Full LST-1 Data Reconstruction Using Convolutional Neural Networks

Jakub Juryšek[1], Etienne Lyard[1], and Roland Walter[1]

[1]*University of Geneva, Astronomy Department, Geneva, Switzerland;*
`jakub.jurysek@unige.ch`
On behalf of the CTA LST project

Abstract.   The Cherenkov Telescope Array (CTA) will be the world's largest and most sensitive ground-based gamma-ray observatory in the energy range from a few tens of GeV to tens of TeV. The LST-1 prototype, currently in its commissioning phase, is the first of the four largest CTA telescopes, that will be built in the northern site of CTA in La Palma, Canary Islands, Spain. In this contribution, we present a full-image reconstruction method using a modified `InceptionV3` deep convolutional neural network applied on non-parametrized shower images. We evaluate the performance of optimized networks on Monte Carlo simulations of LST-1 shower images, and compare the results with the performance of the standard reconstruction method. We also show how both methods work on real-data reconstruction.

1.   Introduction

When primary gamma-ray photons or cosmic-ray particles enter the atmosphere, they interact with atomic nuclei and produce showers of secondary particles emitting Cherenkov radiation. In ground-based gamma-ray astronomy, this radiation can be observed by Imaging Atmospheric Cherenkov Telescopes (IACTs), forming roughly elliptical images of the shower in the camera plane of a focusing telescope. In typical IACT observations, the trigger rate from diffuse cosmic-ray background is about 1000x higher than the trigger rate from gamma-ray photons even for the brightest gamma-ray sources in the sky, and therefore a strong background suppression is necessary. Besides reducing this background, the goal of the image analysis is to reconstruct the energy and incoming direction of the primary gamma-ray. In this study, we compare the reconstruction performance of standard Random Forests (RFs) (Breiman 2001) with a novel method using a deep convolutional neural network (CNN) on Monte Carlo (MC) simulations (Heck et al. 1998; Bernlöhr 2008) and first data from the LST-1 telescope prototype for CTA (CTA Consortium 2019).

2.   Methods

We used a modified `InceptionV3` CNN architecture (Fig. 1) (Szegedy et al. 2016; Lyard et al. 2020) consisting of a variable number of inception modules followed by five sequential convolutional layers, whose capability of gamma/hadron separation was

previously proven by Lyard et al. (2020). In this contribution, we study the performance of this network for full event reconstruction. Our networks were trained on MC gamma and MC proton shower images from a standard CTA MC production data set (Prod5) for LST-1. We used `adadelta` and `rmsprop` optimizers for gamma/hadron separation and energy/direction reconstruction, respectively. Unlike RFs, which need image parameters on input leading to a loss of information, we trained CNN directly on shower images and times of maxima of waveform in each pixel which follow a temporal gradient, calibrated in `lstchain v0.6.3` [1]. In order to get a standard rectangular input image for the `Keras` Python library, hexagonal images from the telescope had to be reshaped following the method introduced in Lyard et al. (2020). Our results suggest that such reshaping doesn't introduce any position reconstruction bias if more than 3 inception modules are used. We also applied relatively weak selection cuts on MC events (Intensity[2] > 50 p.e., leakage[3] < 0.2), to create a MC sample that was used for training/testing of CNNs and RFs consistently. The number of inception modules and dropout were optimised for each reconstruction task on a subset of 10^5 training events (left Fig. 2). Parameters providing the best results in terms of Area Under the Curve (AUC)[4], energy and angular resolution were used for final training of each CNN on 5×10^5 events.
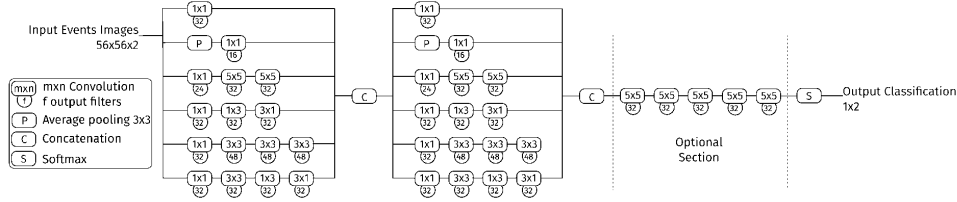

Figure 1. Architecture of modified `InceptionV3` used in this study. Number of inception modules was optimized for each reconstruction task. Figure from Lyard et al. (2020).

3. Reconstruction performance

Figure 3 (left) shows receiver operating characteristic curve (ROC) for gamma/hadron separation trained on simulated diffuse gammas and diffuse protons. A comparison with the separation power of RFs trained/tested on Hillas parameters extracted from the same MC sample is also shown, and it can be clearly seen that CNN significantly outperform RFs. Angular and energy resolution curves are shown in Figure 2 (right) and Figure 3 (right), respectively. For the mid-energy range the angular resolution

---

[1]`https://github.com/cta-observatory/cta-lstchain/releases/tag/v0.6.3`

[2]The total integrated signal in pixels containing Cherenkov light emitted by the shower particles, after removal of the background pixels.

[3]The fraction of the shower signal contained in the two outermost pixel rings of the camera.

[4]Area under the ROC curve, usually being used as a measure of the ability of a classifier to distinguish between classes.
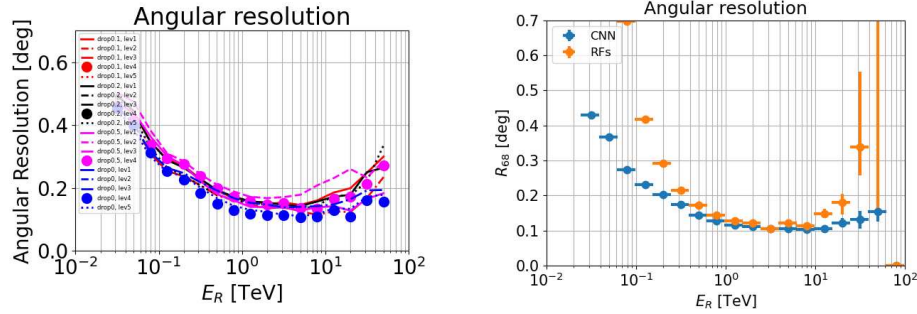
Figure 2.   *Left:* An example of angular resolution evaluated on testing point gammas for different parameters of CNN. *Right:* Angular resolution of reconstruction by optimized CNN compared with Random Forests.

improvement gained from CNN is not significant, but it's very prominent for low and high energies, which can be a hint that CNN are particularly good in the reconstruction of noise-dominated low energy shower images, or showers cut by the camera edges. In energy reconstruction, RFs outperform CNNs by 5%. The worse reconstruction power of CNN in this case could be caused by image normalisation in the preprocessing step and needs to be addressed in a following study.
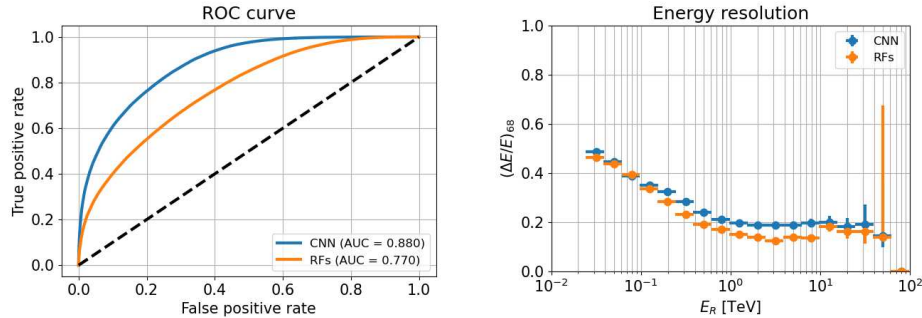

Figure 3.   *Left:* ROC for gamma/hadron separation evaluated on on-axis point gammas and diffuse protons. *Right:* Energy resolution of CNN reconstruction compared with Random Forests.

4.   Analysis of real data

Despite its good performance on simulations, CNN often fails when applied to real data due to subtle differences between the two (e.g. different night sky background level, or stars present in the real field of view), and thus the evaluation of the performance on the real data is essential. We used CNN and RF trained on the same MC events to reconstruct events from Crab Nebula observations from November 2020 (19 mins on-source and 15 mins off-source), and reached a similar detection significance of 6.2σ for both methods after carefully matching background levels using an adaptive gammaness

cut on the CNN reconstructed data (gammaness cut for RFs processed data was 0.8), and common $\Theta^2 < 0.05$ deg^2 cut on the source region. A significance map for the CNN reconstructed data is shown in Figure 4 (left). The Θ^2 distribution for both methods (right Fig. 4) shows that contrary to the performance evaluated on MC, the angular resolution of the CNN on data is worse than we expected, which is probably caused by MC-data differences, which will be addressed in a following study.
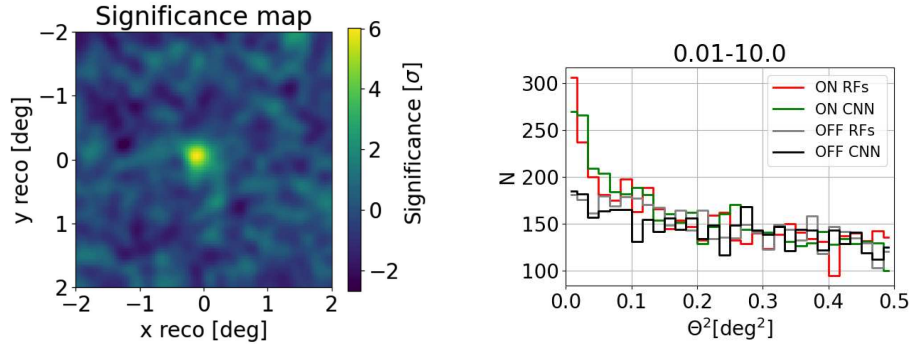

Figure 4.   *Left:* Significance map of events reconstructed with CNN. *Right:* Θ^2 distribution of events in energy range 0.01- 10 TeV.

5.   Conclusions

In this contribution, we demonstrated a potential of CNNs to be used not only for gamma/hadron separation, but also for full event reconstruction with a competitive performance to the standard RFs reconstruction. Even though the CNN performance on real data seems to be slightly worse than expected from the full MC study, there is much room for improvement, particularly in MC-data tuning.

References

Bernlöhr, K. 2008, Astroparticle Physics, 30, 149. `0808.2253`

Breiman, L. 2001, Machine Learning, 45, 5. URL `https://doi.org/10.1023/A:1010933404324`

CTA Consortium 2019, Science with the Cherenkov Telescope Array (WORLD SCIENTIFIC). `https://www.worldscientific.com/doi/pdf/10.1142/10986`, URL `https://www.worldscientific.com/doi/abs/10.1142/10986`

Heck, D., Knapp, J., Capdevielle, J. N., Schatz, G., & Thouw, T. 1998, CORSIKA: a Monte Carlo code to simulate extensive air showers. (TIB Hannover)

Lyard, E., Walter, R., Sliusar, V., & Produit, N. 2020, in Journal of Physics Conference Series, vol. 1525 of Journal of Physics Conference Series, 012084. `1907.02428`

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. 2016, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)