Reconstruction of stereoscopic CTA events using deep learning with CTLearn

T. Miener[1], D. Nieto[1], A. Brill[2], S. Spencer[3], J. L. Contreras[1]
for the CTA Consortium[4]

[1]Universidad Complutense de Madrid, [2]Columbia University, [3]University of Oxford,
[4]See www.cta-observatory.org

cherenkov telescope array

CTLearn

ABSTRACT

The Cherenkov Telescope Array (CTA), conceived as an array of tens of imaging atmospheric Cherenkov telescopes (IACTs), is an international project for a next-generation ground-based gamma-ray observatory, aiming to improve on the sensitivity of current-generation instruments a factor of five to ten and provide energy coverage from 20 GeV to more than 300 TeV. Arrays of IACTs probe the very-high-energy gamma-ray sky. Their working principle consists of the simultaneous observation of air showers initiated by the interaction of very-high-energy gamma rays and cosmic rays with the atmosphere. Cherenkov photons induced by a given shower are focused onto the camera plane of the telescopes in the array, producing a multi-stereoscopic record of the event. This image contains the longitudinal development of the air shower, together with its spatial, temporal, and calorimetric information. The properties of the originating very-high-energy particle (type, energy and incoming direction) can be inferred from those images by reconstructing the full event using machine learning techniques. In this contribution, we present a purely deep-learning driven, full-event reconstruction of simulated, stereoscopic IACT events using CTLearn. CTLearn is a package that includes modules for loading and manipulating IACT data and for running deep learning models, using pixel-wise camera data as input.

Supported by:

IPARCOS COLUMBIA UNIVERSITY

Introduction

The Cherenkov Telescope Array (CTA) is the next-generation ground-based gamma-ray observatory, aiming to improve on the sensitivity of current-generation instruments by a factor of five to ten and provide energy coverage from 20 GeV to more than 300 TeV. In this contribution, we explore how deep learning algorithms like deep convolutional neural networks (DCNs) can be utilized to perform full-event reconstruction in single-telescope and multi-telescope mode.

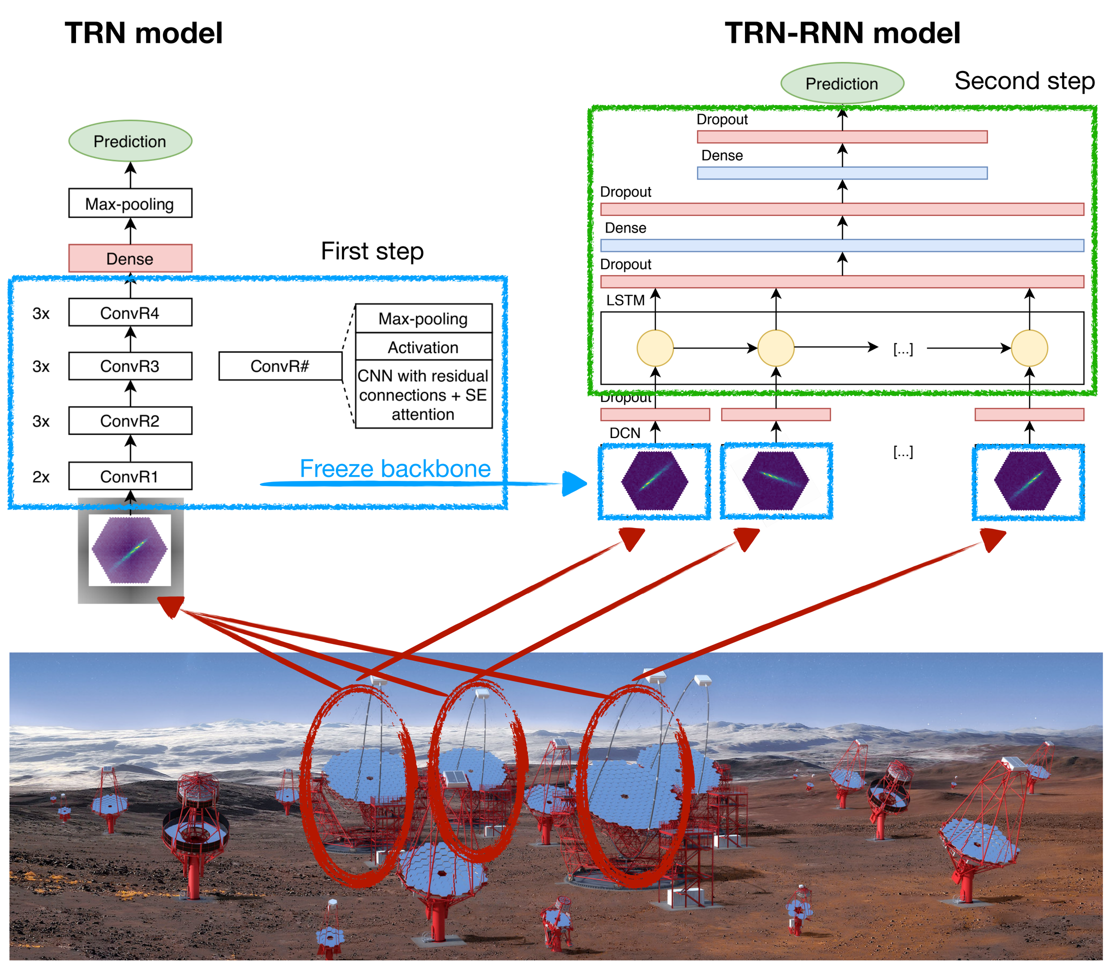

Fig. 1: Diagram depicting the main layers of the *TRN* (left) and the *TRN-RNN* model (right) used in this work.

CTA analysis workflow with CTLearn

The training of the deep learning models and their prediction, the actual full-event reconstruction, are performed with *CTLearn*[1]. *CTLearn* is a high-level, open-source *Python* package providing a backend for training deep learning models for IACT event reconstruction using *TensorFlow*[2]. This work focus on the *TRN* model (left), a deep DCN-based architecture with residual connections adapted from a thin ResNet and on the *TRN-RNN* model (right), multiple *TRN* blocks connected via a recurrent neural network.

Dataset

The subarray layout *M5C5*, consists of 13 medium-size telescopes (MSTs) and 40 small-size telescopes (SSTs), and 4 large-size telescopes (LSTs), which will potentially be added to the array in the future, are considered in this work. We used the CTA South (20° north pointing) reference dataset, processed with *ctapipe*[3]. Data loading and pre-processing, especially designed for deep learning purposes, are managed using the *DL1-Data-Handler*[4] package.

Results

The *TRN* model is trained on ~ 200k batches of 64 images for each telescope type, validating periodically. As expected, the LST is the most sensitive telescope type at the lowest energies; the sensitivity of an MST is best where this telescope type will be responsible of the full-array sensitivity; the SST is providing competitive sensitivity at the highest energies. The LST, MST, and SST provide excellent energy resolutions of ~ 13%, 9%, and 10% at their best, angular resolutions of 0.12°, 0.13°, and 0.1° at their best and area under the ROC curve (AUC) of 0.89, 0.944, and 0.959, in their entire energy ranges, respectively.

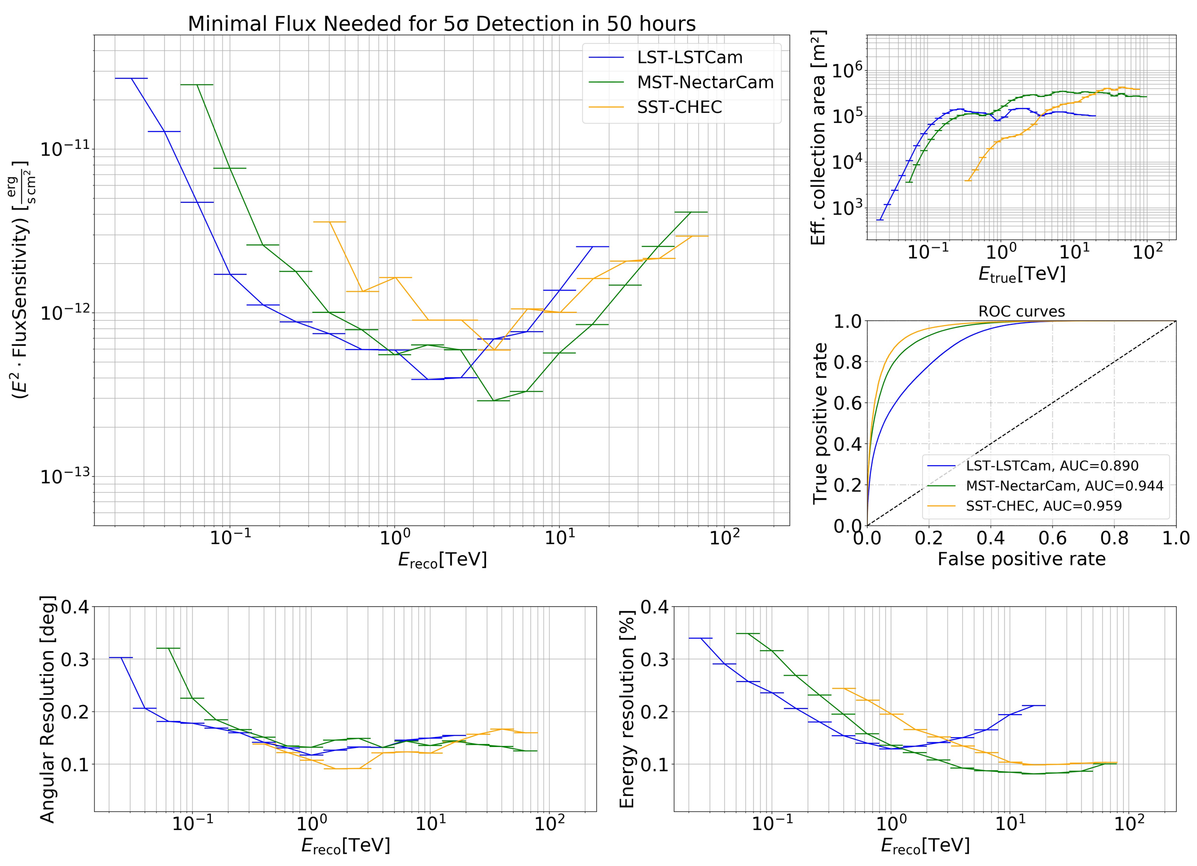

Fig. 2: Single telescope IRFs and sensitivities obtained with the *TRN* model.

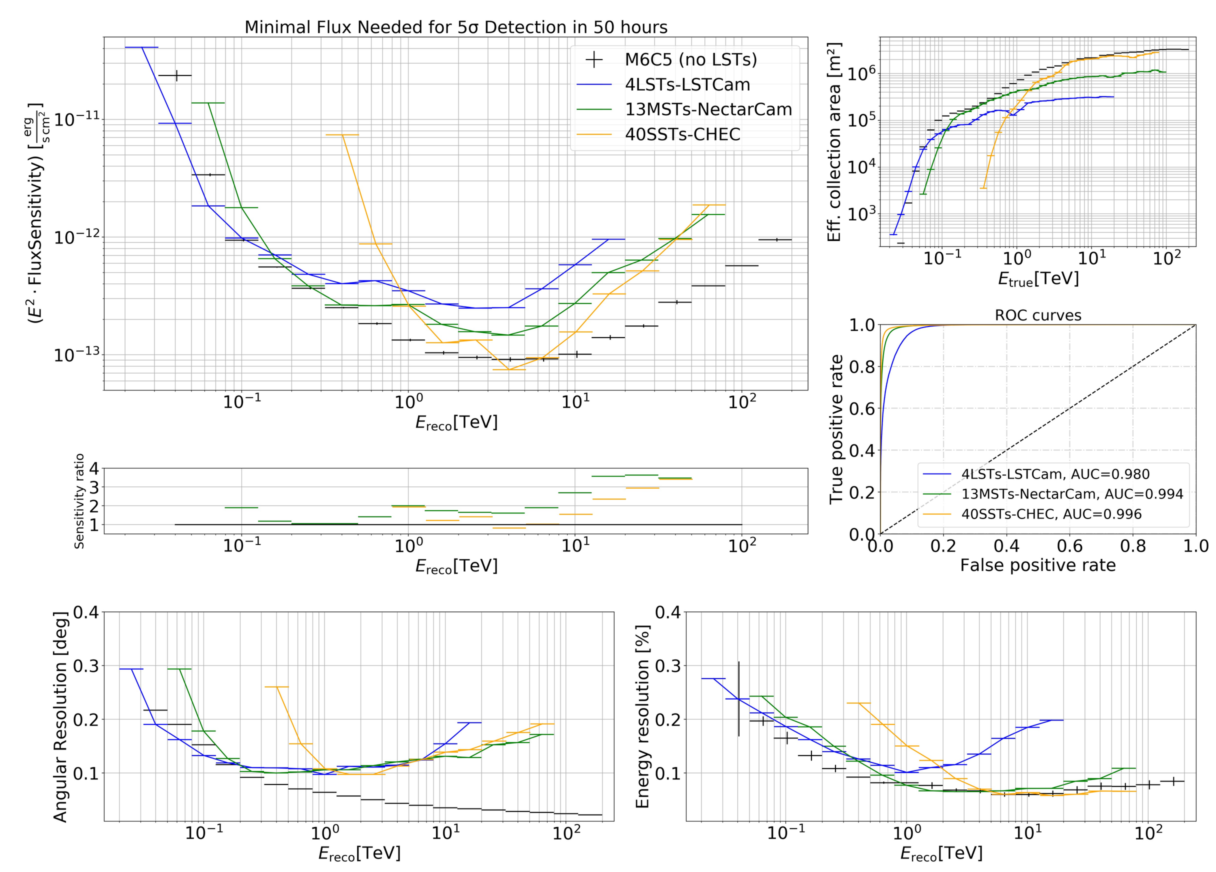

Fig. 3: Multi telescopes IRFs and sensitivities obtained with the *TRN-RNN* model.

The particle classification, performed by the *TRN-RNN* model, works well, with an AUC score of 0.98, 0.994, and 0.996 for the subarrays of 4 LSTs-LSTCam, 13 MSTs-NectarCam, and 40 SSTs-CHEC, respectively. The three subarrays reach promising top values for the energy resolution of ~ 10%, 7%, and 6% for LSTs, MSTs and SSTs, respectively. The *TRN-RNN* model performs poorly on the reconstruction of the arrival direction. The angular resolution for the highest energies differs from the conventional analysis significantly, which translates to the sensitivity curves, causing a deficit of performance especially at energies above 10 TeV. A fair comparison to the conventional analysis (black curves in Fig. 3) is not feasible at this stage of the development, because no LSTs are considered in the conventional analysis, and it is not limited to a per-telescope-type analysis.

Conclusion

This contribution shows for the first time that DCN-based full-event reconstruction works for all sizes of CTA telescopes, in both single-telescope and stereo modes. The performance of the *TRN* and the *TRN-RNN* models for the particle classification and the energy estimation is promising. Tackling the arrival direction reconstruction task via DCNs requires additional modifications and improvements to the existing stereoscopic deep learning models to reach the requirements of CTA. Future developments of CTLearn will include the combination of different telescope types to evaluate the full-array performance of CTA North and South with deep learning models. Multitask learning experiments, where one single model performs the IACT specific tasks, as well as applying DCNs model to observational data, are also planned.

ACKNOWLEDGEMENTS

We gratefully acknowledge financial support from the agencies and organizations listed in this link.
TM acknowledges support from PID2019-104114RB-C32. DN and JLC acknowledges partial support from The ESCAPE Research Infrastructures funded by the European Union's Horizon 2020 research and innovation program under Grant Agreement no. 824064. This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council (STFC, www.dirac.ac.uk).This work used IRIS computing resources funded by the STFC. SS acknowledges an STFC PhD studentship. We acknowledge the support of NVIDIA Corporation with the donation of a Titan X Pascal GPU used for part of this research.

[1]https://github.com/ctlearn-project/ctlearn [2]https://www.tensorflow.org/ [3]https://github.com/cta-observatory/ctapipe [4]https://github.com/cta-observatory/dl1-data-handler