New method for Gamma/Hadron separation in HAWC using neural networks

T. Capistrán[a], I. Torres[a], L. Altamirano[a] and for the HAWC Collaboration[b]

[a]Instituto Nacional de Astrofísica, Óptica y Electrónica, Luis Enrique Erro 1, Tonantzintla, Puebla 72840, México
[b]For a complete author list, see www.hawc-observatory.org/collaboration/icrc2015.php

tcapistran@inaoep.mx, ibrahim@inaoep.mx, robles@inaoep.mx

Abstract

The High Altitude Water Cherenkov (HAWC) gamma-ray observatory is located at an altitude of 4100 meters in Sierra Negra, Puebla, Mexico. HAWC is an air shower array of 300 water Cherenkov detectors (WCD's), each with 4 photomultiplier tubes (PMTs). Because the observatory is sensitive to air showers produced by cosmic rays and gamma rays, one of the main tasks in the analysis of gamma-ray sources is gamma/hadron separation for the suppression of the cosmic-ray background. Currently, HAWC uses a method called compactness for the separation. This method divides the data into 10 bins that depend on the number of PMTs in each event, and each bin has its own value cut. In this work we present a new method which depends continuously on the number of PMTs in the event instead of binning, and therefore uses a single cut for gamma/hadron separation. The method uses a Feedforward Multilayer Perceptron net (MLP) fed with five characteristics of the air shower to create a single output value. We used simulated cosmic-ray and gamma-ray events to find the optimal cut and then applied the technique to data from the Crab Nebula. This new method is tuned on MC and predicts better gamma/hadron separation than the existing one. Preliminary tests on the Crab data are consistent with such an improvement, but in future work need to be compared with the full implementation of compactness with selection criteria tuned for each of the data bins.

Introduction

The High Altitude Water Cherenkov (HAWC) gamma-ray observatory is composed of 300 water Cherenkov detector (WCD). On the bottom of each WCD there are 4 photomultiplier tubes (PMTs) that detect the Cherenkov light. This light is produced by secondary particles in air shower generated by the interaction between atmosphere and primary particle (as for example gammas rays, protons, among other particles). The rate of cosmic rays (CR) is bigger than the gamma rays (GR) so it is critical to find a technique to remove the CR without losing the signals of GR.

Currently, HAWC has a method called compactness for distinguishing those primary particles. For doing this, the data is divided into 10 bins (see Table 1) depending on $nHit$, that is the number of PMTs that have a signal in the event. The compactness depends upon the charge distribution deposited by the secondary particles of the shower on PMTs of the array. In this work, a new method is presented, using a Neural Network (NN) for the gamma/hadron separation without dividing the data into bins. Five characteristics are computed for feeding a NN that computes a value (θ_{NN}) to distinguish between CR and GR. Another method in development can be found in [1].

Training stage

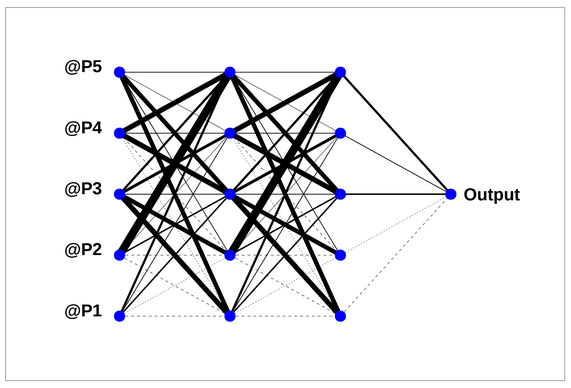

Figure 1: Architecture of NN with 5 neurons as inputs, two hidden layers with 5 neurons and one neuron as output. The width of each connection line between neurons is proportional to the weight of the NN.

The NN used in this work is a Feedforward Multilayer Perceptron [2] with an architecture of 5-5-5-1 (see Figure 1). A target value is defined as 1 for gamma ray event and 0 for hadron event.
We used proton as hadron. The conditions for selecting training events for each set are:

- The difference between the core reconstruction and simulation does exceed 5 m.
- The core falls inside the HAWC array.
- The event with nHit between 30 and 1200.

The characteristic inputs are:

- $P1 = nHit$ is the number of PMTs with at least one photoelectron (PE).
- $P2 = DisMax$ that is the largest distance between any of the pair of tubes passing the next selection: first all the PMTs in the event are sorted by their PEs detected and we summed this value for each PMT from higher to lower until the sum is less that $(SumPE - MaxPE) * k(nHit)$, where $MaxPE$ is the number of PEs in any PMT in the event, and "k" is a factor that depends linearly of nHit, the PMTs involved in that sum are the selected ones.
- $P3 = Log_{10}(\frac{nHit}{\sum_n PE_i * R_{PE_i}})$ where $R_{PE_i} > 30\ m$.
- $P4 = CxPE_{30}/MaxPE$ where $CxPE_{30}$ is the maximum charge outside a exclusion radius of 30 m in the event.
- $P5 = Log_{10}(|CxPE_{30} * R_{CxPE_{30}} - PE_{maxint} * R_{PE_{maxint}}|)$ where $R_{CxPE_{30}} > 30\ m$, and $R_{PE_{maxint}} < 30\ m$

The specifications for training are:

- Stochastic minimization as learning method.
- 500 epoch.

The result of the NN is shown in Figure 2. One threshold is defined for distinguishing between primary particles, this is θ_{NN}.

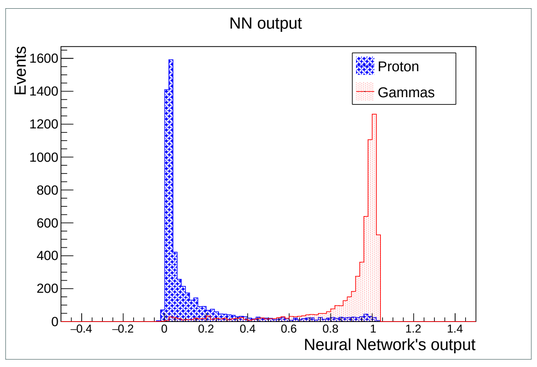

Figure 2: The histogram of NN's outputs for gammas and hadrons in the learning stage. The majority of gamma events have an output close to one, and protons are close to 0

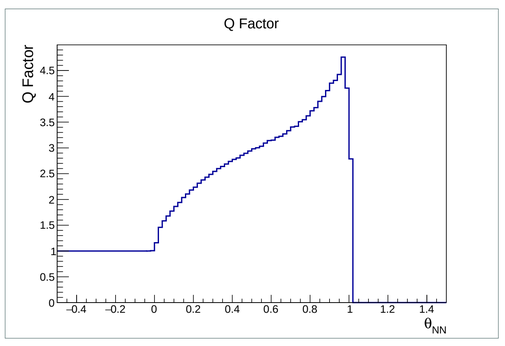

Figure 3: The Q Factor of NN's outputs. The largest Q factor is at 4.76 when the output threshold is around 0.98.

To find the θ_{NN}, we used the Q factor.

$$\text{Q factor} = \frac{\epsilon_{gamma}}{\sqrt{\epsilon_{hadron}}}$$

Where
ϵ_{gamma} gamma efficiency
ϵ_{hadron} hadron efficiency
The Q value estimates the factor by which the significance will be increased by the classification.
In the figure 3 shows the highest values are close to 0.98. After some analysis we found that optimal cut is 0.96.

Testing stage

Simulation

Bin	nHit range	θ_c
-1	30 - 54	-
0	55 - 87	4.6
1	88 - 138	6.3
2	139 - 216	9.8
3	217 - 323	12.7
4	324 - 457	17.6
5	458 - 606	19.5
6	607 - 754	18.5
7	755 - 889	17.1
8	890 - 1000	15.0
9	1001 - 1200	12.4

Table 1: nHit range and gamma/hadron cut in each bin for HAWC-300, θ_c is the compactness cut value.

For comparing the two methods we use the bin called "total" is computed using all events from bin 0 to bin 9. The results are shown in Figure 4 where we can see that for the Q Factor the NN has a better result than using the compactness method.

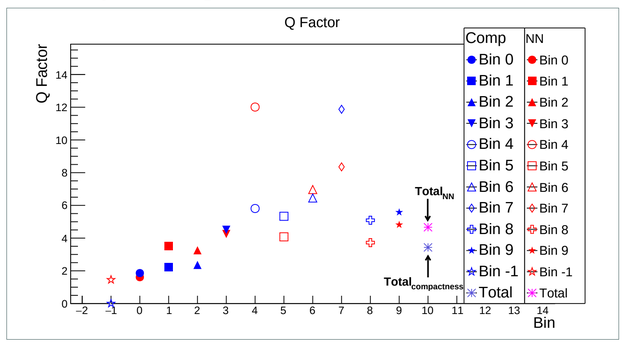

Figure 4: The Q factor is calculated for each bin and the total (bin 0 to 9) with $\theta_{NN} = 0.96$. This shows that for the Q factor in some bins, the NN is better than compactness but for others does not.

Parameter	NN	compactness	Increase (%)
Q Factor	4.663	3.432	**35.889**
gamma efficiency	0.606	0.536	**13.129**
hadron efficiency	0.017	0.024	**-30.693**

Table 2: Difference between methods with simulation.

Data

θ_c	NKG	Gauss
10.0	3.4706	4.4649
12.0	4.3142	4.4703
14.0	**5.2777**	**4.6895**
16.0	3.9327	4.3406
18.0	4.3170	4.3613

Table 3: Significance using the compactness variable with a single cut value for all bins.

θ_{NN}	NKG	Gauss
0.92	5.8842	4.9889
0.94	5.7042	5.4144
0.96	**5.9217**	**5.5096**
0.98	3.7534	4.6703
1.00	4.0977	3.1792

Table 4: Significance using NN Vs NN threshold.

Method	NKG	Gauss
compactness	5.2777	4.6895
NN	5.9217	5.5096
Increase (%)	**12.202**	**17.488**

Table 5: Difference between methods with data.

Conclusions

In this work, we propose a new method for gamma/hadron separation that used a Multilayer Perceptron fed with 5 characteristics. The NN's output is continuous and has a value targeting 1 for gamma events and 0 for hadron events. In the analysis, we found an optimal cut value for the NN output $\theta_{NN} = 0.96$. With this value the NN has better performance than compactness. The Q Factor increases approximately 36%, because the gamma efficiency increased about 13% and a decrease of 30% in hadron efficiency. In the case of data we also obtained a better significance using NN instead of a simplified version of compactness where the compactness cut was constrained to be the same for all nHit bins. In future work we will compare with the full compactness implementation.

References

[1] **HAWC** Collaboration, Z. Hampel, *Gamma Hadron Separation using Pairwise compactness Method with HAWC*, in *Proc. 34th ICRC*, (The Hague, The Netherlands), August, 2015.

[2] Christophe Delaere, *class TMultiLayerPerceptron*, 2013 (accessed January, 2014). http://root.cern.ch/root/html/TMultiLayerPerceptron.html.

[3] V. Grabski, L. Nellen, and A. Chilingarian, *Gamma/hadron separation study for the HAWC detector on the basis of the multidimensional feature space using non parametric approach*, in *Proc. 32th ICRC*, vol. 9, (Beijing, China), p. 207, August, 2011.

[4] T. Fawcett, *An introduction to ROC analysis*, *Pattern Recognition Letters* **27** (2006) 861–874.

[5] **HAWC** Collaboration, J. Pretz, *Enabling Fast Efficient Crab Verification*, tech. rep., internal Document, 2014.

[6] **HAWC** Collaboration, F. S. Greus, *Observations of the Crab Nebula with Early HAWC Data*, in *Proc. 34th ICRC*, (The Hague, The Netherlands), August, 2015.

[7] C. M. BISHOP, *Neural Networks for Pattern Recognition*. Oxford University Press, Inc. New York, NY, USA, 1995.

Acknowledgements

We acknowledge the support from: the US National Science Foundation (NSF); the US Department of Energy Office of High-Energy Physics; the Laboratory Directed Research and Development (LDRD) program of Los Alamos National Laboratory; Consejo Nacional de Ciencia y Tecnología (CONACyT), Mexico (grants 239762, 260378, 55155, 105666, 122331, 132197, 167281, 167733); Red de Física de Altas Energías, Mexico; DGAPA-UNAM (grants IG100414-3, IN108713, IN121309, IN115409, IN111315); VIEP-BUAP (grant 161-EXC-2011); the University of Wisconsin Alumni Research Foundation; the Institute of Geophysics, Planetary Physics, and Signatures at Los Alamos National Laboratory; the Luc Binette Foundation UNAM Postdoctoral Fellowship program.