Gamma Hadron Discrimination with a Neural Network in the ARGO-YBJ Experiment

Presented by S. Bussino[1] for the ARGO-YBJ Collaboration
[1]*Dipartimento di Fisica and INFN, Universitá di Roma Tre, I-00146 Roma, Italy*

Abstract

The structure of a neural network developed for the gamma hadron separation in the ARGO-YBJ detector is presented. The discrimination power in the full ARGO-YBJ energy range is shown in detail and the improvement in the detector sensitivity is also discussed.

1 Introduction

The ARGO-YBJ detector (Abbrescia et al., 1996), described elsewhere at this conference (D'Ettorre et al., 1999), is a full coverage $71 \times 74\ m^2$ RPC carpet that will be installed at YanBaJing Cosmic Ray Laboratory (Tibet, China) at an altitude of $4300\ m$ a.s.l. Due to the high altitude of the site and to the high granularity of the carpet, made of 14040 pads $56 \times 62\ cm^2$ each (dead space $\simeq 5\ \%$), the electronic shower image can be used to discriminate between gamma and proton initiated showers. In order to develop a neural alghoritm able to perform such a discrimination more then $7 \cdot 10^5$ showers, generated using the CORSIKA package, were propagated through the detector and the neural net was structured around the peculiarities of each class of showers. The neural net performances have been studied for different sets of simulated data and the results lead to a relevant increase in the sensitivity of the detector.

2 The Monte Carlo data sample

The events have been generated by using CORSIKA code 5.61 (Heck, 1998) which provides a complete simulation of the shower development in the earth's atmosphere. The electromagnetic part of the shower simulation is realized by the EGS4 code (Nelson, Hirayama and Rogers, 1985), while for the adronic component the data have been processed by VENUS code (Werner, 1993) for the high energy hadronic interactions and by GHEISHA code (Fesefeldt, 1985) for the low energy hadronic interactions (Knapp, Heck and Schatz, 1996). The data have been generated in the energy range $100\ GeV - 10\ TeV$ with an energy distribution given by:

$$N(E)dE = N_0\, E^{-2.7} dE \tag{1}$$

Data have been generated in different energy intervals, as shown in table 1. The Monte Carlo data have been processed through a code simulating the ARGO-YBJ detector, taking into account the energy loss in the material, the ionization process in the RPC gas and the RPC efficiency ($94\ \%$) as measured during the test at YBJ (Mari et al., 1999). Also the background has been simulated and a basic trigger logic has been implemented, requiring at least 25 signals inside the trigger time window. For vertical showers simulated with the core at the center of the apparatus, the number of events passing the trigger logic are reported in table 1.

Mean Primary C.R. Energy	Number of generated showers		Number of detected events	
(Tev)	*gamma*	*proton*	*gamma*	*proton*
$0.100 < E < 0.500$	360000	360000	37000	6000
$0.500 < E < 1$	12000	12000	11000	10200
$1 < E < 3$	5000	5000	4700	4200
$3 < E < 10$	770	770	750	735

Table 1: Number of generated showers in different window energy of the primary particle.

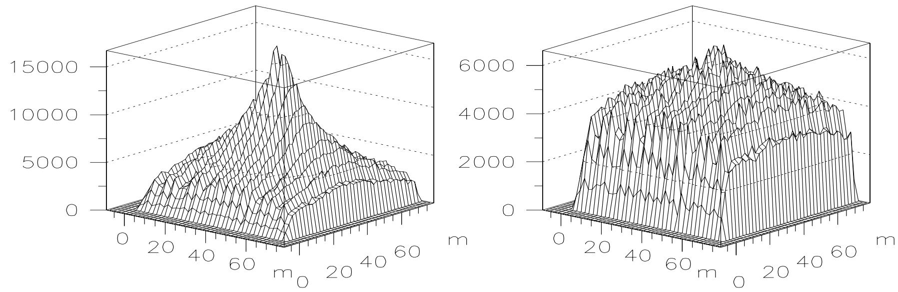

Figure 1: The hits distribution in the detector, for vertical events with the core at the centre of the carpet, for gamma (left) and protons (rights). Data are averaged over all the generated events.

In order to plan the neural network architecture, we must know the general behavior of gamma and hadron events, as a function of the observables detected in the apparatus. The two classes of showers can be distinguished by the different radial profile and from the presence, in the hadronic showers, of local fluctuations induced by small electromagnetic subshowers produced by neutral pion decay. The distribution of the hits number as a function of the coordinates is shown in fig. 1: the two classes of events, averaged over the whole energy range and over all the generated events, are clearly different. The fluctuations too are used as a tool in distinguishing gamma- from hadron-induced showers by evaluating the quantity $F(x, y)$ defined as:

$$F(x,y) \;=\frac{\mid (N_{hits} - < N_{hits} > \mid}{\sqrt{< N_{hits} >}} \tag{2}$$

3 The neural network setup

The distribution shown in the previous picture (see fig. 1) are averaged over many events: the purpose of the neural network is to distinguish between the two classes event by event, and not only when all the fluctuations are smoothed by the statistics. For this reason the network was planned looking for functions of observables which can discriminate the two classes and then the net architecture was implemented in such a way to combine the discrimination power of these functions to obtain the final rejection factor. We implemented a very simple neural network, taking into account that the number of free parameters is strongly correlated to the number of Monte Carlo events needed to teach the network.

The neural network (see fig. 2) is a standard three layer perceptron. The input layer consists of four neurons, each reading four different functions of the observables: the intensity of the signal in each pad and its square, the quantity F described by eq. 2 and its local maxima. Each neuron has a number of inputs corresponding to the pad number in the detector, plus a bias signal to set the threshold. We divided the electronic image of each event in $1\ m$ wide circular coronas, centered around the core of the shower and symmetrized the neural weights within each corona: all the inputs connected to the pads belonging to the same corona have the same weight. In this way, based upon the intrinsic symmetry of the whole physical process, the weight number is reduced to a maximum of 145 (corresponding to a shower with the core centered in one corner of the detector) instead of $\simeq\ 14000$ (one for each pixel). The aim of the hidden and of the output layer is to improve the rejection factor by combining the outputs of the input layer neurons. We implemented a two neurons hidden layer, which seems a good compromise between flexibility and power. The

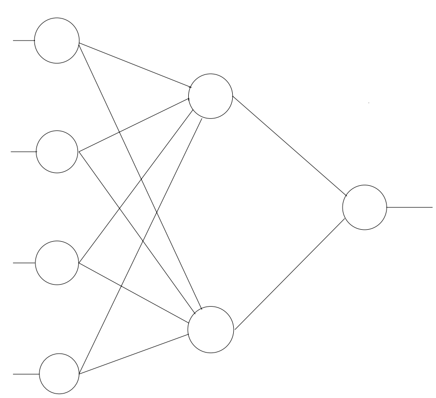

Figure 2: The setup of the network.

Hits Number	Total Number of Events	Central Core Events ϵ_γ	ϵ_h	Edge Core Events ϵ_γ	ϵ_h
$50 \leq n \leq 75$	4000	75 %	74 %	75 %	75 %
$75 \leq n \leq 100$	1800	75 %	73 %	77 %	76 %
$100 \leq n \leq 150$	1600	80 %	75 %	78 %	77 %
$150 \leq n \leq 200$	700	81 %	79 %	80 %	81 %

Table 2: Identification power ϵ for gamma and protons showers hitting in the centre of the detector (third and fourth colummns), and on the edge of the detector (last two colummns), for different pad-number intervals.

output neuron furnishes the final result and the classification of the shower (i.e. 0 for gamma induced showers and and 1 for proton induced showers).

4 Results on gamma hadron separation

The Monte Carlo events have been classified in several groups according to intervals in the number of hits recorded in the apparatus, which is an observable. For each class of events of a different hit number, we could then tune the neural network in order to maximize its discrimination power. The data sample was processed using half of the available data to teach the net and the other half to evaluate the rejection power of the net. The descrmination power of the network was initially evaluated for vertical showers with the core at the centre of the detector. The results are summarized in the third and fourth colummns of table 2, where it is shown the fraction of particle rightly identified. We tested the effect of the uncertainities on the core reconstruction, using very simples alghorithms for the core position determination, and we found that the discrimination factors ϵ are of the same order of those obtained without taking into account the error reconstruction. This result can be easily understood looking at the radial profile of the weights of the input layer neurons. The slope is smooth and the hit density is symmetric with respect to the shower core. Because of that, a small shift due to the core position reconstruction has a small influence in the network result. An example of the output of the neural network is given in figure 3, for events with a number of pads in the range $100 \div 150$.

In order to study the effects of the core position, we generated a set of events with the core on the edge of the detector. Also in this case the neural network rejection power is slightly affected, and this fact can be understood by underlining that even for events hitting the apparatus at one edge, a fraction of the hit can be detected at great distances from the center and this, at least for high populated events, can balance the loss of information due to the part of the shower hitting outside the detector. The effect of the core reconstruction leads now to a reduction of the discrimination powers ϵ of the order of few percent.

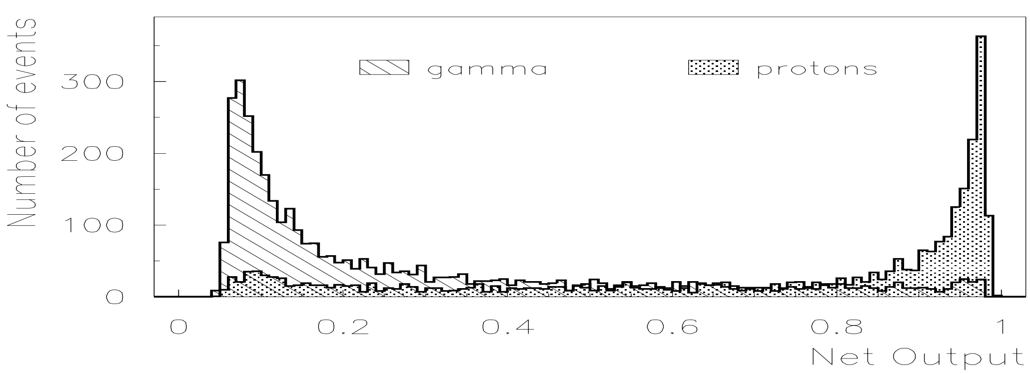

Figure 3: The answer of the neural network is shown for events with $100 < n_{pad} < 150$.

We underline, however, that the discrimination power is generally of the order of $75\ \%$ or more, a value high enough to obtain a substancial enhancement in the sensitivity of the detector. If we consider the number of data detected in a given direction, the sensitivity of the apparatus results from the ratio between the gamma flux (unknown) and the fluctuation of the hadron showers background (a well known value). In a simplified form, this ratio is given by:

$$\frac{n(\gamma)}{\sqrt{n(h)}} = \frac{\phi(\gamma)A_\gamma T_{obs}d\Omega}{\sqrt{\phi(h)A_h T_{obs}d\Omega}} \simeq \frac{\phi(\gamma)}{\sqrt{\phi(h)}}\sqrt{AT_{obs}d\Omega} \tag{3}$$

With the aid of the neural network rejection, the number of protons is suppressed by a factor $(1-\epsilon_h)$ (neglecting, of course, the contribution from the gamma showers), while the gamma signal is reduced by a factor ϵ_γ. The previous equation becomes:

$$\frac{n(\gamma)}{\sqrt{n(h)}} \simeq \frac{\phi(\gamma)}{\sqrt{\phi(h)}}\frac{\epsilon_\gamma}{\sqrt{(1-\epsilon_h)}}\sqrt{AT_{obs}d\Omega} \tag{4}$$

This corresponds to an effective area given by:

$$\frac{A_{eff}}{A} \simeq \frac{\epsilon(\gamma)^2}{1-\epsilon_h} \tag{5}$$

Taking, from table 2, $\epsilon_\gamma \simeq \epsilon_h \simeq 0.8$, we obtain:

$$\frac{A_{eff}}{A} \simeq 3 \tag{6}$$

This means that a rejection factor of the order of $80\ \%$ furnishes an increase in the sensitivity corresponding to an apparatus three times bigger, while a rejection factor of the order of $70\ \%$, which represents our worst result, corresponds to an apparatus more than two times bigger. We underline that the rejection power of the neural network can be tuned in order to fit the particular problem under analysis. In the field of gamma astronomy, the relevant quantity is that of eq. 4. In fig. 4 we show the behavior of this quantity, as a function of ϵ_γ for different values of $\epsilon_\gamma + \epsilon_h$ which represents an estimation of the global identification power of the neural network.

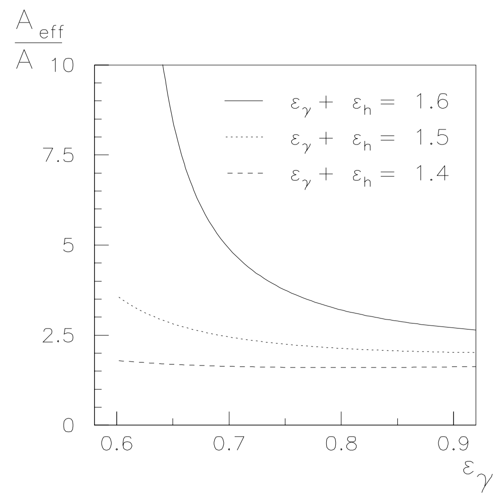

Figure 4: The enhancement in the gamma ray sensitivity, as a function of ϵ_γ, for three different values of the global identification power of the network: 80% (full line), 75% (dotted line) and 70% (dashed line).

Therefore, the use of this neural network improves the sensitivity of the apparatus, lowering the total observation time required to detect a gamma ray source above the cosmic ray background.

References

Abbrescia, M. et al., 1996, Astroparticle physics with Argo (unpublished). This document can be downloaded at the URL: http://www1.na.infn.it/wsubnucl/cosm/argo/argo.html

D'Ettorre, B. et al., 1999, Proc. 26th ICRC (Salt Lake City, 1999)

Fesefeldt, H., 1985, PITHA Report 85-02 (Reading: RWTH Aachen)

Heck, D. et al., 1998, FZKA Report 6019 (Reading: Forschungszentrum Karlsruhe GmbH)

Knapp, J., Heck, D. & Schatz, G., 1996, FZKA Report 5828 (Reading: Forschungszentrum Karlsruhe GmbH)

Mari, S.M. et al. 1999, Proc. 26th ICRC (Salt Lake City, 1999)

Nelson, W.R., Hirayama, H. & Rogers, D.W., 1985, SLAC Report 265 (Reading: Stanford National Acceleration Center)

Werner, K., 1993, Phys. Rep. 232, 87