



HAL
open science

First Full-Event Reconstruction from Imaging Atmospheric Cherenkov Telescope Real Data with Deep Learning

Mikaël Jacquemont, Thomas Vuillaume, Alexandre Benoit, Gilles Maurin,
Patrick Lambert, Giovanni Lamanna

► **To cite this version:**

Mikaël Jacquemont, Thomas Vuillaume, Alexandre Benoit, Gilles Maurin, Patrick Lambert, et al.. First Full-Event Reconstruction from Imaging Atmospheric Cherenkov Telescope Real Data with Deep Learning. International Conference on Content-Based Multimedia Indexing (CBMI), Jun 2021, Lille, France. 6 p., 10.1109/CBMI50038.2021.9461918 . hal-03241475v2

HAL Id: hal-03241475

<https://hal.science/hal-03241475v2>

Submitted on 28 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

First Full-Event Reconstruction from Imaging Atmospheric Cherenkov Telescope Real Data with Deep Learning

Mikaël Jacquemont*, Thomas Vuillaume*, Alexandre Benoit†, Gilles Maurin*,

Patrick Lambert† and Giovanni Lamanna*, for the CTA LST project

*Univ. Savoie Mont Blanc, CNRS, LAPP - IN2P3, Annecy, France

Email: {firstname.lastname}@lapp.in2p3.fr

†LISTIC, Univ. Savoie Mont Blanc, Annecy, France

Email: {firstname.lastname}@univ-smb.fr

Abstract—The Cherenkov Telescope Array is the future of ground-based gamma-ray astronomy. Its first prototype telescope built on-site, the Large Size Telescope 1, is currently under commissioning and taking its first scientific data. In this paper, we present for the first time the development of a full-event reconstruction based on deep convolutional neural networks and its application to real data. We show that it outperforms the standard analysis, both on simulated and on real data, thus validating the *deep* approach for the CTA data analysis. This work also illustrates the difficulty of moving from simulated data to actual data.

Index Terms—deep learning, gamma astronomy, multitasking, learning bias, model generalization to real data

I. INTRODUCTION

Astronomy, as many other domains requiring performant data analysis methods, is giving more and more attention to deep neural networks. Gamma-ray astronomy is the observation of the most energetic photons produced by violent astrophysics phenomena, such as supernova explosions, neutron star mergers or the environment of black holes. The analysis of such phenomena will be enhanced by the Cherenkov Telescope Array (CTA), an Imaging Atmospheric Cherenkov Telescope (IACT) representing the next generation of ground-based gamma-ray telescopes, currently under construction. It will be composed of two arrays, one in each hemisphere to allow for maximal sky coverage, and will consist of tens of telescopes. The telescope prototype being built on-site on the Canary island of La Palma, the first Large Size Telescope (LST1) has started taking commissioning data recently. As shown in Fig. 1, its method of observation, as is standard for such telescopes, relies on the observation of the Cherenkov

light stimulated by the development of the atmospheric shower of particles occurring when a gamma-ray photon enters the atmosphere. This emitted light is collected by a mirror, and captured by a specific ultra-fast and very sensitive optical camera [1]. Gamma-ray-initiated showers then appear as roughly elliptical images on the camera plane.

The goal of an IACT data analysis is to reconstruct the interesting physical parameters - the particle type (gamma-ray photon vs. cosmic-ray background), its energy and incoming direction (altitude, azimuth). Several methods have been developed to achieve this event reconstruction in current generation IACTs. The most common one relies on the characterization of the moments of the image [2]. The extracted parameters are then combined with multivariate analysis methods, such as boosted decision trees or random forests [3]. This approach, referred to as Hillas+RF in the rest of this paper, is robust, but has limited sensitivity, especially at low energies. On the other hand, state-of-the-art methods rely on template matching between the captured images and a huge dictionary, with the help of a likelihood function [4]. However, these methods are slow [5], and consequently might not be practical for CTA data analysis in their current implementation. Indeed, to achieve its best sensitivity, CTA requires the most physically accurate analysis methods, but also methods that are computationally efficient to be able to deal with the large volume of data that will be produced (several PB per year to be analyzed again every year to benefit from calibration and analysis improvements). Deep neural networks may thus offer a solution to this challenge.

In this paper, we present the recent developments made to address the event reconstruction problem using deep learning, and the new results obtained both on simulated and real data from LST1. We first compare the performance of a deep multitask architecture and a widespread analysis method on the most recent production of simulated data, which takes into account the updated knowledge of the LST1. We also use an adaptation method for the simulation to better correspond to the real data. Then, we realize the first ever full-event reconstruction from IACT data with deep learning, highlighting the

We gratefully acknowledge financial support from the agencies and organizations listed here: www.cta-observatory.org/consortium_acknowledgment. This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreements No 653477 and No 824064, and from the Fondation Université Savoie Mont Blanc. This work has been done thanks to the facilities offered by the Univ. Savoie Mont Blanc - CNRS/IN2P3 MUST computing center and HPC resources from GENCI-IDRIS (Grant 2020-AD011011577) and computing and data processing resources from the CNRS/IN2P3 Computing Center (Lyon - France). We gratefully acknowledge the support of the NVIDIA Corporation with the donation of one NVIDIA P6000 GPU for this research.

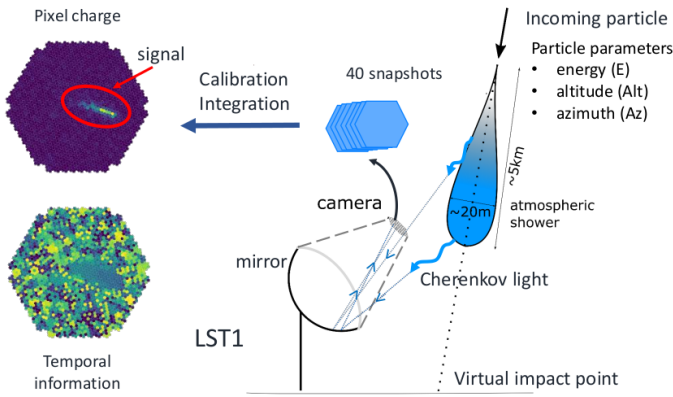


Fig. 1. Large-Sized Telescope 1 (LST1) principle.

complexity of transferring the good performance obtained on simulations to real data.

II. RELATED WORK

Applying deep learning techniques to the analysis of real IACT data is recent. In the High Energy Stereoscopic System (H.E.S.S.)¹ experiment, two studies have been carried out using stereoscopic data (i.e., composed of the data from the four telescopes of H.E.S.S. I), addressing only a part of the gamma-ray event reconstruction. Shilon *et al.* [6] address the gamma/proton classification and direction regression tasks. They propose a combination of a convolutional neural network (CNN) and a Recurrent Neural Network, denoted CRNN, to handle the gamma/proton classification task. The role of the recurrent part of the model is to combine the information coming from the different telescopes. For the direction regression, they adopt a different strategy. They incorporate the data coming from the different telescopes as the different channels of a unique image, and use a shallow CNN to complete the task. Analyzing real data, they observe a decrease of the performance compared to the one on simulated data for the direction regression. Starting from the CRNN architecture, Parsons *et al.* [7] propose to address the gamma/proton classification task by combining IACT images and standard method parameters. They observe that the obtained architecture is sensitive to the level of Night Sky Background (NSB) present in the data.

In the context of CTA, deep learning approaches have been explored, all analysis of simulated data. In [8], Nieto *et al.* probe very deep networks for gamma/proton classification from single telescope images. Reference [9] presents a shallower CNN to address, from stereoscopic data, gamma/proton classification, and energy and direction regression tasks. For the LST event reconstruction, the *TRN-single-tel* model proposed in [10] consists of a custom ResNet [11] augmented with the Squeeze-and-Excitation attention mechanism [12]. Each task (i.e., gamma/proton classification, energy and direction regression) is addressed with a different model. This strategy introduces several limits, increasing the computational cost

both at training and inference time while not taking into account the interdependence between the parameters to estimate.

III. PRESENT WORK

Different from these approaches, γ -PhysNet DA presented in depth in [13] is a deep multi-task network that performs the full-event reconstruction from LST1 data with a single architecture. It is composed of a ResNet encoder and a physically inspired multi-task block. More precisely, the encoder is the convolutional part of the ResNet-56 implemented with indexed convolutions [14], and augmented with Dual Attention mechanism [15]. As LST1 produces hexagonal pixel images, indexed convolutions allow applying convolution directly to the input data, avoiding additional preprocessing steps. In [13], γ -PhysNet DA has proven to outperform the Hillas+RF method on all the tasks of the event reconstruction on simulated data. Also, this work highlights the appeal of the multi-task approach, thus taking into account the dependencies between the target predictions. However, these preliminary experiments were limited to simulated data.

In this paper, we extend the work of [13] to evaluate how the performance obtained on simulations transfers to real data, highlighting the learning bias due to the difference between simulated and real data. In a first step, following the standard analysis procedure, we train γ -PhysNet DA and the Hillas+RF method on a simulated dataset (called Prod5, see Section IV-A). Compared to the data used in [13] this data set is new, and integrates the updated knowledge of the LST1 brought by the ongoing commissioning phase. Both models are then applied to the real data. However, differences remain between these updated simulations and the real data analyzed in this paper, such as the acquisition conditions or the NSB level. As emphasized in Section II, and as we have observed in a preliminary analysis, the NSB level has a strong impact on the performance of the gamma/proton classification task with deep learning approaches. Therefore, in a second step, we measure the difference in NSB level between the simulations and the real data. We then adapt the simulations as explained in Section IV-C to reduce this difference in NSB level, retrain both analysis models, and reapply them to the real data.

On both the simulations and the real data the Hillas+RF method trained on standard Prod5 simulations, representing the standard approach, serves as a baseline.

IV. DATA SETS

A. Simulated Data

Ground-truth labels are impossible to obtain from IACT real data. To overcome this issue and train the reconstruction models, accurate Monte Carlo simulations have been developed to simulate the atmospheric shower development [16] and the LST1 response [17]. These simulations are used to develop analysis pipelines, train machine learning algorithms, and test the reconstruction predictions. In this work, we use the Prod5 LST1 mono trigger simulation dataset produced by the LST consortium which, for now, remains a private dataset. Compared to the simulations used in [13], the Prod5 relies on

¹<http://www.mpi-hd.mpg.de/hfm/HESS/pages/about/telescopes>

the latest knowledge of the LST1 acquired during the first months of the telescope commissioning for the simulation of its response at the pointing of 20° zenith.

B. Real Data Samples of the Crab Nebula

The real dataset is composed of two observation runs from the Crab nebula, a standard candle for gamma-ray astronomy, taken in February 2020. The observation #2013 (28.9° zenith), denoted ON run in the rest of the paper, was taken with the telescope pointing to the gamma source direction, and the observation #2012 (21.4° zenith), denoted OFF run, was taken with the telescope pointing to a dark patch of the sky. The runs contain respectively $10.9 M$ and $10.4 M$ events. As detailed later in Section VI, the OFF run allows for the estimation of the background noise that has to be subtracted from the events recorded from the source direction to obtain the gamma-ray excess.

C. Data Preparation

Simulated and real raw data are made of spatiotemporal data cubes of 40 samples of 1 ns each, called waveforms. The waveforms first undergo a calibration and integration phase (the same in both cases) using `lstchain v0.6.3` [18] in its standard configuration. This results in integrated images composed of two channels: the first, referred to as charge channel, integrating the number of photoelectrons per pixel, the second containing their mean arrival time (see Fig. 1), thus compressing the temporal information. These images will be used as input in the later stages.

Besides, although we benefit from high quality updated simulations to prepare the model, there still exist discrepancies between the training data and the real data produced by the LST1 that make the real data analysis challenging. The telescope pointing direction differs slightly between real data and simulations, but more importantly, the NSB level is different. Fig. 2 illustrates the difference of NSB level between the simulations and the real data analyzed in this work. To overcome this issue, we add a Poisson noise to the simulated data. In the simulations, the NSB is indeed defined as a Poisson distribution. Following Raikov’s theorem [19], we determine the parameter λ of the Poisson distribution as the difference between the average noise pixel charge of the simulation and the one of the real data. The models trained on this adapted dataset are denoted + *Poisson noise* in the rest of the paper.

D. Data Selection

Finally, these images are further selected following quality cuts common in the field and required by the Hillas+RF method: the integrated signal in the charge channel must be superior to 50 photoelectrons, the ratio of the signal on the camera edges to the total signal must be lower than 20%, and the charge channel must pass a two-level filter, called cleaning operation, requiring pixels to be above a threshold of 6 photoelectrons and to have a neighbor above 3 photoelectrons. Image samples for which none pixel survives the cleaning are discarded. To allow for a fair comparison, the exact same

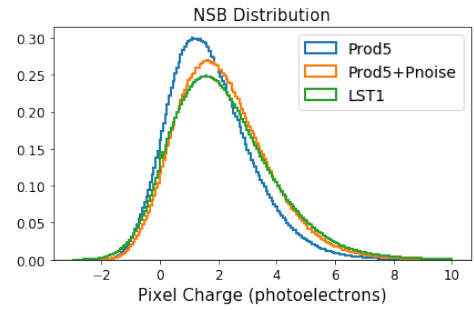


Fig. 2. Distribution of NSB pixels in the simulated data (Prod5), compared to the one of the Crab data (LST1). Prod5 + Pnoise corresponds to the addition of a Poisson noise to the simulation data.

image samples are thus used with γ PhysNet DA. However, the images themselves are not cleaned from the noise in the case of γ PhysNet DA, in order to keep the information from all pixels.

After this preparation and selection phase, the test set is composed of:

- for simulated data: 993 *k* gamma-ray images from a point-like source located at the center of the field of view, 510 *k* proton images and 773 *k* electron images
- the real data as described above, for a total of 12.36 *M* images

It is worth noticing that the models are trained on gamma rays with direction isotropically distributed within 6° of the telescope pointing, denoted diffuse gamma rays, so as to reconstruct events coming from any direction within the field of view. In addition, no selection is applied to the training set, leaving 1.25 *M* diffuse gamma-ray images and 836 *k* proton images.

V. PERFORMANCE ON SIMULATED DATA

To evaluate how both γ -PhysNet DA and the Hillas+RF method adapt to the simulated data, we compare their performance on the simulation test set, with (denoted + *Poisson noise*) and without the addition of Poisson noise for NSB level discrepancy reduction.

A. Performance Metrics

The **sensitivity** is a measure of the overall performance of the model. This metric is specific for gamma astronomy, and is regarded as the most relevant one because of its global aspect. It represents, per energy bin, the gamma-ray flux that an observed point-like source should emit to allow a detection with significance [20] of 5 standard deviations (denoted σ) above the background fluctuations for a 50-hour observation, with a gamma-ray excess of at least 10 events, and of at least 5% of the residual background, per energy bin. In this paper, we present the sensitivity of the different methods relatively to the performance of the standard Hillas+RF method trained on standard Prod5 simulations.

The **energy resolution** represents the performance of the model for the energy reconstruction task. It is computed as

the 68% containment of the relative error of the model for the energy regression task, per energy bin.

The **angular resolution** is a measure of the performance of the model for the direction regression task. It represents the angular separation in which 68% of the reconstructed gamma rays fall, per energy bin.

For the resolution curves, lower values indicate better performance. All these metrics have been computed using the *pyirf* v0.4.0 package [21] developed by the CTA community.

B. Results

The observation of the ratio of sensitivity over the Hillas+RF method trained on standard Prod5 simulations presented in Fig. 3 demonstrates the overall superiority of the deep multi-task approach on simulated data. The γ -PhysNet DA architecture improves the sensitivity compared to the Hillas+RF method by decreasing the minimal detectable flux by a factor of 1.25 to 8 below 80 GeV. However, we observe that the addition of Poisson noise to the data slightly degrades the performance of γ -PhysNet DA, while the Hillas+RF is less affected. The energy resolution curves presented in Fig. 4 show that γ -PhysNet DA, the Hillas+RF and Hillas+RF + *Poisson noise* obtain similar results above 100 GeV for the energy regression task, while γ -PhysNet DA + *Poisson noise* underperforms above 2 TeV. However, below 100 GeV, both γ -PhysNet DA (with and without noise) clearly outperform the standard method. These results should be considered with regards to the required energy range of the LST1 being [20 GeV ; 3 TeV].

Analogously, the results for the direction regression task shown in Fig. 5 are comparable for the four models above 200 GeV, albeit the performance of γ -PhysNet DA slightly degrades with the addition of noise. Again, γ -PhysNet DA (with and without noise) obtains a better performance for lower energies, highlighting its greater ability to handle difficult cases, and to extend the analysis sensitivity compared to the Hillas+RF method.

From a general perspective, we observe that the performance obtained by γ -PhysNet DA slightly degrade with the addition of Poisson noise to the simulation, especially above 2 TeV. This emphasizes the learning bias introduced into the models by the simulated data. On the contrary, the Hillas+RF method achieves similar performance with and without noise, highlighting the robustness of this method to the NSB level. Besides, we also observe that the deep multi-task approach constantly outperforms the Hillas+RF method for lower energies. However, it is important to stress that transferring the good performance obtained on these simulated data to real data is challenging. The improvement brought by the γ -PhysNet architecture over the standard Hillas+RF method then has to be confirmed on real data.

VI. REAL DATA ANALYSIS

We analyze in this section the two observations of the 2020 LST1 Crab campaign described in Section IV-B with the four models: γ -PhysNet DA, γ -PhysNet DA + *Poisson noise*,

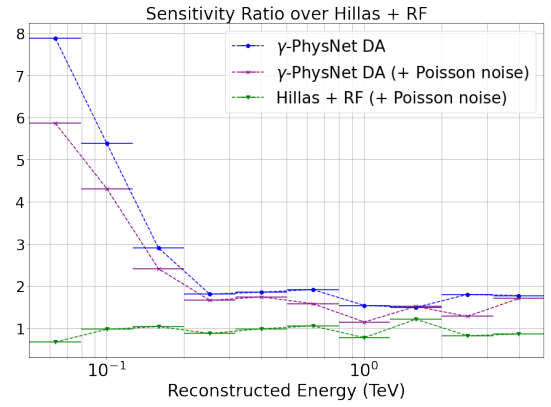


Fig. 3. Ratio of sensitivity between γ -PhysNet and the Hillas+RF method.

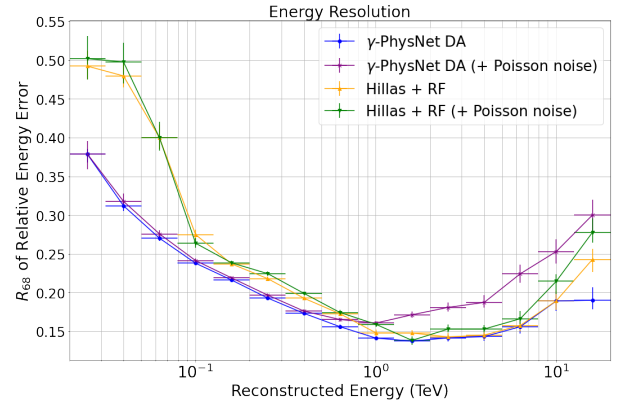


Fig. 4. Energy resolution of γ -PhysNet and the Hillas+RF method (lower is better).

the Hillas+RF method and the Hillas+RF method + *Poisson noise*. As it follows the standard approach, the analysis with the Hillas+RF method serves as a baseline for performance comparison.

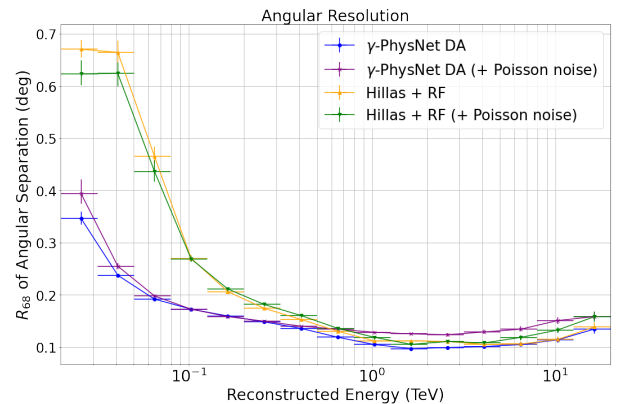


Fig. 5. Angular resolution of γ -PhysNet and the Hillas+RF method (lower is better).

A. Gamma-Ray Event Detection

To select in both ON and OFF runs the events detected as gamma, we define gammaness cuts, the threshold of the gamma class output of the model. We observe that gammaness distributions are different depending on the analysis method and on the run analyzed. Therefore, to allow for a fair comparison of the detection performance between the four models, we define the following procedure:

- 1) a unique gammaness cut for all energy bins is defined for the Hillas+RF method for the ON run,
- 2) for each energy bin, the gammaness cut of the three other models – γ -PhysNet DA, γ -PhysNet DA + *Poisson noise* and Hillas+RF + *Poisson noise* – for the ON run is tuned so that the level of background identified as signal is equivalent for all the methods,
- 3) for the OFF run, the expected background (detected in the ON run by the baseline Hillas+RF) in every energy bin is first normalized to take into account the difference in acquisition duration and rate between the ON and the OFF runs. Then, the previous procedure is applied for background matching.
- 4) Finally, because the acquisition conditions are different between both runs, we need to globally normalize the number of events detected in the OFF run, relying on the background level. Thanks to the previous steps, the ON-OFF normalization coefficient is similar for every analysis model.

To ensure that the results obtained are reproducible, and not dependent on a particular gammaness cut chosen for the baseline, we repeat the analysis with 7 different cuts within the range [0.5 ; 0.8], exhibiting a common trend for the relative performance of the four models. For clarity, we only present the results obtained with a gammaness cut of 0.65.

In both runs, the background is estimated by counting the events classified as gamma rays in an area ranging from 0.32 to 0.55° from the telescope pointing direction. This area is outside the source location, even in the ON run.

The theta squared curves shown in Fig. 6 represent the distributions of squared angular separation between the source location and the reconstructed direction. We observe that the gamma-ray event detection procedure leads to a similar level of background events (between 0.1 and 0.3 deg^2) in the ON and the OFF runs for all the models.

B. Gamma-Ray Source Detection

The last step of the analysis is the detection of the source itself in the ON run, based on the gamma-ray events detected in the previous step. As classically done in the field, we apply a final cut on the angular separation to the detected gamma-ray events in both runs. More precisely, events with an angular separation greater than 0.22° are discarded. Indeed, in the analysis carried out in this paper, the source location is known and events with an angular separation greater than 0.22° are more likely to be badly reconstructed.

From the surviving gamma-ray events, the following quantities are computed:

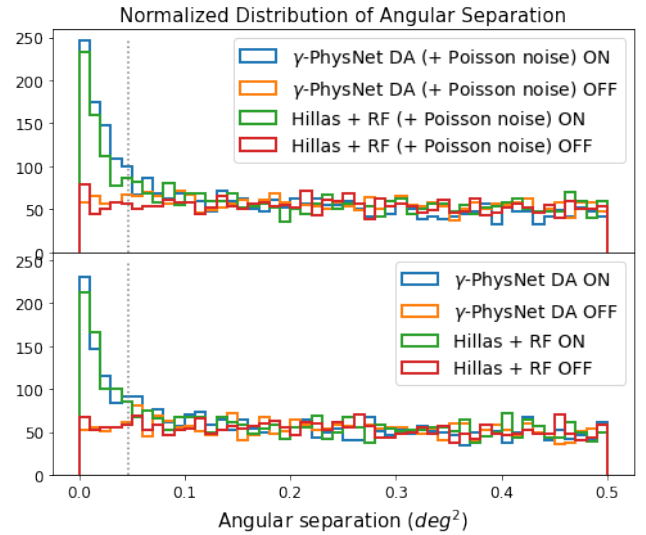


Fig. 6. Gamma-ray event distribution in both observations as a function of the squared angular separation between the reconstructed and the true directions.

- the number of **background** events classified as gamma ray in the OFF run,
- the total gamma-ray **excess** in the ON run. Because analysis models are not perfect, some particles classified as gamma ray in the ON run may be background events (protons and electrons). The excess is computed by subtracting the number of background events to the number of events detected in the ON run.
- The statistical **significance** of the source detection.

As shown in Table I, the Crab nebula detection by all models is clear with a significance ranging from 11.9 to 14.3 σ when the required value for a source detection is 5 σ . The γ -PhysNet DA models achieve similar or better performance than the Hillas+RF ones, highlighting the value of neural networks for gamma-ray astronomy.

Besides, γ -PhysNet DA + *Poisson noise* model obtains the best excess and significance, emphasizing the importance of simulating as accurately as possible the observation conditions. The addition of Poisson noise to the training data to reduce the NSB level difference with the real data helps lower the learning bias.

This is confirmed by the distribution of the excess of gamma-ray events per energy bin presented for all methods in Fig. 7. As for simulated data, the gain brought by γ -PhysNet DA + *Poisson noise* architecture is more important at low energies (thanks to a greater excess for similar backgrounds). This is especially interesting for the study of extragalactic sources and transient phenomena. Such results make it possible to exploit the full potential of the built instrument. However, these results also show a lower improvement brought by the γ -PhysNet architecture over the Hillas+RF method on real data than on simulations. This was expected as neural networks, such as γ -PhysNet, process directly the pixels of the images, while the Hillas+RF relies on extracted parameters

TABLE I
CRAB NEBULA SOURCE DETECTION BY BOTH MODELS

| | Excess | Significance | Background |
|---|------------|---------------------------------|------------|
| Hillas+RF | 379 | 12.0 σ | 308 |
| Hillas+RF + Poisson noise | 376 | 11.9 σ | 305 |
| γ -PhysNet DA | 395 | 12.5 σ | 302 |
| γ -PhysNet DA + Poisson noise | 476 | 14.3 σ | 317 |

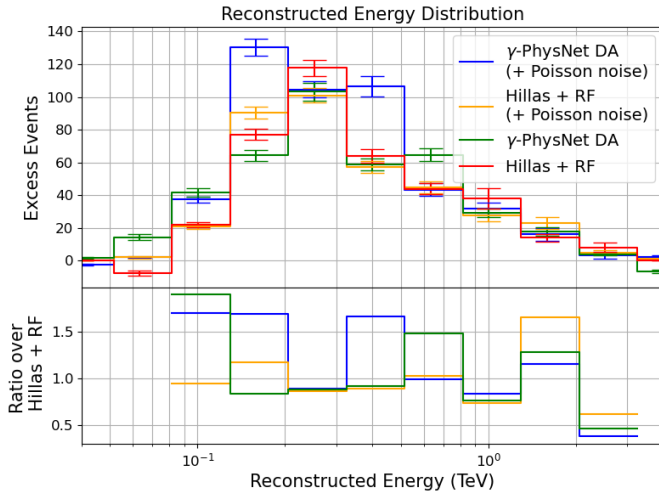


Fig. 7. Excess of detected gamma rays per energy bin. The lower part of the plot represents the ratio of gamma-ray excess per energy bin detected by the models over the one detected by the baseline Hillas+RF method.

from denoised images, thus being more robust to differences between simulations and real data. This robustness of the Hillas+RF method is underlined by the similar results obtained with or without the addition of noise to the training data.

VII. CONCLUSION

In this paper, we propose the first ever full-event reconstruction from IACT real data with deep learning. With a deep multi-task architecture, γ -PhysNet DA, we achieve a clear detection of the Crab nebula with a statistical significance of 14.3 σ , outperforming the standard Hillas+RF method. To obtain this result it was necessary to adapt the simulated data used to train the model by adding noise to the images, and thus taking into account the NSB level difference with the real data. A future work will focus on finding a more elegant solution to adapt to the NSB level that varies over the observation and over the field of view, and more generally to adapt to the data variability problem (e.g., using domain adaptation techniques).

Besides, the high-level analysis part that was beyond the focus on this study will be strengthened in a following study. In particular, an energy-dependent study of the Crab Nebula shall be made, introducing energy-dependent gammaness cuts for the baseline and deriving the spectrum of the source and comparing it to the spectrum measured by other instruments.

REFERENCES

- [1] Y. Inoue *et al.*, “Development of the camera for the large size telescopes of the Cherenkov Telescope Array,” in *Advances in Optical and Mechanical Technologies for Telescopes and Instrumentation*, R. Navarro, C. R. Cunningham, and A. A. Barto, Eds., vol. 9151, International Society for Optics and Photonics. SPIE, 2014, pp. 1384 – 1391. [Online]. Available: <https://doi.org/10.1117/12.2054619>
- [2] A. Hillas, “Cherenkov light images of eas produced by primary gamma,” in *International Cosmic Ray Conference*, vol. 3, 1985.
- [3] A. Fiasson, F. Dubois, G. Lamanna, J. Masbou, and S. Rosier-Lees, “Optimization of multivariate analysis for IACT stereoscopic systems,” *Astroparticle Physics*, vol. 34, 2010.
- [4] M. de Naurois and L. Rolland, “A high performance likelihood reconstruction of γ -rays for imaging atmospheric Cherenkov telescopes,” *Astroparticle Physics*, vol. 32, no. 5, pp. 231–252, 2009.
- [5] R. Parsons, T. Murach, and M. Gajdus, “Hess ii data analysis with impact,” in *The 34th International Cosmic Ray Conference*, vol. 236. SISSA Medialab, 2016, p. 826.
- [6] I. Shilon *et al.*, “Application of deep learning methods to analysis of imaging atmospheric cherenkov telescopes data,” *Astroparticle Physics*, vol. 105, pp. 44–53, 2019.
- [7] R. D. Parsons and S. Ohm, “Background rejection in atmospheric Cherenkov telescopes using recurrent convolutional neural networks,” *European Physical Journal C*, vol. 80, no. 5, p. 363, May 2020.
- [8] D. Nieto Castaño, A. Brill, B. Kim, T. B. Humensky, and C. Consortium, “Exploring deep learning as an event classification method for the Cherenkov Telescope Array,” in *35th International Cosmic Ray Conference (ICRC2017)*, ser. International Cosmic Ray Conference, vol. 301, Jan. 2017, p. 809.
- [9] S. Mangano, C. Delgado, M. I. Bernardos, M. Lallena, and J. J. Rodríguez Vázquez, “Extracting gamma-ray information from images with convolutional neural network methods on simulated cherenkov telescope array data,” in *IAPR Workshop on Artificial Neural Networks in Pattern Recognition*. Springer, 2018, pp. 243–254.
- [10] D. Nieto Castaño, T. Miener, A. Brill, J. L. Contreras, T. B. Humensky, and R. Mukherjee, “Reconstruction of iact events using deep learning techniques with clearn,” in *ADASS XXX*, ser. Astronomical Society of the Pacific Conference Series, 2020, in press.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [12] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.
- [13] M. Jacquemont, T. Vuillaume, A. Benoit, G. Maurin, and P. Lambert, “Multi-task architecture with attention for imaging atmospheric cherenkov telescope data analysis,” in *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, INSTICC. SciTePress, 2021.
- [14] M. Jacquemont *et al.*, “Indexed operations for non-rectangular lattices applied to convolutional neural networks,” in *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 5: VISAPP*, INSTICC. SciTePress, 2019, pp. 362–371.
- [15] J. Sun, F. Darbeha, M. Zaidi, and B. Wang, “Saunet: Shape attentive u-net for interpretable medical image segmentation,” *arXiv preprint arXiv:2001.07645*, 2020.
- [16] D. Heck, J. Knapp, J. Capdevielle, G. Schatz, T. Thouw *et al.*, “CORSIKA: A monte carlo code to simulate extensive air showers,” *Report fzka*, vol. 6019, no. 11, 1998.
- [17] K. Bernlöhr, “Simulation of imaging atmospheric Cherenkov telescopes with CORSIKA and sim_telarray,” *Astroparticle Physics*, vol. 30, no. 3, pp. 149–158, 2008.
- [18] R. Lopez-Coto *et al.* for CTA LST project, “lsthchain: An analysis pipeline for LST-1, the first prototype Large-Sized Telescope of CTA,” in *Astronomical Data Analysis Software and Systems XXX*, ser. Astronomical Society of the Pacific Conference Series, 2021.
- [19] D. Raikov, “On the decomposition of poisson laws,” in *Dokl. Akad. Nauk SSSR*, vol. 14, no. 1, 1937, pp. 9–12.
- [20] T.-P. Li and Y.-Q. Ma, “Analysis methods for results in gamma-ray astronomy,” *The Astrophysical Journal*, vol. 272, pp. 317–324, 1983.
- [21] M. Nöthe *et al.* (2020, Nov.) pyirf. [Online]. Available: <https://doi.org/10.5281/zenodo.4304466>