MASTER OF SCIENCE
IN ENGINEERING

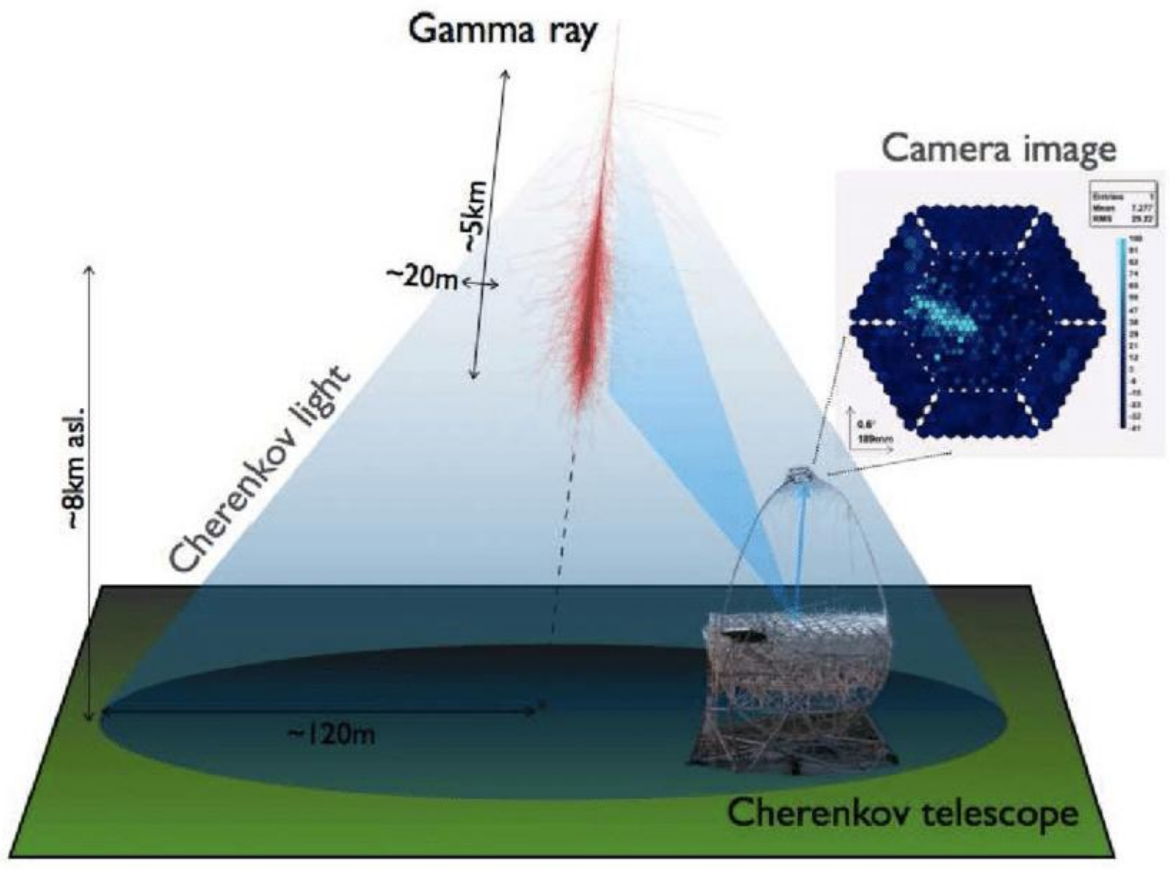
# Machine Learning for Cherenkov Telescope SST-1M


**Hugo Varenne**
Master of Science HES-SO in Engineering
Major in Data Science
Under the supervision of Upegui Posada Andres
In collaboration with HEPIA and UNIGE


## Introduction

Gamma rays are produced when cosmic rays interact with matter, radiation fields, or magnetic fields in extreme astrophysical environments such as supernovae or black holes. When these gamma rays enter the Earth’s atmosphere, they cannot be detected directly and instead generate extensive air showers, emitting faint Cherenkov light as secondary particles travel faster than the speed of light in air. This Cherenkov light is captured by Imaging Atmospheric Cherenkov Telescopes (IACTs), which record its spatial and temporal patterns as images and waveforms. These images are then analyzed using machine learning and deep learning techniques, notably within the CTLearn framework, to reconstruct the primary particle type, energy, and arrival direction.

**Figure 1. Gamma ray interaction with the atmosphere captured by a Cherenkov telescope**




## Material and methods

**CTLearn** is a Python library that uses deep learning to analyze images from Cherenkov telescopes. It helps reconstruct gamma-ray events by predicting the particle type, energy, and arrival direction from the recorded Cherenkov light.

**MLOps** (Machine Learning Operations) is a set of practices that makes machine learning models easier to develop, reproduce, and deploy. It focuses on automation, experiment tracking, and scalable execution, which is especially important when training complex models on large datasets and HPC infrastructures.

An **HPC cluster** is a high-performance computing system that uses multiple CPUs and GPUs to run large and computationally intensive tasks in parallel, making it possible to efficiently train deep learning models on large datasets.

**Data:** Simulated gamma-ray events → stored in HDF5 → includes calibrated waveforms, image features, reconstructed events & ground truth → used to train models for particle type, energy & direction.

## Tools & Concepts

**MLOps Tools:** Modular tasks for training, testing, report generation, and model comparison enable independent execution and fault tolerance. Configuration files manage experiment parameters for reproducibility. Reports and visualizations provide metrics and performance insights. The setup supports automation, monitoring, and scalable, production-ready model management.

**Custom Model:** A template was implemented to define custom models in CTLearn, separating backbone (high-level features) and head (task-specific predictions). This structure allows fine-tuning, reusability, and integration of pretrained weights while ensuring compatibility with the library.

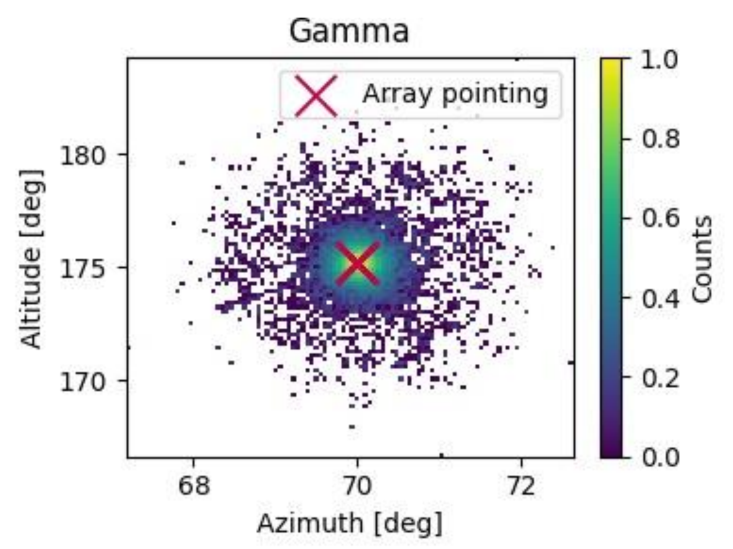



**Figure 2. Graphic from the model report about Alt-Az distribution**

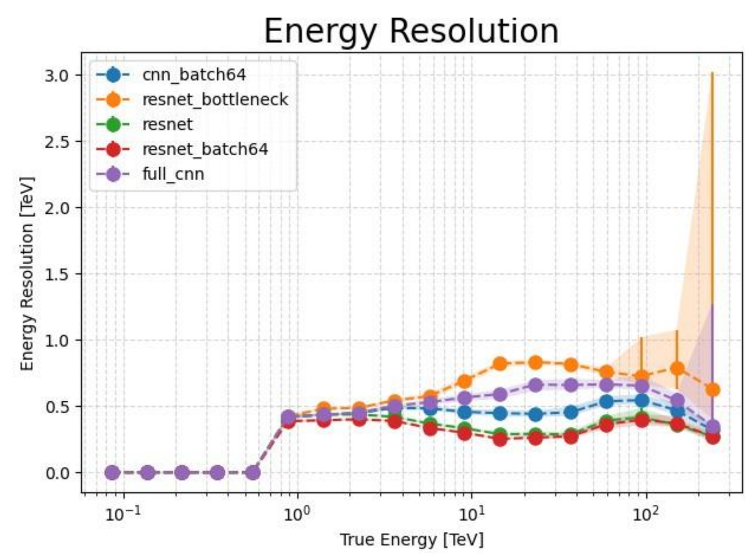



**Figure 3. Graphic from the comparison of model for the energy resolution**

## Conclusion

This work demonstrates how modern MLOps practices can significantly improve deep learning workflows for gamma-ray reconstruction with Imaging Atmospheric Cherenkov Telescopes. By integrating automation, configuration-driven experiments, standardized reporting, and scalable HPC workflows into the CTLearn framework, the thesis bridges astrophysics expertise with machine learning engineering. Beyond model performance, the results highlight that reproducibility, scalability, and maintainability are critical for complex deep learning architectures in scientific environments. This project serves as a proof-of-concept showing that MLOps is a key enabler for sustainable and efficient model development in VHE gamma-ray astronomy and more especially in the CTLearn library.


## Acknowledgments

I would like to thank all people involved in the CTLearn library, especially Andres Upegui Posada, Matthieu Heller, Laurent Gantel and Jakub Kvapil, who contributed to this project with their guidance, feedback and support. Their contributions had a positive impact on this thesis and its results.


## References


1.CTAO, https://www.ctao.org/, website for Cherenkov imaging domain
2. CTLearn, https://arxiv.org/abs/1912.09877, Research paper about the Python library
3. Ultra-High-Energy Gamma-Ray Astronomy, **https://doi.org/10.1146/annurev-nucl-112822-025357**


Hes·so Master