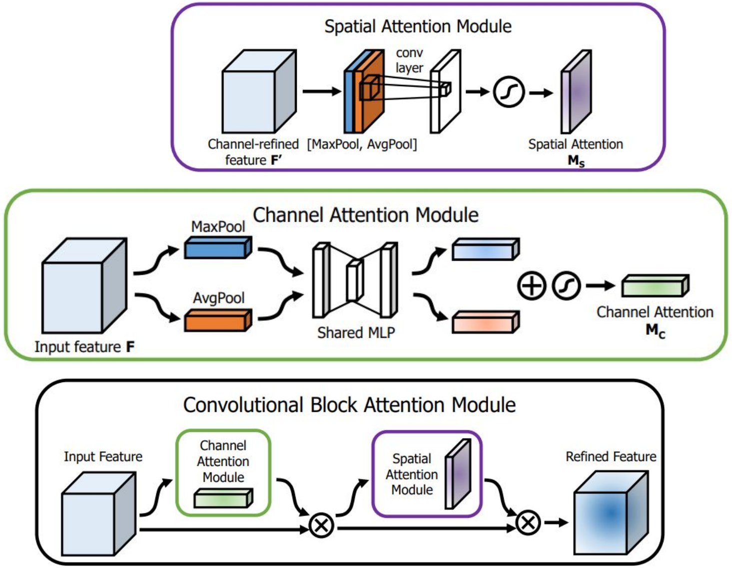


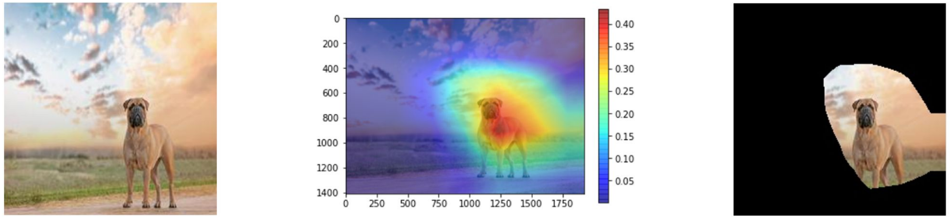
CTLearn - Transfer Learning - Test



Spatial attention (pixels) and channel attention (time, image)

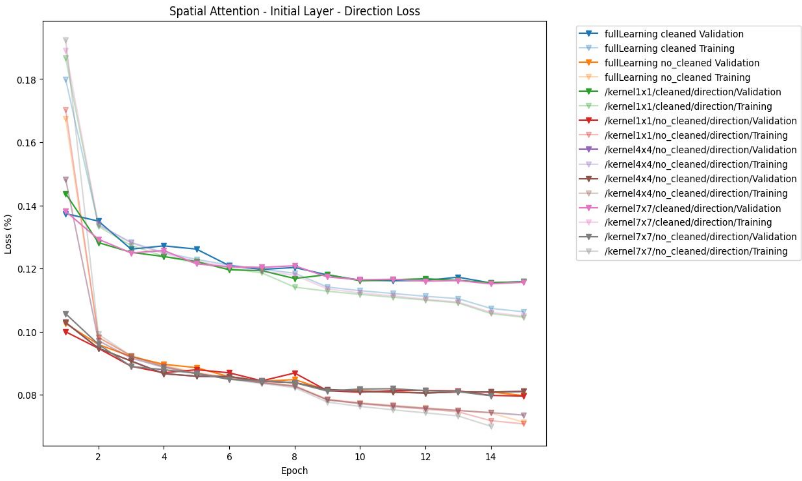
- Variants
  - Structure (Initial layer, between convolutional blocks)
  - Data (Clean; not clean)
  - Kernel (1x1, 4x4, 7x7)
- Goals
  - Understanding CNN’s approaches when predicting
  - Trying to avoid cleaning process

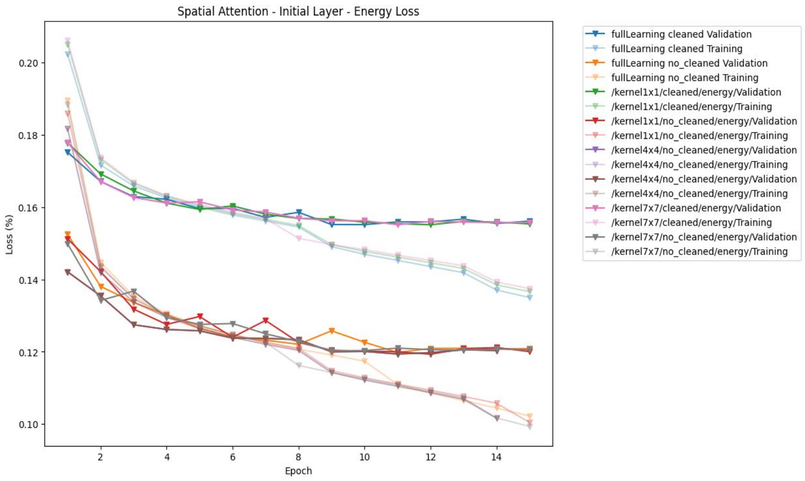


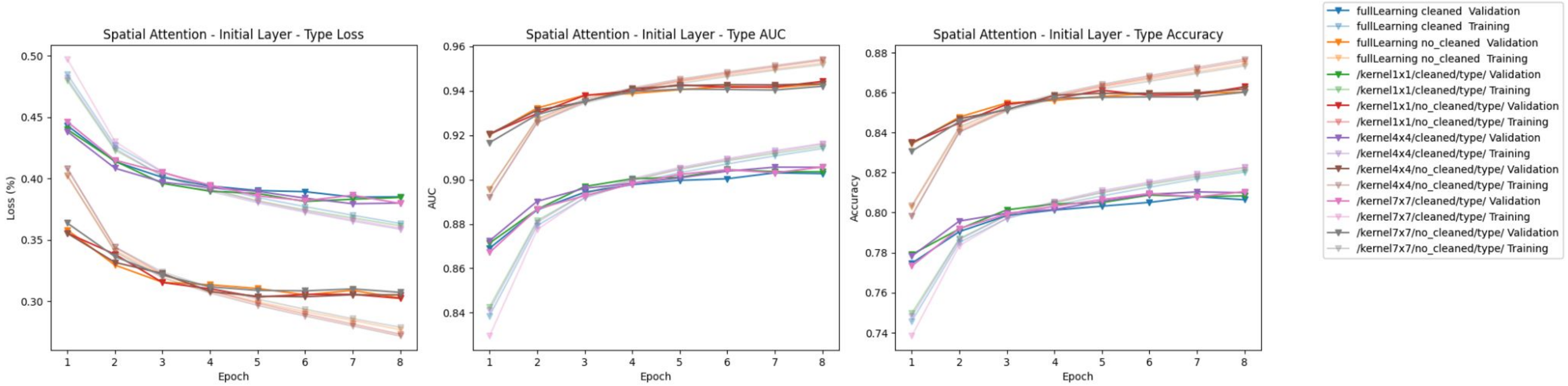


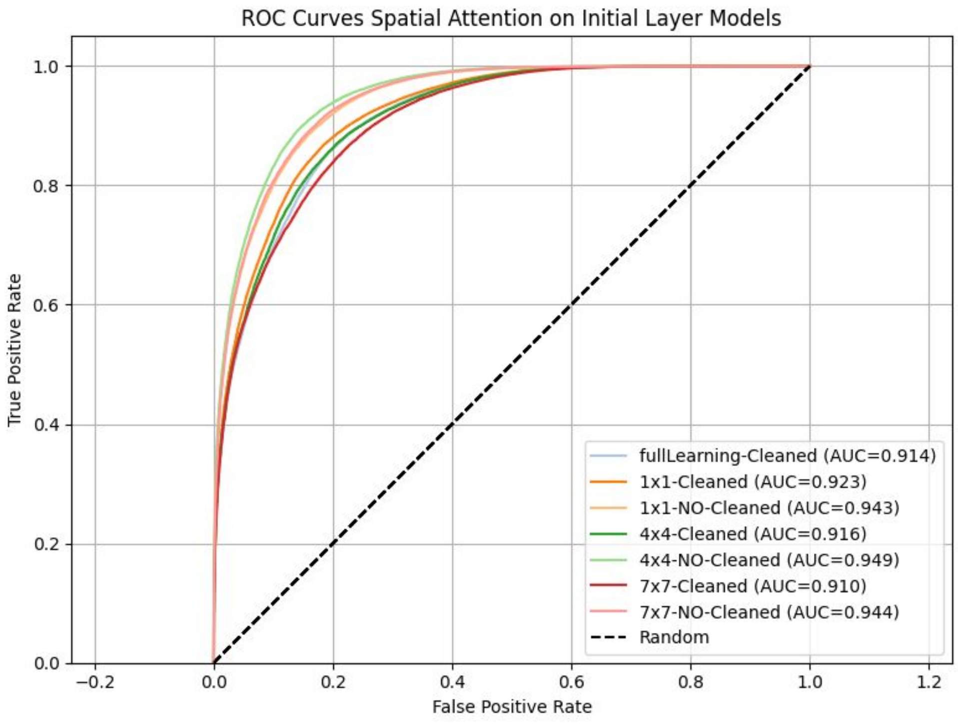
(a) Original image (b) Attention map (c) Visual explanation map

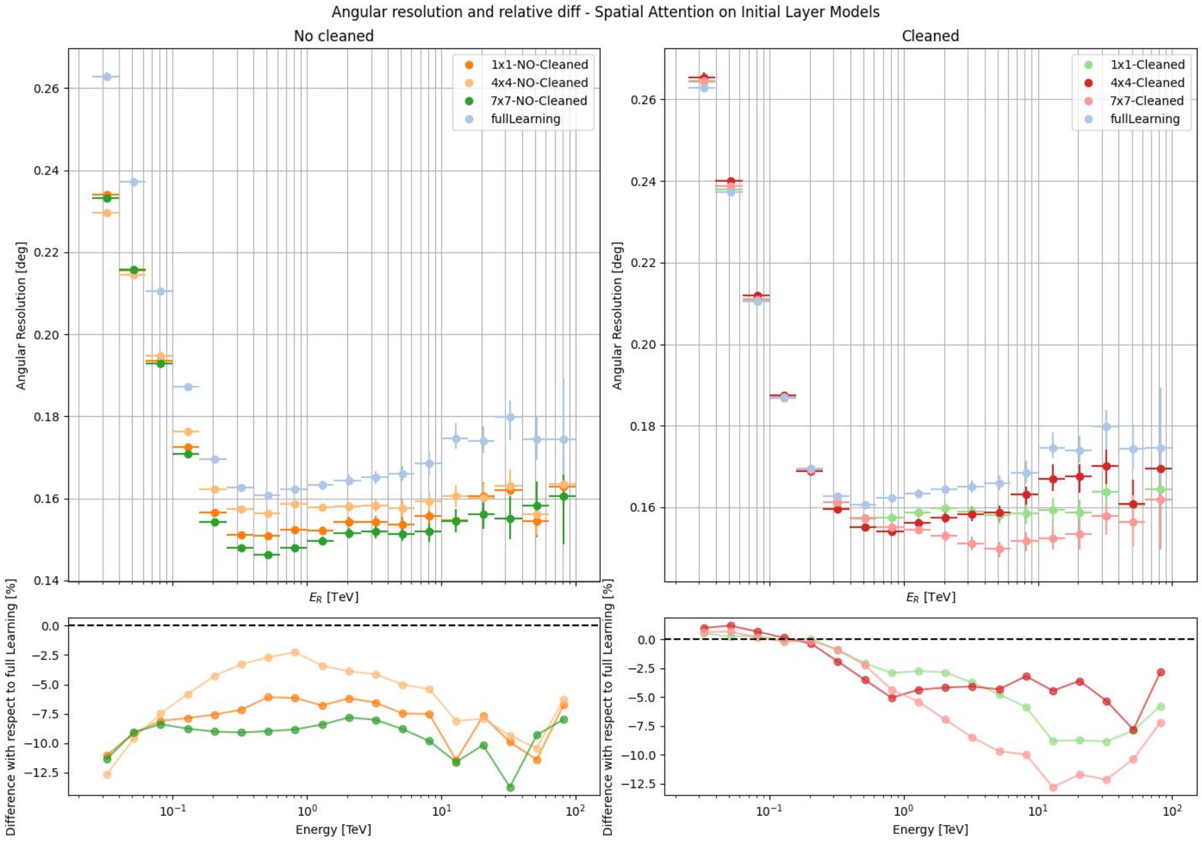
Artemisa
ARTificial Environment for ML and Innovation in Scientific Advanced Computing


Spatial Attention - Initial Layer - Direction Loss
Loss (%)
Epoch
fullLearning cleaned Validation
fullLearning cleaned Training
fullLearning no_cleaned Validation
fullLearning no_cleaned Training
/kernel1x1/cleaned/direction/Validation
/kernel1x1/cleaned/direction/Training
/kernel1x1/no_cleaned/direction/Validation
/kernel1x1/no_cleaned/direction/Training
/kernel4x4/no_cleaned/direction/Validation
/kernel4x4/no_cleaned/direction/Training
/kernel4x4/no_cleaned/direction/Validation
/kernel4x4/no_cleaned/direction/Training
/kernel7x7/cleaned/direction/Validation
/kernel7x7/cleaned/direction/Training
/kernel7x7/no_cleaned/direction/Validation
/kernel7x7/no_cleaned/direction/Training

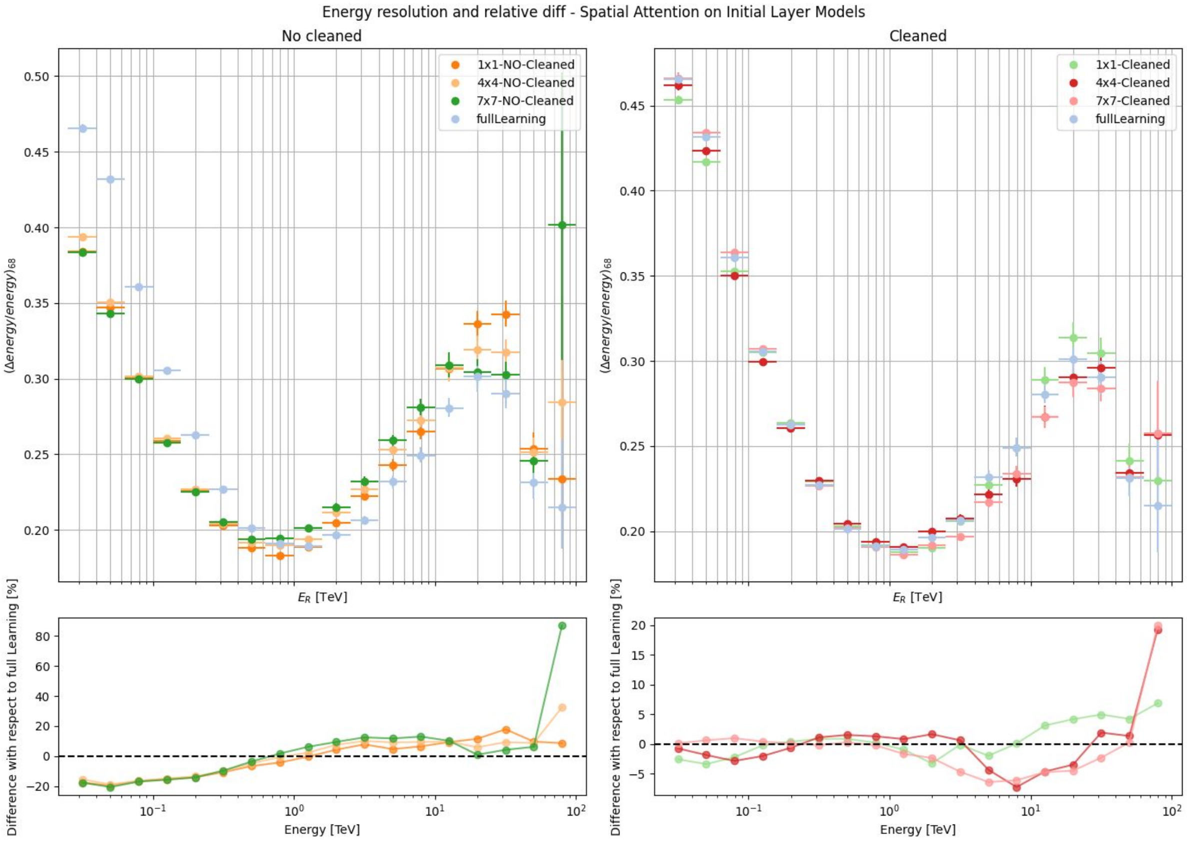

Spatial Attention - Initial Layer - Energy Loss
Loss (%)
Epoch
fullLearning cleaned Validation
fullLearning cleaned Training
fullLearning no_cleaned Validation
fullLearning no_cleaned Training
/kernel1x1/cleaned/energy/Validation
/kernel1x1/cleaned/energy/Training
/kernel1x1/no_cleaned/energy/Validation
/kernel1x1/no_cleaned/energy/Training
/kernel4x4/no_cleaned/energy/Validation
/kernel4x4/no_cleaned/energy/Training
/kernel4x4/no_cleaned/energy/Validation
/kernel4x4/no_cleaned/energy/Training
/kernel7x7/cleaned/energy/Validation
/kernel7x7/cleaned/energy/Training
/kernel7x7/no_cleaned/energy/Validation
/kernel7x7/no_cleaned/energy/Training


Spatial Attention - Initial Layer - Type Loss
Loss (%)
Epoch
Spatial Attention - Initial Layer - Type AUC
AUC
Epoch
Spatial Attention - Initial Layer - Type Accuracy
Accuracy
Epoch
fullLearning cleaned Validation
fullLearning cleaned Training
fullLearning no_cleaned Validation
fullLearning no_cleaned Training
/kernel1x1/cleaned/type/ Validation
/kernel1x1/cleaned/type/ Training
/kernel1x1/no_cleaned/type/ Validation
/kernel1x1/no_cleaned/type/ Training
/kernel4x4/cleaned/type/ Validation
/kernel4x4/cleaned/type/ Training
/kernel4x4/no_cleaned/type/ Validation
/kernel4x4/no_cleaned/type/ Training
/kernel7x7/cleaned/type/ Validation
/kernel7x7/cleaned/type/ Training
/kernel7x7/no_cleaned/type/ Validation
/kernel7x7/no_cleaned/type/ Training


ROC Curves Spatial Attention on Initial Layer Models
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
fullLearning-Cleaned (AUC=0.914)
1x1-Cleaned (AUC=0.923)
1x1-NO-Cleaned (AUC=0.943)
4x4-Cleaned (AUC=0.916)
4x4-NO-Cleaned (AUC=0.949)
7x7-Cleaned (AUC=0.910)
7x7-NO-Cleaned (AUC=0.944)
Random


Angular resolution and relative diff - Spatial Attention on Initial Layer Models
No cleaned
Cleaned
1x1-NO-Cleaned
4x4-NO-Cleaned
7x7-NO-Cleaned
fullLearning
1x1-Cleaned
4x4-Cleaned
7x7-Cleaned
Angular Resolution [deg]
E_R [TeV]
Difference with respect to full Learning [%]
Energy [TeV]

CTLearn - Attention experiments - Test



THANK YOU

rcervino@ucm.es

https://apod.nasa.gov/apod/ap200724.html r.com

Acknowledgment

The research here presented has been partially supported by the former Spanish Ministry of Economy, Industry, and Competitiveness / ERDF grants FPA2015-73913-JIN and FPA2017-82729-C6-3-R, the Spanish Ministry of Science and Innovation grant PID2019-104114RB-C32, PID2022-138172NB-C42, PDC2023-145839-I00, NSF awards PHY-1229205, 1229792, and 1607491, the Community of Madrid grant C. M. (2023) PEJ-2023-AI, and the European Science Cluster of Astronomy & Particle Physics ESFRI Research Infrastructures funded by the European Union's Horizon 2020 research and innovation program under Grant Agreement no. 824064. The authors acknowledge support from Google LLC through the Google Summer of Code program and NVIDIA Corporation with the donation of a Titan X Pascal GPU used for part of this research.