Cherenkov Shower Detection Combining Probability Distributions from Convolutional Neural Networks

Mauricio Araya,[1,2] Francisco Casas,[2] and Rodrigo Cáceres[2]
[1]*Centro Científico Tecnológico de Valparaíso, Valparaíso, Chile;* `mauricio.araya@usm.cl`

[2]*Departamento de Informática, UTFSM, Valparaíso, Chile*

Abstract. Ground-based gamma-ray observatories such as the Cherenkov Telescope Array presents new challenges for astronomical data analysis. The dynamics of the atmosphere and the complexity of Cherenkov shower are two uncertainty sources that needed to be embraced rather than corrected. As each telescope only has access to a separated patch, the partial information of each one has to be combined. For instance, when blots can be identified on the images, the application of Hillas parameters allows to identify the approximate direction to the projection's center. This information can be combined for several telescopes using stereoscopic reconstruction to converge on a single point. The limitation of this technique however is that it performs regressions to a predefined blot shapes, not using all the information contained in the images. Thus, deep learning techniques based on Convolutional Neural Networks have been applied with promising results. However, they rely on very large networks that process all the telescope images at once, which might not scale properly when dealing with large arrays. We propose to run several separate instances of an smaller network for each telescope, but that are able to retrieve a probability distribution instead of approximate coordinates for the sought point. This probability distribution can be arranged by the network so it can express certainty about the direction or the distance to the center of the projection separately. The distributions retrieved by all the telescopes can be combined to get a final probability distribution. Preliminary results shows the viability of this approach to identify the center and assign a confidence value to the result.

1. The Gamma-ray Reconstruction Problem

When a very-high energy gamma-ray photon hits the Earth's atmosphere, it produces a fast shower of particles, some of them travelling at ultra-relativistic speed, and therefore emitting Cherenkov light that can be detected by using Imaging Atmospheric Cherenkov Telescopes (IACTs) (Völk & Bernlöhr 2009). The analysis of the images produced by the telescopes, allow the reconstruction of the penetration depth in the atmosphere, its direction and energy. This will be done on a large scale and with high precision by the Cherenkov Telescope Array (CTA), to be installed in the next few years in Chile and Spain (Actis et al. 2011). The estimation of physical parameters from the images is a complex inverse problem that requires computing-intensive methods for simulations, inference and validation. The accuracy, performance and reliability of the algorithms involved in particle discrimination and gamma-ray reconstruction, are essential for science with IACTs, because the sensitivity and confidence on the whole observation depends on them. There are classical approaches for reconstructing gamma-rays from IACTs images, such as exploiting the known geometric properties of the showers (Hillas 1985), likelihood maximization of statistical models (De Naurois & Rolland 2009) or using fast monte-carlo simulations (Parsons & Hinton 2014). However, recent advances are strongly based on machine learning methods, such as random forests, boosted decision trees, and more recently, deep learning (Shilon et al. 2018; Mangano et al. 2018). The use of these data-driven techniques impose new challenges in terms of computational performance, correctness demonstration and uncertainty propagation.

2. Uncertain Multi-Observer Neural Network Assembly (UMONNA)

The general framework that we propose, called UMONNA, is formalized as follows. Let $T{=}\{t_1, t_2, ..., t_n\}$ be the set of images (and possibly additional data) received by a set of n observers, and y a value that wants to be predicted on the domain D from this images. The conventional machine learning approach would be to train the parameters W of a model f that retrieves an approximation $\hat{y}$ of y from the set T, i.e., $\hat{y} = f(T|W)$. We propose a model that retrieves a Probability Density Distribution (PDF) for y on the domain D, based on the data of each observer t_i independently. For this, let M be a PDF parametrized by Σ_i that follows a model f:

$$\Sigma_i = f(t_i|W)$$

Then, the PDF for y is:

$$M(z|\Sigma_i) \qquad \text{where } z \in D$$

This allows each observer i to retrieve wide probability distributions when their uncertainty is high, which may be the case when t_i doesn't contain enough information, this also allows them to express uncertainty in particular components of the domain D, e.g. distance but not direction. The PDFs retrieved from the observers are then merged within the domain D:

$$M^*(z) = \frac{\sqrt[n]{\prod_i M(z|\Sigma_i)}}{\int_D \sqrt[n]{\prod_i M(z|\Sigma_i)}dz}$$

The approximation $\hat{y}$ of y can be computed from the distribution directly, for example by using the MAP criterion:

$$\hat{y} = \arg\max_{z \in D} M^*(z|\Sigma_i).$$

2.1. Proof of Concept: Simple UMONNA

As a proof of concept we generated a "dummy shower": synthetic blots distributed with a clear pattern around a shower center and try to predict its position. We take spread patches of this canvas as the images that each telescope receives. We trained a Convolutional Neural Network (CNN) (Goodfellow et al. 2016) that retrieves a probability density distribution where each output of the M output neurons represents the probability of the center being at a certain angle (see Figure 1.

One of the main challenges of this idea is to choose the right loss function so that the retrieved PDFs can be evaluated correctly. When the PDF displays a high degree of certainty, this loss must be higher on errors and lower on right predictions than when it doesn't.

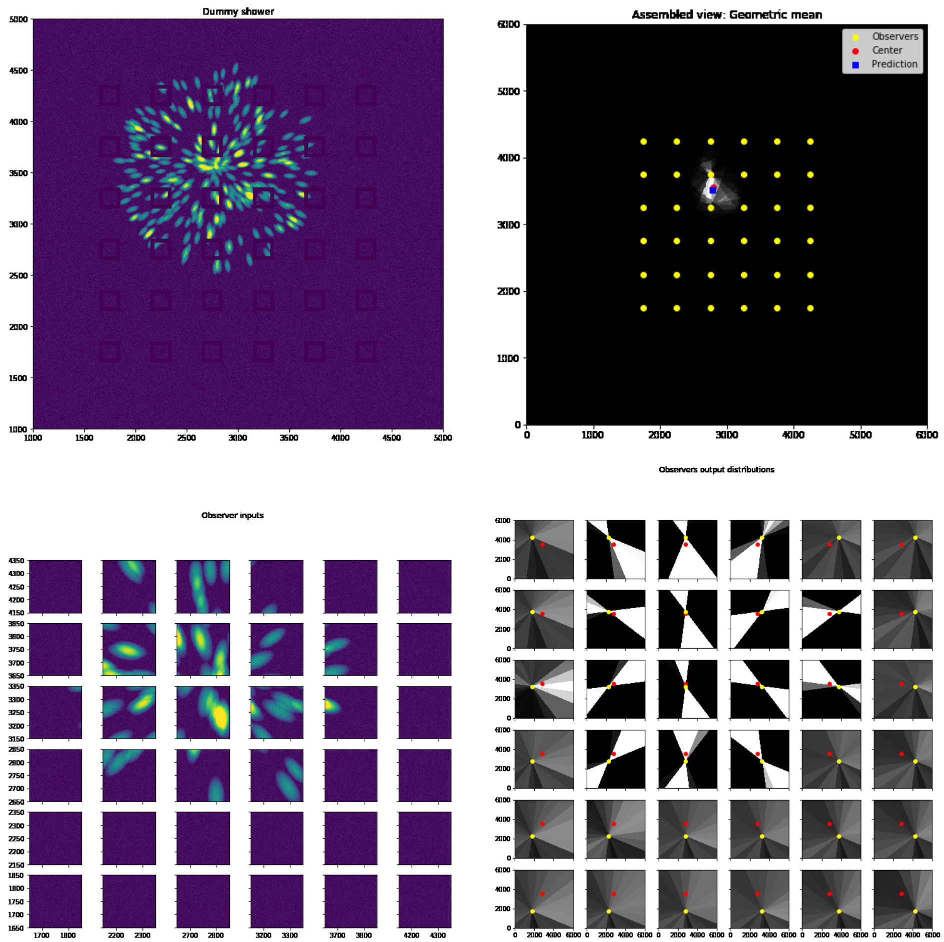

Figure 1. **Simple UMONNA**. The top-left image shows the synthetic shower over the domain D. The bottom-left image simulates the scattered view that the telescopes will observe. The bottom-right image presents the result of each network: the sought-point probability distribution based only on a single image reconstruction. At last, the top-right image shows the joint probability distribution and the maximum-a-posteriori (MAP) sought-point.

For this proof of concept, we selected a loss function that evaluates y in the output distribution to obtain an error value:

$$L(y, \Sigma_i) = \left(1 - \min\left\{1, \frac{M(y|\Sigma_i)}{C_{ap}}\right\}\right)^2$$

where C_{ap} is a limit for the PDF density to ensure that it is spread on other points of the domain D.

3. Conclusions and Future Work

The UMONNA approach separates a problem in smaller ones and provides a way to merge the results (i.e., divide and conquer strategy). This allows to work with simpler learning models, furthermore, by separating the observations, it increases the number training of samples by a factor of n. Another benefit of this approach is that it allows to measure the level of certainty of the assembly on different points of a continuous domain D. We believe that his could be used to detect simultaneous γ-ray phenomena occurrences through finding more than one local maximum in $M^*(.)$.
In terms of future work, we are beginning to train and test with simulated data using the Monte-Carlo simulators provided by CTA (Bernlöhr 2008). Also, we plan to estimate the energy and direction of the γ-rays, and not only the sought point as in this paper. On the model side, we would like to get smoother PDFs through interpolation, retrieve multivariate Gaussian distributions through deconvolutions, and include automatic derelativation for each telescope.

References

Actis, M., Agnetta, G., Aharonian, F., Akhperjanian, A., Aleksić, J., Aliu, E., Allan, D., Allekotte, I., Antico, F., Antonelli, L., et al. 2011, Experimental Astronomy, 32, 193
Bernlöhr, K. 2008, Astroparticle Physics, 30, 149
De Naurois, M., & Rolland, L. 2009, Astroparticle Physics, 32, 231
Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. 2016, Deep learning, vol. 1 (MIT press Cambridge)
Hillas, A. M. 1985
Mangano, S., Delgado, C., Bernardos, M. I., Lallena, M., Vázquez, J. J. R., Consortium, C., et al. 2018, in IAPR Workshop on Artificial Neural Networks in Pattern Recognition (Springer), 243
Parsons, R., & Hinton, J. 2014, Astroparticle physics, 56, 26
Shilon, I., Kraus, M., Büchele, M., Egberts, K., Fischer, T., Holch, T. L., Lohse, T., Schwanke, U., Steppa, C., & Funk, S. 2018, arXiv preprint arXiv:1803.10698
Völk, H. J., & Bernlöhr, K. 2009, Experimental Astronomy, 25, 173

Acknowledgments. This work has been partially funded by CONICYT PIA/Basal FB0821, CONICYT - PFCHA/Magister Nacional/2018 - folio 22182114, and DGIIP PI-L-18-13.