CLASSIFICATION METHODS FOR MAGIC TELESCOPE IMAGES ON A PIXEL-BY-PIXEL BASE

CONSTANTINO MALAGÓN[1]
1 Departamento de Ingeniería Informática, Univ. Antonio de Nebrija, 28040 Madrid, Spain.
JUAN ABEL BARRIO[2] , DANIEL NIETO [2]
2 Departamento de Física Atmica, Universidad Complutense, 28040 Madrid, Spain

Abstract:The problem of identifying gamma ray events out of charged cosmic ray background (so called hadrons) in Cherenkov telescopes is one of the key problems in VHE gamma ray astronomy. In this contribution, we present a novel approach to this problem by implementing different classifiers relying on the information of each pixel of the camera of a Cherenkov telescope, rather than using common Hillas parameter analysis. Separation between gamma-like and hadron-like events is tested by using Monte Carlo data samples of both types of events.

Keywords: Gamma Ray Astronomy, Cherenkov telescopes, gamma rays detection

1 Introduction

Since Hillas parameter analysis was developed back in 1985 to separate between gamma-like and hadron-like events as recorded by Cherenkov telescopes [1], many techniques have been used for gamma/hadron separation based on such parameters. Bock et al. [2] performed a case study for most of these techniques, to be later applied to MAGIC telescope gamma event selection. However, all these techniques might not be using the whole potential of a Cherenkov telescope, as they use Hillas parameters (second moments of image in telescope camera) as input.
In this work we propose to apply usual machine learning techniques (some of them, mentioned in [2]) to the full image recorded by a Cherenkov telescope, on a pixel-by-pixel base. We will demonstrate the method using images produced by the MAGIC telescope simulation and reconstruction package [3].
Moreover, current techniques fail to efficiently discriminate beteween primaries (gamma rays and cosmic rays) for energies of the primary roughly below 100[1] - 200 GeV (Giga

[1]This is the so-called *software* energy threshold of a telescope

ElectronVolts). And it is precisely below that energy that very important physics results are expected.
Possible advantages of this approach are the use of the full information in the camera, and a more natural way to treat fluctuations in the image, thus permitting a relaxation of the image cleaning and a likely reduction of the telescope software threshold. Some inconvenients about this approach come from the fact that certain effects very difficult to simulate (like bright stars in the Field of View or non-uniformities in atmosphere), can play a certain role in a pixel-by-pixel analysis, while only a minor effect on a Hillas parameter analysis.

2 Data sample

For this study, we have used gamma and proton events (as the latter represents the majority of hadronic cosmic rays) simulated with Corsika code [6], plus MAGIC Reflector and Camera reconstruction standard software. Each event consists of an image based on the calibrated photoelectron content in each pixel of the MAGIC telescope camera. The pixels whose signal is likely to originate from NSB or electronic noise are removed from the image using the so-called image cleaning procedure [7], with 10 photoelectron threshold for core pixels and 5 photoelectron threshold for boundary pixels.
Gamma and proton samples consist of 28750 events. Image total photoelectron spectrum of gamma sample resembles that of typical cosmic sources. Corresponding spectrum of proton sample is forced with similar slope to avoid biasing the selection procedure.

3 Experimental evaluation

In what follows, we will briefly enumerate the different classifiers used in the experiment, composed by individual or ensemble classifiers.
We have used three individualclassifiers for our experiments: decision trees using Quinlan's C4.5 algorithm [8]), a multilayer perceptron [9] and K Nearest Neighbour algorithm [10], with K=11 and euclidean distance metric. Regarding ensemble classifiers, which will only be applied to decision trees classifier, we will deal with multiple classifier systems generally described as voting classification algorithms, i.e., techniques in which we use several individual classifiers that output a particular prediction or label for each of the examples of the test data set. These predictions are then combined to produce a single output, the output of the ensemble, by majority voting decision. We used two different voting algorithms: Adaboost algorithm [11]

implemented in Weka [12] and Bagging, the classical ensemble method developed by Breiman [13].

4 Results

All these machine learning methods have been fed with above described gamma and proton samples, using a typical holdout validation method (two thirds of them for training and the rest for testing purposes).
Results from the different methods, presented in terms of ROCc (Receiver Operating Characteristic curves) for different classification methods, are shown in figures below: vertical axis shows gamma acceptance while horizontal one represents hadron acceptance. Figure 1 shows the results when classifying the whole data set (i.e. without distinction on the energy of the event). Boosting trees show results comparable with those in Bock et al. [2], graphically displayed using point style.
Figure 2 show the ROC curve for events with energies lower than 100 GeV, which is

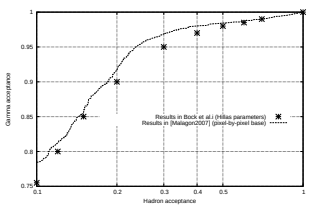

Figure 1: ROCc for the whole data sample

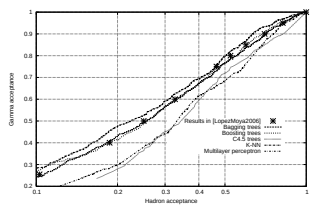

Figure 2: ROCc for events with energies lower than 100 GeV

compared with results in López [14] (equivalent to results in Bock et al. [2], but using MAGIC telescope real data instead of Motecarlo simulations for hadron background estimation), graphically displayed using point style.
As discussed above, this energy threshold is a key point in the discussion of the experimental results. When approaching a threshold in the low GeV domain, typically few hundreds of GeV, one is confronted with a new background situation. Information of gamma-ray images provided by Hillas parameters is degraded, and difficulties arise in gamma selection process, as gamma-ray images become similar to hadron images.

5 Conclusions and outlook

Separation between simulated gamma-like and hadron-like events (as reconstructed by the MAGIC Cherenkov Telescope) is performed using several machine learning

techniques applied to pixel-by-pixel defined images. Both ensembles of classification trees and K Nearest Neighbours show similar performance as for Hillas parameters defined images [2] (without any restriction on the energy of the primaries).
These classifiers were also discussed in terms of event energies, showing promising results for events with energies below 100 GeV. This is an unexplored energy region which will be a major issue in future explorations.

Acknowledgements: We want to thank the MAGIC collaboration for kindly providing us with the simulated gamma and proton sample that we have used to test the method. We also want to thank R. Bock for fruitful discussions.

References

[1] Hillas A. M.: Cherenkov light emission of extensive air showers produced by primary gamma rays and by nuclei, Proc. 19th ICRC, La Jolla, 4, 445 ,1985.
[2] Bock, R.K. et al.: Methods for multidimensional event classification: a case study using images from a Cherenkov gamma-ray telescope, Nucl. Instrum. MethA516, 511-528 , 2004.
[3] Bretz, T. (MAGIC Coll.): Procs. 29th ICRC, Pune, India, 2005.
[4] Gaug, M., et al (MAGIC Coll.): Procs. 29th ICRC, Pune, India, 2005.
[5] Le Bohec, S. et al.: A new analysis method for very high definition Imaging Atmospheric Cherenkov Telescopes as applied to the CAT telescope. Nucl.Instrum.Meth. A416, 425-437,1998.
[6] D. Heck et al., CORSIKA: A Monte Carlo Code to Simulate Extensive Air Showers, Forschungszentrum Karlsruhe Report FZKA 6019, 1998.
[7] D.J.Fegan: Gamma/hadron separation at TeV energies, Topical Review, J.Phys.G: Nucl. Part. Phys. 23, 1013, 1997.
[8] Quinlan, R.: Induction of Decision Trees. Machine Learning,1, 81-106, 1986.
[9] Rumelhart D., McClelland J. And the PDP Research Group. Learning internal representations by error propagation. Pararell Distributed processing: Explorations in the microestructure of cognition. 1-3. MIT Press, 1986.
[10] Aha, D., and Kibler, D.: Instance-based learning algorithms. Machine Learning, 6, 37-66, 1991.
[11] Freund, Y., Schapire, R. E.: Experiments with a new boosting algorithm, In Proc. 13th International Conference on Machine Learning. San Francisco: Morgan Kaufmann. 148-146, 1996.
[12] Witten, Ian H., Eibe Frank.: Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations Morgan Kaufmann, San Francisco, 2000.
[13] Breiman, L.: Bagging predictors. Machine Learning 24, 123-140, 1996.
[14] Lopez M., PhD. thesis, UCM, Madrid 2006