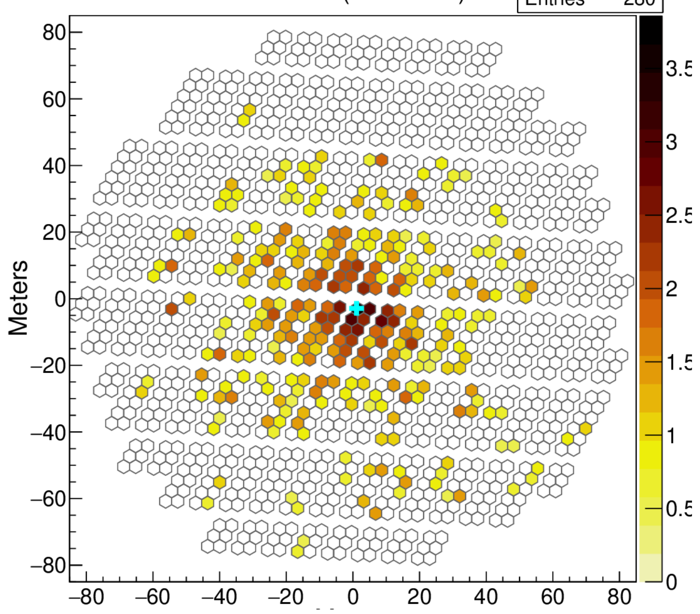
Studies of Gamma-Ray Shower Reconstruction Using Deep Learning

Tomas Bylund,[a,*] Gašper Kukec Mezek,[a] Mohanraj Senniappan,[a] Yvonne Becherini,[a] Michael Punch,[a,b] Satyendra Thoudam[c] and Jean-Pierre Ernenwein[d]

*Presenter; [a]Linnaeus University, Department of Physics and Electrical Engineering; [b]Université de Paris, CNRS, Astroparticule et Cosmologie; [c]Khalifa University, Department of Physics; [d]Aix Marseille Univ, CNRS/IN2P3, CPPM.

The Cosmic Multiperspective Event Tracker[1] (CoMET) R&D project aims to optimize the techniques for the detection of soft-spectrum sources through very-high-energy gamma-ray observations using particle detectors (called ALTO detectors), and atmospheric Cherenkov light collectors (called CLiC detectors). In this contribution, we leverage Convolutional Neural Networks (CNNs) using only particle detectors, aiming to improve reconstruction performance at lower energies (< 1 TeV) as compared to the SEMLA analysis procedure, which uses a more conventional method using manually derived features.

Figure 1



The ALTO particle detector array consists of 1242 detector units, distributed in a circular area of ~ 160 m in diameter, with each unit belonging to a hexagonally packed cluster combining 6 detector units, as can be seen in the Figure 1 on the right, also showing a 1.1 TeV event.

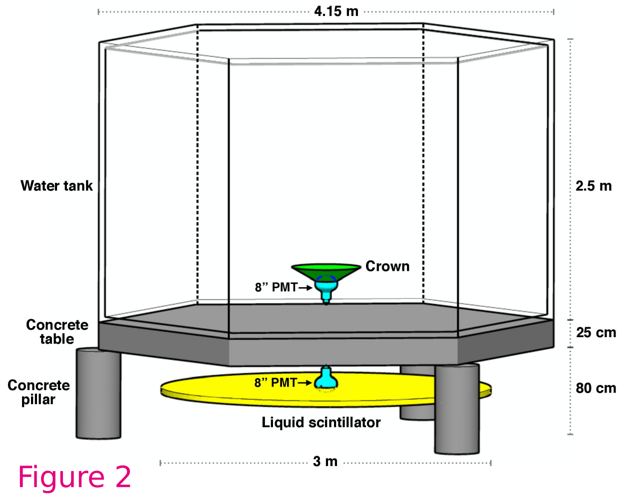


Figure 2

Each ALTO unit consists of a hexagonal Water Cherenkov Detector (WCD) positioned on a concrete slab, with a cylindrical liquid Scintillator Detector (SD) underneath, see Figure 2 on the left. Both the WCD and the SD in one unit are instrumented by an 8" Hamamatsu Photo Multiplier Tube (PMT) recording the photons emitted by the passage of charged particles.

The WCDs are used for the detection of particles in the air shower, while the SDs are conceived for signal over background discrimination through muon tagging.

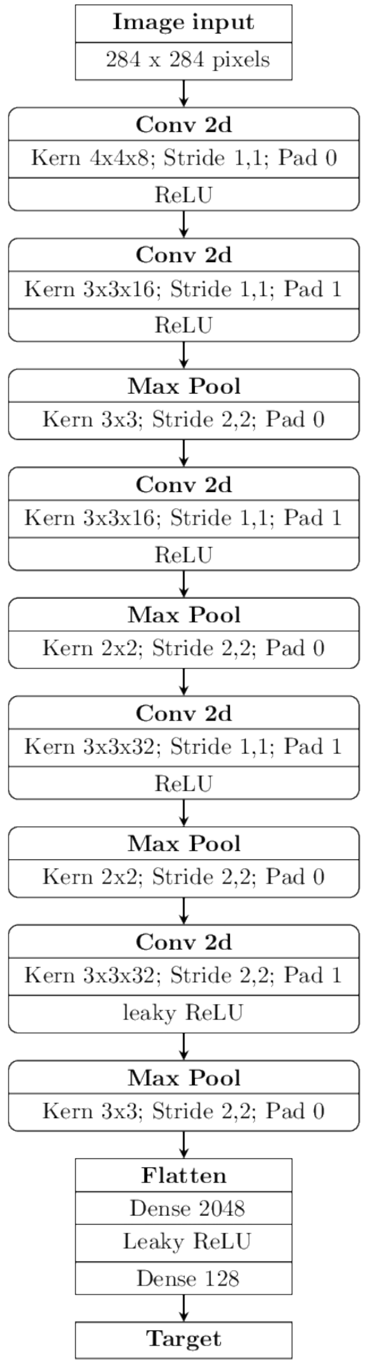


Figure 3

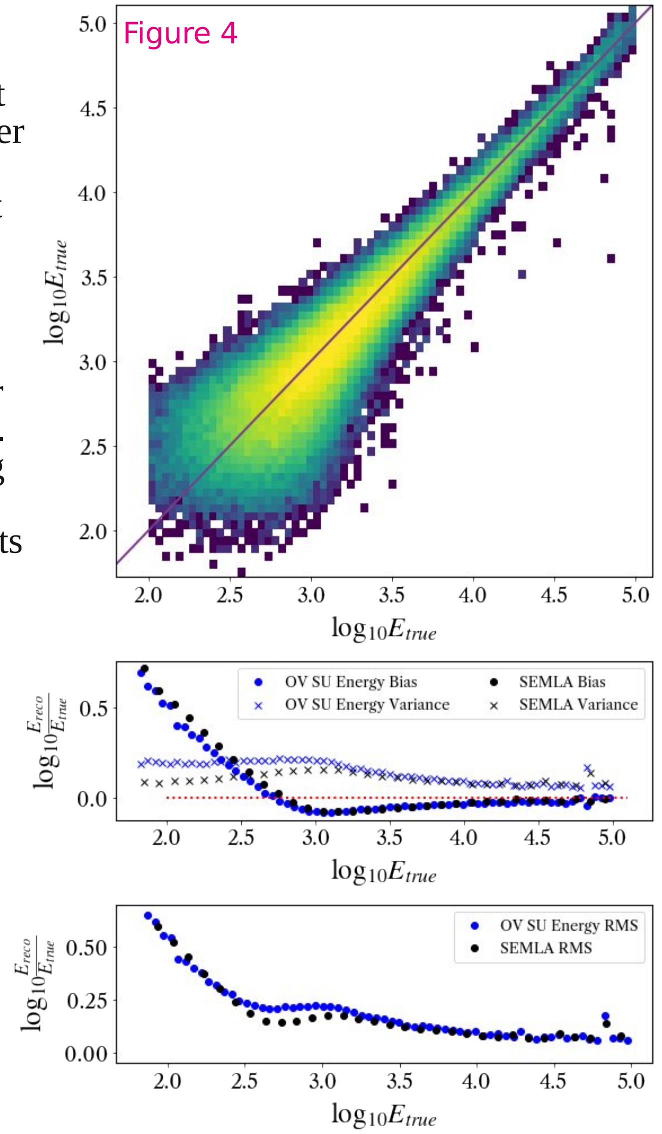
To convert events to a format the CNN can use, the complete array is subdivided into numerous square pixels of 60 cm to a side, resulting in a 285 by 285 pixel image that preserves the spatial distribution of the water tanks.

The layout used here is a simple series of gradual downsampling using unpadded convolutions combined with sharp resizing through max pooling, and is illustrated in detail in Figure 3 to the left. The network was implemented in PyTorch (1.4.0) using the PyTorch-lightning (1.2.8) framework and trained using the Adam optimiser with default settings and a fixed learning rate of 0.001, in batches of 16, targeting either $\log_{10}$ E (results shown in Figure 4 below) or $\log_{10} X_{max}$ (results shown in Figure 5 to the right).

CoMET aims at the detection of soft-spectrum sources, but the impact of the effective area results in very few events at 200 GeV. This means that we have to train our CNN on data that is imbalanced over our region of interest. To offset this we explore two variants of "**data pre-processing**": random under- or surplus sampling. In undersampling you randomly throw away events until reaching a desired balance; surplus sampling is the converse idea: instead of discarding events, the sampling is performed by randomly repeating events[2].

In total the CNN was trained using three different datasets: the One Pass normal train-validation split, the SMall undersampled dataset, and the larger SUrplus sampled dataset.

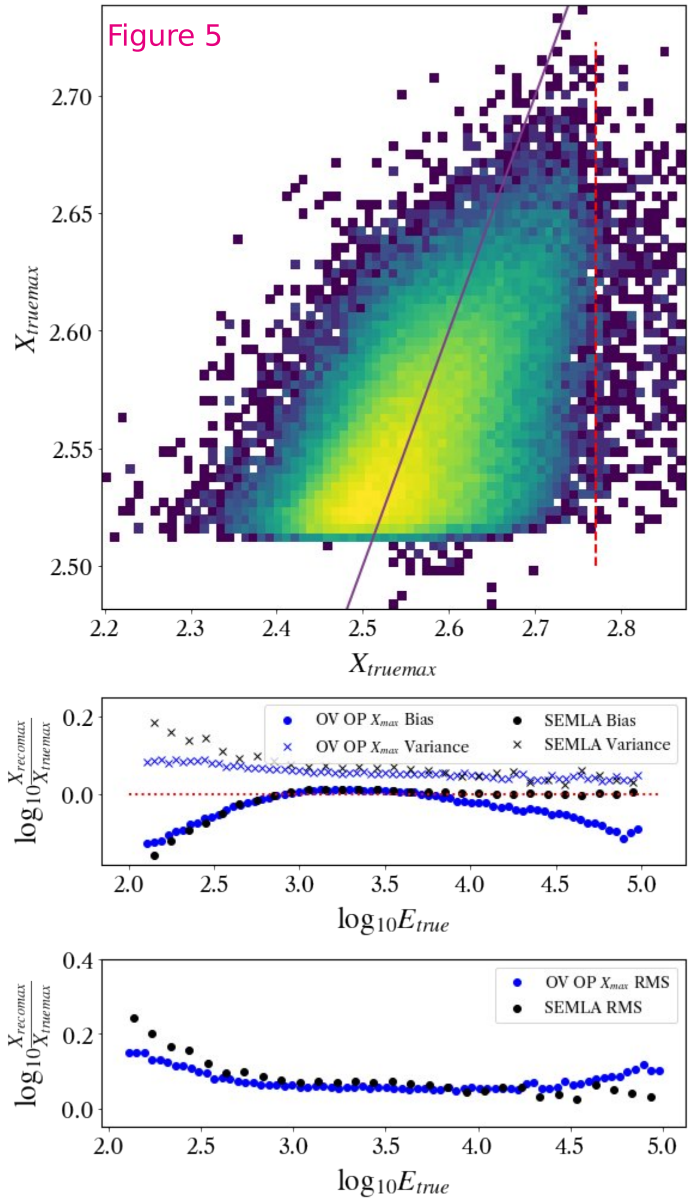
Figure 4



We **measure the perfomance** using the RMS (around zero) of $\log_{10}(a_{reco}/a_{truth})$ in bins of energy, with "a" representing either energy or shower maximum.

Results of the best **shower maximum model** (highest accuracy and quality), trained with the One Pass sample, is shown in the figure on the right, along with results for SEMLA. Here the CNN achieves a better performance. In the transfer matrix in Figure 5 (on the right) the purple line indicates perfect correlation, and red dashed line shows the depth of the detector. We see that the (shallow) high-altitude events are correctly reconstructed far away from the detector location, and this information might be useful in a future version of the ALTO reconstruction.

Figure 5



In Figure 4 (on the left) we show a comparison of the best performing **energy regression** model, trained on the surplus sampled dataset, along with the SEMLA result. In this case, the traditional approach proves hard to beat, as the CNN is unable to match its overall performance (see bottom left panel).

Conclusions

We have presented attempts to reconstruct the Energy of VHE gamma-rays, and X_{max} of the resulting air showers, using only the total charge seen in the ALTO particle detector surface array with convolutional neural networks.

- It is challenging to reconstruct low energies with particle detectors; some models did reach better RMS perfomance but at the cost of larger systematics
- The CNN proved uniformly better at reconstructing X_{max} compared to the traditional machine learning approach of SEMLA, indicating that in the data there are unexplored correlations between the observables and X_{max}
- There is some gain in preprocessing unbalanced datasets, but it does not appear to be essential to achive good results.

References:

1. G. Kukec Mezek et al.,*The COMET multiperspective event tracker for wide field-of-viewgamma-ray astronomy*, PoS ICRC2021(2021) (at this conference)

2. Branco et al., *A survey of predictive modeling on imbalanced domains*, ACM ComputingSurveys (CSUR) 49 (2016)

Acknowledgements:

https://alto-gamma-ray-observatory.org/acknowledgements