ELSEVIER

Astroparticle Physics 15 (2001) 65–77

Astroparticle Physics

www.elsevier.nl/locate/astropart

Gamma–hadron discrimination in extensive air showers using a neural network

S. Bussino *, S.M. Mari

Dipartimento di Fisica, Università di Roma Tre, INFN, Sezione di Roma Tre and INFN, Via della Vasca Navale 84, I-00146 Rome, Italy

Received 17 May 1999; received in revised form 21 April 2000; accepted 21 July 2000

Abstract

A neural algorithm was developed to separate electromagnetic and hadronic showers detected with an air shower array. The requirements on the detector performance are very general, so that the results of the calculation can be applied to a wide set of detectors, actually operating or planned for the future. More then 700 000 showers were generated using the Corsika package and were propagated through an ideal pixel-like detector. The peculiarities of each class of showers are presented in detail and it is shown how the neural net architecture is structured around them. The neural net performances were studied for different sets of simulated data. The physics relevance of the gamma–hadron separation is also discussed. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Extensive air showers; Gamma astronomy; Neural networks; Cosmic rays

1. Introduction

The problem of separating gamma-induced showers from proton-induced showers is a well-established topic in the physics of cosmic rays [1–7]. In the field of gamma astronomy there is wide interest concerning the effort to extend toward the region of 1 TeV the search for signals of unknown point-like gamma-ray sources. The data coming from conventional extensive air showers (EAS) array detectors are above 50 TeV, because of their high thresholds, while measurements outside the atmosphere cannot reach the TeV region because of the gamma low-flux due to the steep energy spectrum of the sources. Cherenkov detectors reach TeV energies, but with a low duty cycle and a small solid angle. For these reasons, the detection of gamma ray sources in the TeV region requires specific experiments, like Milagro in New Mexico [8] or ARGO-YBJ in Tibet [9,10]. The performances of these kinds of detectors are discussed in a very general way in Ref. [11]. If this class of detectors could provide also a separation between the gamma ray from the source and the background due to the cosmic ray, the detector sensitivity would increase. A shower array detector seems to be the best choice for this field of gamma-ray astronomy, because of its duty cycle and of its large angle coverage, which permit a continuous all-sky search for unknown sources. In order to lower the threshold, the apparatus must be installed in high-altitude laboratory, like Chacaltaya [12], Yangbajing [13,14] or other proposed sites

* Corresponding author.

E-mail addresses: bussino@fis.uniroma3.it (S. Bussino), mari@fis.uniroma3.it (S.M. Mari).

0927-6505/01/$ - see front matter
PII: S0927-6505(00)00152-3

even at higher altitudes [15]. Focusing the interest on EAS detectors, the discrimination between gamma- and hadron-induced showers can be performed essentially in three ways:

(i) by a different trigger sensitivity to showers induced by gamma or by hadrons;
(ii) by identifying particles typical of the hadronic shower (i.e. muons or neutrons);
(iii) by studying the different topological structures of the gamma- or hadron-induced showers.

The first point is a quality factor, generally known as Q or R_γ (see Section 7.2 for details), intrinsic to each detector, which slightly increases the overall sensitivity. The second method, with special regard to muons, was widely used by array detectors [4,16–18] but is not practicable in all the situations, at least for two reasons:

(i) the technique requires additional hardware, which cannot always easily equip the shower detector or that can be added only in a very small fraction of the full apparatus;
(ii) the number of specific hadronic shower particles is too low for the energy region which concerns our work.

These two points are strictly linked one to the other: if we want to detect very low fluxes of specific particles contained in the hadron-induced EAS, the upgrade of the apparatus should be substantial. It has to be considered that the number of muons for hadron-induced showers of 1 TeV is of the order of 1 μ over 150×150 m^2 at sea level [11]. Also the neutrons are secondary particles produced in the hadron-induced showers [19]: however their detection seems difficult to achieve, because of the low flux and because the neutrons induced signal is delayed with respect to the shower. Using a pixel-like shower array, the collected data reproduce the spatial structure of the front of the shower; the position of the hits on the detector and their density are two fundamental observables which every apparatus can store. Using this electronic shower image it is possible to separate gamma and hadron showers with high degree of reliability: a neural algorithm can be planned to distinguish, event by event, one class from the other by means of different topological behaviors that are evident in an averaged sample. A similar use of a neural net was implemented in the past at TeV energy for Cherenkov detectors [3] or at energy above 30 TeV for air shower arrays [4]. Also, for balloon-borne experiments [20] and calorimetric detectors at the LHC energy [21], neural net techniques were applied. The aim of the present work is to exploit the possibility of using the neural net algorithm for air shower arrays around and below the TeV energy region. For these reasons, we chose to sample the shower front at a slant depth of 600 g cm^{-2}, corresponding to high altitude sites (i.e. Chacaltaya and Tibet), closer to the shower maximum, in order to obtain an electronic image of the shower as detailed as possible and to lower the energy threshold down to the hundreds of GeV. Further on we will briefly describe the characteristics of an ideal detector and the requirements concerning its reconstruction power. Later we will discuss the Monte Carlo data sample and the characteristics of the two classes of events will be presented in detail. In the third part, based upon the first two sections' assumptions and their results, we will present the structure of the neural network and the philosophy behind it, in relation to certain peculiarities of the whole problem. The fourth part will be devoted to a detailed analysis of the performance of the neural network, for different sets of events classified on the basis of observable quantities. The conclusion will discuss the improvement of the detector sensitivity in the field of gamma astronomy.

2. The ideal detector performance

An ideal detector [11,22,23] should be able to furnish the electronic image of each event: we have in mind a pixel-like detector, with each active element (scintillator, streamer tubes, RPC, and so on) of the order 50×50 cm^2. These active elements can concretely be the physical substructures of the whole detector or can be logically designed during the event reconstruction; for the purpose of computation we need only to know the particle

density averaged over each elementary surface. We suppose also that this density can be modeled on a digital scale spanning from 0 to 7 (3 bit word). An important element in the gamma–hadron discrimination is the possibility to measure the particle density with great accuracy over the largest possible area: for this reason the apparatus is supposed to be a full coverage detector. We have considered in our computation a continuous square made of 20 000 logical pixels (corresponding to about 70×70 m^2) and with a dead space of the order of 5%. With such a detector, by using a very rough algorithm for the core reconstruction, based on the center of gravity of the hits distribution, we obtained a spatial resolution spanning from 2.5 to 7.0 m, depending on the characteristic of the shower. The core reconstruction will be improved by a factor of $\approx$30% by means of a fitting procedure of the radial shape of the hit density. We have tested that the identification power of the net is lowered by a few percent by such a level core reconstruction uncertainty (see Table 6). For these reasons and because we want to focus on the gamma–hadron separation problem, we will present the results requiring that the detector should be able to reconstruct the position of the core with an uncertainty of 3 m along each direction. The spread is supposed to be distributed along a Gaussian with such a width and we also suppose that each coordinate is independent from the other. By using a fitting algorithm to evaluate the core coordinates, their estimation is roughly independent from the core position on the carpet. We have simulated a trigger based on a majority logic, requiring a coincidence of at least 25 fired pixels in a time window of 300 ns. The time window size was estimated considering the time delay between two particles hitting two opposite corners of the detector for incoming showers at a maximum zenith angle of 60°. The trigger efficiency affects only the population of each subsample (gamma or hadrons-induced shower) without influence on the geometrical structure of the detected electronic image of the shower. On the contrary, the efficiency for each logical subdetector is relevant, because in principle it could distort the geometrical structure of the event. We suppose that the detector has an overall efficiency of the order of 95%. This is the most simple trigger that can be implemented for an ideal detector, without taking into account the peculiarities of each apparatus. We stress that these requirements fit very well with the conceptual design of a new generation shower array detector [11]; however, the results can be modified or scaled over the more realistic structure of existing [8] or planned detectors [24,25].

3. The Monte Carlo data sample

3.1. The events generation

The events were generated by using Corsika [26] code 5.61 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 [27,28], while for the hadronic component several options are available: the Monte Carlo data generated for this paper were processed by VENUS code [29] for the high energy hadronic interactions and by GHEISHA code [30] for the low energy hadronic interactions [31]. The data were generated in the energy range 100 GeV to 10 TeV with the usual energy distribution given by

$$N(E)\mathrm{d}E = N_0 E^{-2.7}\mathrm{d}E. \quad (1)$$

The relevant parameters used in the events generation are reported in Table 1. In order to avoid an accumulation of simulated data in the lower energy bins, data were generated in different energy intervals, according to the energy spectrum given by Eq. (1), as shown in Table 2. The image of

Table 1
Parameters used for the Monte Carlo events

Primary particle	Spectral index	Energy range (TeV)	Energy threshold
Proton	2.7	0.1–10	$E_{\mu} = 5$ MeV $E_{\gamma,e^{+-}} = 30$ KeV $E_{\mathrm{hadron}} = 5$ MeV
Gamma	2.7	0.1–10	$E_{\mu} = 5$ MeV $E_{\gamma,e^{+-}} = 30$ KeV $E_{\mathrm{hadron}} = 5$ MeV

Table 2
Number of generated showers in different energy intervals of the primary particle

Primary energy interval (TeV)	Number of generated showers	
	Gamma	Proton
$0.100 < E < 0.500$	360 000	360 000
$0.500 < E < 1$	12 000	12 000
$1 < E < 3$	5000	5000
$3 < E < 10$	770	770
Overall ($0.100 < E < 10$)	377 770	377 700

each shower was sampled at the altitude of 600 g cm^{-2} and was recorded and prepared for the detector simulation.

3.2. Detector simulation and data processing

The Monte Carlo were processed through a code simulating the expected performance of a detector which satisfies the general requirements (high duty cycle, large effective area, high altitude, fast timing) and the geometrical setup (20 000 pixels covering about 70×70 m^2) discussed in Section 2. For a realistic simulation of such a detector we have implemented the logical pixels as a single layer of RPC chambers, equipped with strips or pads as pickup electrodes, which can easily fit the requirements we have discussed. The detector simulation takes into account the energy loss in the material, including the RPC, the ionization process in the RPC gas and the RPC efficiency. The background generated by each logical pixel has been simulated and added. The trigger described in Section 2 was applied. The number of events passing the trigger logic are reported in Table 3. We chose to put the shower core on the center of the detector, in order to analyze the performance of the net in a situation in which the shower is well reconstructed. The effect of the core position on the net identification power will be discussed later (see Section 6.2).

Table 3
Number of events surviving the general cuts due to thresholds and trigger

Mean primary CR energy (TeV)	Number of detected events	
	Gamma	Proton
$0.100 < E < 0.500$	36 720	5980
$0.500 < E < 1$	10 612	10 060
$1 < E < 3$	4470	3814
$3 < E < 10$	700	731
Overall ($0.100 < E < 10$)	52 502	20 585

The core of each event is put at the center of the detector (see text for details).

4. Characteristics of gamma and hadron showers sampled by a pixel-like detector

The pattern of the hits distribution can be used to distinguish between electromagnetic and hadronic showers, and for this reason we have studied in detail the geometrical behavior of the two classes of events, sampled at 600 g cm^{-2}. 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 by the presence, in the hadronic showers, of local fluctuations induced by small electromagnetic subshowers produced by neutral pion decay. We stress that the two different topological behaviors, enhanced by the very high altitude we assumed, can produce useful information for discrimination. These different topological behaviors should be used in order to separate gamma- from hadron-induced showers. In Fig. 1 we show the distribution of the hits number as a function of the pixel coordinates for gamma- and proton-induced showers. The two distributions show that the two classes of events, averaged over the whole energy range and over all the generated events, are clearly different. Another tool useful in distinguishing gamma- from hadron-induced showers is the distribution of the fluctuations of the number of hits detected by each pixel unit. We consider the quantity $F(x, y)$ defined as:

$$F(x,y) = \frac{|N_{\text{hits}} - \langle N_{\text{hits}} \rangle|}{\sqrt{\langle N_{\text{hits}} \rangle}}, \tag{2}$$

where the mean $\langle N_{\text{hits}} \rangle$, computed event by event, is a function of the distance from the core of the

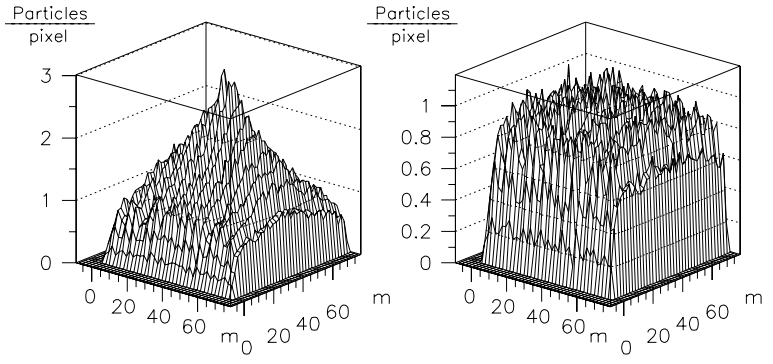

Fig. 1. The hits distribution in the detector, for vertical events with the core at the center of the carpet, for gamma (left) and protons (right). Distances are measured in meters with respect to the corner of the detector. The dimension of the pixel is 50×50 cm^2. Data are averaged over all the generated events.

shower. Details of the computation are discussed in Section 5.2. Because of the presence of absolute value, $F(x,y)$ can assume very different values also for events with the same number of fired pixels, due to the fluctuations around the mean value. The distribution of this quantity for electromagnetic and hadronic showers is shown in Fig. 2. The Monte Carlo events have been classified in several groups according to intervals in the number of hits recorded in the apparatus. As can be immediately seen from Table 3, the yield of the events which trigger the detector is different for gamma and hadron in the same energy range. In order that events generated in a very wide energy range with yields different for gamma and hadron in an energy dependent way not be processed as a whole sample, we classified the data by using the hits number, which is an observable. For each class of events of a different hits number, we could then tune the neural network in order to maximize its discrimination power. As an example, we show in Fig. 3 the behavior of the quantity plotted in Fig. 2 with regard to the two classes of the hits number: 75–100 and 150–200 hits per event. As is evident, the selection of hit intervals reduces the contamination level: the distributions referring to the same

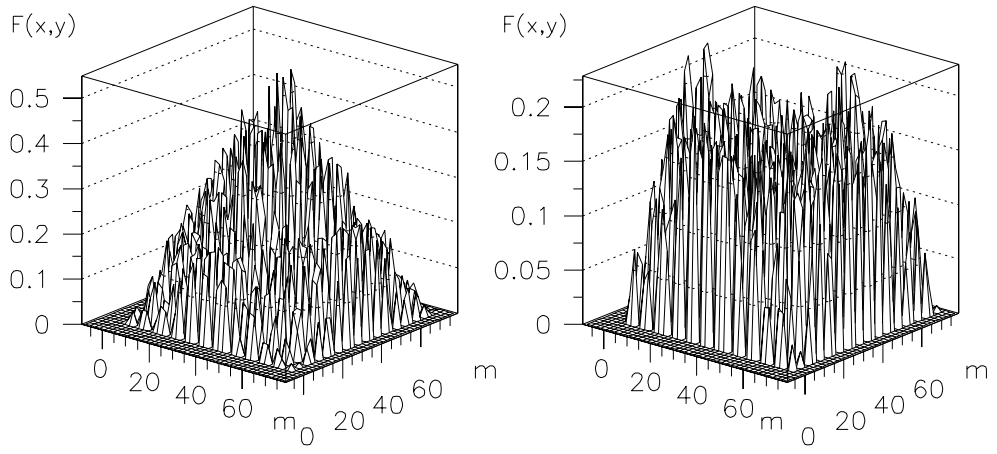

Fig. 2. The fluctuation $F(x,y)$ in the hit distribution, as discussed in the text, is shown for vertical events with the core at the center of the carpet (gamma: left, proton: right). Distances are measured in meters with respect to the corner of the detector. The units of the z-axis come from Eq. (2), measuring the number of particles over a pixel 50×50 cm^2 wide. Data are averaged over all the generated events.

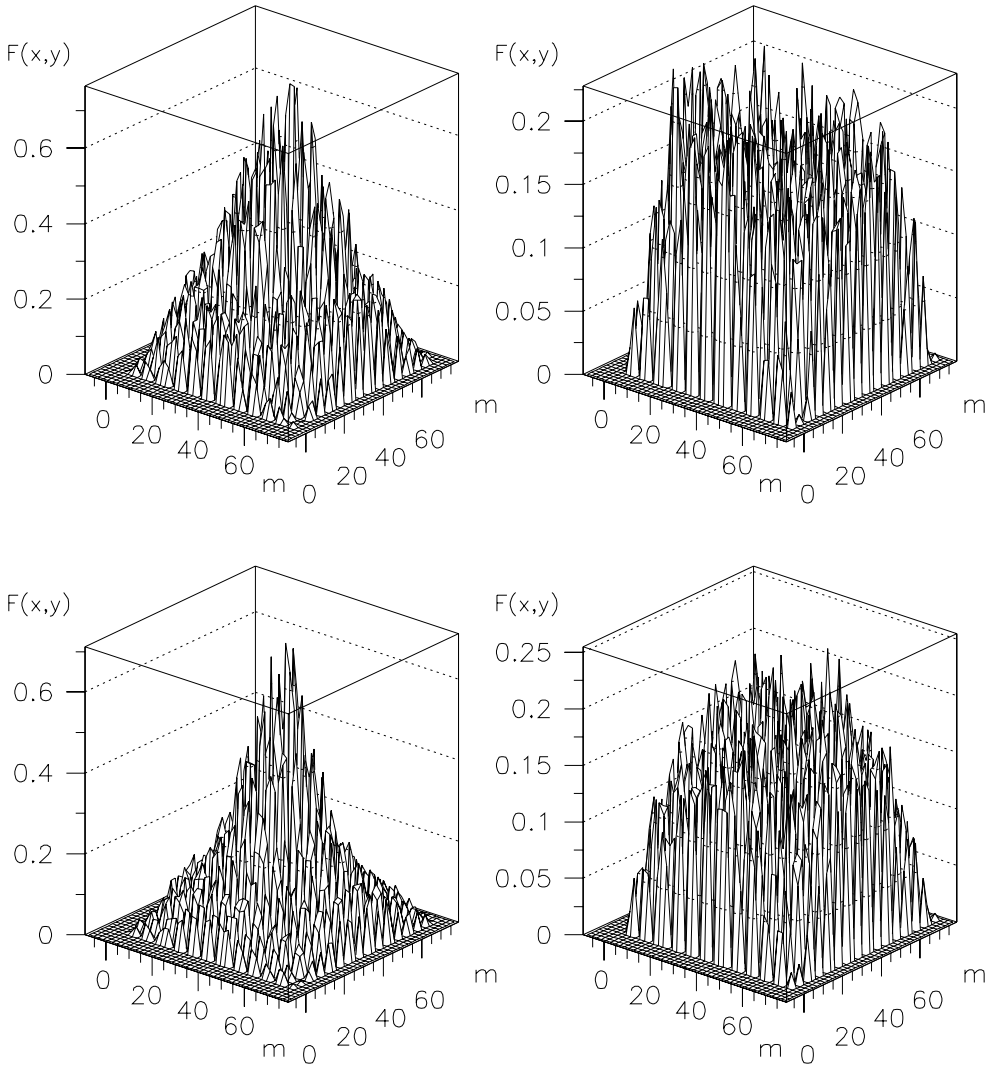

Fig. 3. The fluctuation $F(x,y)$ in the hit distribution in two different hit intervals: 75–100 (down) and 150–200 (up). (gamma: left, protons: right). See the text and Fig. 2 for details.

hit interval (Fig. 3) are more sharply different than the overall one (Fig. 2).

5. The neural network

5.1. General setup

The distribution shown in the previous pictures (see Figs. 1–3) are averaged over all the 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 must be planned with two logical steps in mind:

(i) to identify functions of observables which can discriminate between the two classes event by event;
(ii) to combine the discrimination power of the functions of point (i) in order to obtain the final rejection factor.

The first point should take into account the rotational symmetry intrinsic to this kind of process. The second step involves the optimization of the rejection factors resulting from the different functions of the observables. We have designed the network taking into account that the number of free parameters is closely correlated to the number of Monte Carlo events needed to teach the network. In Sections 5.2 and 5.3 we will describe in detail the structure of the neural network.

5.2. The input layer

The neural network (see Fig. 4) is a standard three layer perceptron, with only one output neuron. The input layer consists of four neurons, each of them reading different functions of the observables. Each neuron has a number of inputs equal to the number of pixels 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. We symmetrized the neural weights within each corona: all the inputs connected to the pixels belonging to the same corona have the same weight. In this way, based upon the intrinsic symmetry of the whole physics 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 20 000 (one for each pixel). The first function, used as input for the first neuron, is simply the intensity of the signal in each pixel. The second neuron reads as input the square number of the intensity of the signal in each pixel, in order to amplify the big signals, which are less sensitive to local fluctuations. So the functions used by the first two neurons reflect the geometrical shape of the shower. The other two neurons of the input layer are sensitive to the fluctuations of the data; the third neuron receives as input the quantity F described by Eq. (2). The mean $\langle N_{\text{hits}} \rangle$ is computed, event by event, over the hits belonging to the same circular corona. The inputs of the fourth neuron are the local maxima of the quantity F: only if the value of $F(x, y)$, computed in a pixel of coordinates (x, y), is greater than the value that F has on the eight contiguous pixels on the boundary, it is used as an input to the fourth neuron. The initial values of the weights of all these four neurons are randomly set to a positive value less than one.

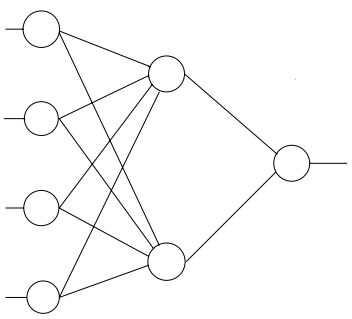

Fig. 4. The basic setup of the network. See text for comments.

5.3. The hidden and the output layers

Each neuron in the input layers has a rejection factor, with respect to the electromagnetic and hadronic shower, which depends on the pixels number and on the function processed by the neuron itself. 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 initial values of the weights of the hidden layer are set in such a way that the first hidden neuron works like a logical OR and the second one like a logical AND. The aim of the output neuron is to furnish the final result and the classification of the shower (i.e. 0 for gamma-induced showers and 1 for proton-induced showers). The initial weights of the output neuron are set in such a way that the neuron discriminates along the diagonal on the plane of the two outputs of the two hidden neurons.

5.4. The learning training

The network training is performed using a classical back-propagation technique, slightly modified in order to avoid the stand-by of the net on local minima or on the edge of the error function. The transfer function of each neuron is a sigmoid, common in a multi-layer network. A very

important requirement is the stability of the network results, i.e. the degree of confidence that we deal with after the learning training. The required result is that the network discriminates between the two sampled classes with the same rejection power during the learning training and during the period when it is required to analyze a new data set. In order to minimize the fluctuations due to non-physics broad changes in the weight profile, the weight w of each input neuron was studied as a function of the distance r from the core. This quantity must be smooth, because it reflects the radial distribution of the shower, averaged over events within the same hit number. For this reason a soft smoothing was applied to this weight profile, in order to avoid non-physics characterization of a single event in the learning data sample. A very usual smoothing algorithm was used:

$$w_{i+1}(r) = \frac{w_i(r-\delta r) + 2 \times w_i(r) + w_i(r+\delta r)}{4}, \quad (3)$$

where δr is the width of each circular corona. The behavior of the weights of the first two input neurons is shown in Fig. 5: the smooth profile ensures us that the network answer is of the same order on both the learning and on the testing samples. From a conceptual point of view, this procedure has the same meaning as a very soft learning training with a very large number of Monte Carlo data. This was tested on a reduced data sample.

6. Results concerning gamma–hadron separation

6.1. Discrimination factor for central core events

As discussed in Section 5, the data sample was classified into different subsets, according to the number of hits in the event. 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 results are summarized in Table 4, where the fraction of particle rightly identified is shown, which is the identification power of the neural network. The results that figure in Table 4 are obtained using the real core of the event. In order to simulate the uncertainties in

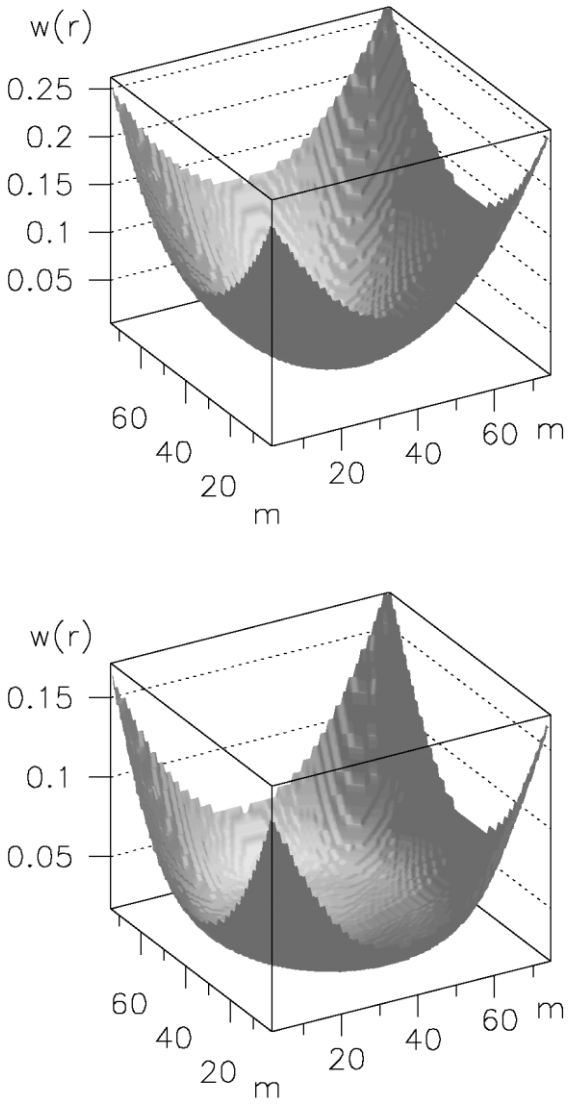

Fig. 5. The values $w(r)$ of the weights of two input-layer neurons (up and down) as a function of the pixel coordinate, measured in meters with respect to the corner of the detector.

the core reconstruction, we added a random error to the real core position and we measured a second time the identification power of the network. The scattering added to the core position was distributed along a Gaussian with a 3 m RMS. The results obtained in this framework are of the same order as those obtained without taking into account the error reconstruction. This result can be easily understood looking at the weight profile, as

Table 4
Identification power ϵ for gamma and proton showers hitting in the center of the detector, for different number of hit intervals

Hits number	Total number of events	True identification	
		ϵ_γ	ϵ_h
$25 \leqslant n \leqslant 50$	10 750	70%	69%
$50 \leqslant n \leqslant 75$	3760	75%	71%
$75 \leqslant n \leqslant 100$	1815	75%	73%
$100 \leqslant n \leqslant 150$	1640	80%	75%
$150 \leqslant n \leqslant 200$	760	81%	79%

shown in Fig. 5: the slope is smooth and the class of events is symmetric with respect to the center. Because of that, a slight shift in the core position has a slight influence on the whole input signal. The discrimination power coming from each neuron was estimated. The contribution to the discrimination factor ϵ due to the first two neurons, which describe the geometrical behavior of the shower, is of the order of 0.65, both for gamma and hadrons. The contribution of the last two neurons, which describe the local fluctuations of the shower, raises the discrimination factors ϵ_γ and ϵ_p up to the values reported in Table 4. This contribution is not negligible, because the enhancement of the detector sensitivity is a steep function of the discrimination factors ϵ_γ and ϵ_p, as shown in Fig. 7.

6.2. Discrimination factor for edge core events

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, as discussed in the section devoted to the Monte Carlo simulation. The distribution of the hits in the detector is shown in Fig. 6. The core position plays a very important role in the discrimination problem in two ways:

(i) the shower sampled by the detector is now asymmetric;
(ii) the effect of the error on the core reconstruction is not reduced by the symmetry of the sampling.

During the analysis the two aspects were separated and the first problem was studied changing

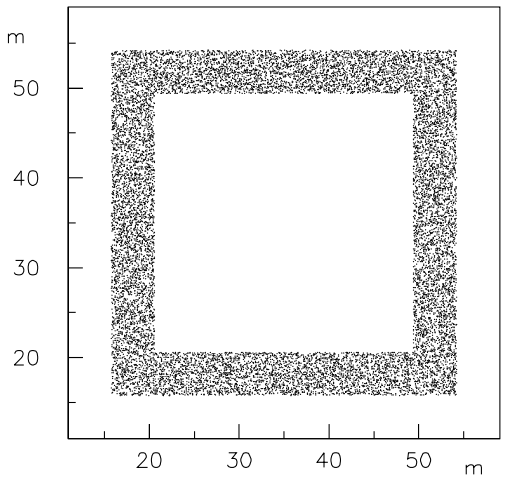

Fig. 6. The geometrical distribution of the core of the events generated on the edge of the detector, as discussed in the text. Distances are measured in meters with respect to the corner of the detector.

only the position of the core and assuming the Monte Carlo real core as the center of the shower. The effects on the discrimination power are shown in Table 5, for the usual data classification by means of the hits number. 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 whole effect due to the position of the core can be studied adding a random

Table 5
Identification power ϵ for gamma and proton showers generated with the core on the edge of the detector, for different number of hits intervals

Hits number	Total number of events	True identification	
		ϵ_γ	ϵ_h
$25 \leqslant n \leqslant 50$	9175	70%	70%
$50 \leqslant n \leqslant 75$	2890	77%	76%
$75 \leqslant n \leqslant 100$	1230	79%	77%
$100 \leqslant n \leqslant 150$	1110	79%	78%
$150 \leqslant n \leqslant 200$	450	80%	83%

Table 6
Identification power ϵ for gamma and protons showers generated with the core on the edge of the detector, for different number of hits intervals

Hits number	Total number of events	True identification	
		ϵ_γ	ϵ_h
$25 \leqslant n \leqslant 50$	9175	70%	67%
$50 \leqslant n \leqslant 75$	2890	75%	75%
$75 \leqslant n \leqslant 100$	1230	77%	76%
$100 \leqslant n \leqslant 150$	1110	78%	77%
$150 \leqslant n \leqslant 200$	450	80%	81%

The uncertainty in the core determination is taken into account in evaluating the identification power of the network. The same Monte Carlo data sample of Table 5 has been used.

error to the core position, using the same algorithm described in Section 6.1. In this case, the results shown in Table 6 suffer a loss in the discrimination power, due to the new topology of the problem which no longer has an intrinsic symmetry capable of compensating for small shifts in the reconstructed core position. We underline, however, that the discrimination power is generally of the order of 75% or more, a value high enough to obtain a substantial enhancement in the sensitivity of the detector, as shown in Section 7.

7. Improvement on gamma ray physics analysis

7.1. Enhancements of the detector sensitivity

In Section 6 we discussed the sensitivity of the network to identify gamma-induced showers. What is important for the gamma astronomy is the statistical significance of the flux estimated by the number $n(\gamma)$ of detected gamma showers. 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 $n(\mathrm{h})$ (a well-known value). 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 required ratio is:

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

where T_{obs} is the observation period, A^{eff} is the effective area of the detector for gamma- or hadron-induced showers, and ϕ_γ and ϕ_h are the fluxes of the source and of the cosmic rays background. The sensitivity of Eq. (4) corresponds to an effective observation period given by:

$$\frac{T_{\mathrm{obs}}^{\mathrm{eff}}}{T_{\mathrm{obs}}} \simeq \frac{\epsilon(\gamma)^2}{1-\epsilon_h} \simeq 3, \tag{5}$$

where $\epsilon_\gamma \simeq \epsilon_h \simeq 0.8$ is assumed. This means that a rejection factor of $\approx$80% furnishes an increase in the sensitivity corresponding to an observation time three times longer, while a rejection factor of $\approx$70%, which represents our worst result, corresponds to an observation time more than two times longer. We underline that the rejection power of the neural network can be tuned in order to fit the particular problem under analysis: for example, in some particular cases we may not necessarily be interested in obtaining a similar value for the two parameters ϵ_γ and ϵ_h, but in reducing the contamination by suppressing one of the two contributions. In the field of gamma astronomy, the relevant quantity is that of Eq. (5). In Fig. 7 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.

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

7.2. The Crab as a standard candle

As an example, we can briefly discuss the sensitivity of this kind of detector by evaluating the observation time required to detect the standard Crab signal over the cosmic rays background. The signal to noise ratio (see Eq. (4) for a simplified expression) in a more accurate way is given by the following equation:

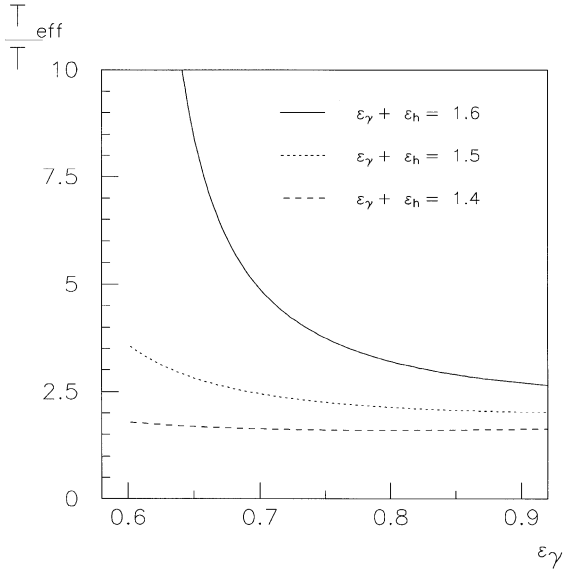

Fig. 7. The enhancement in the gamma ray sensitivity, as a function of ϵ_γ, for three different values of the global identification power of the network: 80% (—), 75% (···) and 70% (- - -). See text for details.

$$\frac{n(\gamma)}{\sqrt{n(\mathrm{h})}} \simeq \sqrt{T_{\mathrm{obs}} f(\delta)}\,\frac{0.72}{\sqrt{2\pi(1-\cos\theta)}} \times \frac{\int A_\gamma^{\mathrm{eff}}\phi_\gamma\,\mathrm{d}E}{\sqrt{\int A_{\mathrm{h}}^{\mathrm{eff}}\phi_{\mathrm{h}}\,\mathrm{d}E}}, \tag{6}$$

where $f(\delta)$ is the time fraction at which the source at declination δ is seen by the detector, θ is the angle of view opened around the source in order to collect a fraction 0.72 of its flux. If the neural network information is used, this sensitivity increases by a factor shown in Eq. (4). The Crab flux [33] and the cosmic ray background flux [34] are given by:

$$\phi_\gamma^{\mathrm{Crab}} \simeq 3.2\times10^{-11}E^{-2.5}\ \mathrm{cm^{-2}\,s^{-1}\,TeV^{-1}},$$
$$\phi_{\mathrm{h}} \simeq 1.1\times10^{-5}E^{-2.8}\ \mathrm{cm^{-2}\,s^{-1}\,sr^{-1}\,TeV^{-1}}. \tag{7}$$

A relevant parameter in Eq. (6) is the effective area of the detector. In order to obtain a realistic evaluation of this parameter, we use the values computed by the Argo [24] and the Milagro collaborations [8] for their two detectors. In both cases the effective area for gamma is quoted as $\simeq$10 000 m^2 at 1 TeV. The geometrical area of these two detectors is of the same order (60×80 m^2 for Milagro and 78×74 m^2 for Argo) and it is comparable with the ideal detector size discussed in Section 2. For these reasons, this value of the effective area for gamma can be chosen as a realistic and safe value. The detector sensitivity as a function of the gamma and hadron effective areas is, in first approximation, proportional to $R_\gamma\sqrt{A_\gamma^{\mathrm{eff}}}$, where $R_\gamma=\sqrt{A_\gamma/A_{\mathrm{h}}}$ is due to trigger effects. This trigger factor R_γ is greater than 1 and is closer to 1 as the energy increases [32]: by assuming the value 1, we will obtain a lower value for the performance of the detector. The angle θ is related to the angular resolution of the apparatus, which can be roughly estimated by taking into account the time resolution of a single pixel ($\simeq$1 ns, see Section 2) and the linear dimension of the full detector ($\simeq$70 m). A value $\sigma_\theta\simeq1^\circ$ can be assumed, including the full chain of data acquisition, as reported also by the two collaboration Argo and Milagro. Finally, the factor $f(\delta)$ is of the order of 0.20–0.25 for a source like the Crab and an apparatus in the Northern hemisphere, around the Tropic line of Cancer. By inserting these values in Eq. (6) and integrating above 500 GeV we can compute the observation time required to detect the Crab with a signal of 5σ above the cosmic ray background:

$$\left(T_{\mathrm{obs}}^{\mathrm{Crab}}\right)_{\mathrm{noNet}}^{5\sigma} \simeq 320\ \mathrm{days}. \tag{8}$$

A more accurate computation was obtained taking into account the energy dependence of the two effective areas A_γ^{eff} and A_{h} as estimated by the Argo-YBJ collaboration [24]. Their behavior was scaled to the dimension of the ideal detector described in Section 2 and inserted in Eq. (6). The result of the numerical integration is:

$$\left(T_{\mathrm{obs}}^{\mathrm{Crab}}\right)_{\mathrm{noNet}}^{5\sigma} \simeq 120\ \mathrm{days}. \tag{9}$$

The difference between this detailed computation and the rough estimation of Eq. (8) can be understood by taking into account the behavior of the factor R_γ, which in the relevant energy region of 0.5–1 TeV is around 2.5 and is 1 above 5 TeV. A one year 5σ sensitivity for a point-like gamma ray source, by using a factor $\sqrt{3}$ from the network

discrimination signal (see Eqs. (4) and (5)) and by assuming the same slope as the Crab one, can be simply obtained by scaling the result of Eq. (9):

$$\left(\phi_{\mathrm{obs}}^{\mathrm{min}}\right)_{\mathrm{Net}} \simeq 10^{-11} E^{-2.5}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}\,\mathrm{TeV}^{-1}. \qquad (10)$$

Without using the network discrimination information, the same source will be detected, with the same 5σ signal above the cosmic rays background, in three years.

8. Conclusions

The gamma–hadron separation has been studied by means of a neural network. A standard three layers perceptron was used to set up the network and particular care was taken in choosing the variables more sensitive to the different topological behavior of the gamma- hadron-induced showers. A fine tuning of the error back propagation was performed. In order to plan and to test the network, more than 700 000 showers were generated in the energy range 100 GeV to 10 TeV using Corsika package 5.61. The identification power of the network (ϵ) was measured for gamma and hadron showers: in Fig. 7 the global enhancement in the gamma-ray sensitivity was reported for a wide interval of the ϵ_γ parameter. The present work can be compared with the neural network discrimination power applied to the HEGRA and WHIPPLE detectors. The results discussed in Ref. [4] are related to the HEGRA air shower array [35], operating above 30 TeV: they obtained a proton identification factor $\epsilon_{\mathrm{p}} \simeq 0.95$, with a gamma identification factor $\epsilon_\gamma \simeq 0.6$. The comparison can be performed by looking at Fig. 7: HEGRA result $\epsilon_\gamma + \epsilon_{\mathrm{p}} \simeq 1.55$ lies between the curve obtained with $\epsilon_\gamma + \epsilon_{\mathrm{p}} \simeq 1.6$ and that one obtained with $\epsilon_\gamma + \epsilon_{\mathrm{p}} \simeq 1.5$. A gamma–hadron separation in the same energy interval was discussed in Ref. [3] by using the WHIPPLE Observatory TeV Imaging Telescope [36]. Also in this case their neural network identification power ($\epsilon_{\mathrm{p}} \simeq 0.99$ and $\epsilon_\gamma \simeq 0.6$) lies between the same two curves in Fig. 7, much closer to the best one. It should be considered that a Cherenkov, by using cuts on the image parameters and because of the high angular resolution of such a technique, can obtain a great enhancement in the sensitivity to a point-like gamma-ray source, for example, for the Crab, 34σ in 65 h of effective data-taking (between November 1988 and March 1989) [3], or 21σ in 20 h of effective data-taking (between November 1995 and January 1996) [6]. If a full sky survey is required, the comparison must take into account that the shower array detector has a large solid angle sensitivity, with a π sr angular acceptance for zenith angles between 0° and 60°, while the Cherenkov detector is a pointing device with a field of view of 4° circa. Moreover a shower array detector has a duty cycle close to 100%, while for a Cherenkov apparatus, the effective data-taking during one year is of the order of 800–1000 h. If we consider all these factors, the ideal pixel detector considered above (see Section 2) can perform a more detailed full sky survey, due also to the advantage of a high duty cycle and the possibility of studying transient phenomena and large time scale structures. The neural network described in this paper can improve the overall sensitivity, as discussed in the text. Other discrimination techniques, such as the identification of specific particles, generated in a hadron-induced shower, are powerless at this energy, mainly because of their low flux, as discussed in Section 1. With the general fashion detector discussed in the present work, the Crab signal can be detected in 120 days at 5σ above the cosmic ray background. For this reason the apparatus can be employed in the search of weak gamma ray sources: this search can benefit from the neural-net signal in order to lower the minimum flux detectable at 5σ during a one year period of data-taking: as shown in Eq. (10), this flux which is of the order of a third of the Crab flux, assuming the same energy spectrum.

References

[1] T. Gaisser, Cosmic Rays and Particle Physics, Cambridge University Press, Oxford, 1990.
[2] J. Matthews, et al., Search for diffuse cosmic gamma rays above 200 TeV, Astrophys. J. 375 (1991) 202–208.
[3] P.T. Reynolds, D.J. Fegan, Neural network classification of TeV gamma ray images, Astropart. Phys. 3 (1995) 137–150.

[4] S. Westerhoff, et al., Separating γ- and hadron-induced cosmic ray air showers with feed-forward neural networks using the charged particle information, Astropart. Phys. 4 (1995) 119–132.
[5] F. Arqueros, et al., Separation of gamma and hadron initiated air showers with energies between 20 and 500 TeV, Astropart. Phys. 4 (1996) 309–332.
[6] H.M. Badran, T.C. Weekes, Improvement of gamma–hadron discrimination at TeV energies using a new parameter, image surface brightness, Astropart. Phys. 7 (1997) 307–314.
[7] H. Cabot, et al., Measurable difference in Cherenkov light between gamma and hadron induced EAS, Astropart. Phys. 9 (1998) 267–276.
[8] B.C. Shen, et al., High energy gamma ray astroparticle physics with Milagro, Nucl. Phys. B (Proc. Suppl.) 71 (1999) 470–477.
[9] C. Bacci, et al., The use of RPC in the ARGO-YBJ project, Nucl. Phys. B (Proc. Suppl.) 78 (1999) 38–43.
[10] C. Bacci, et al., High altitude test of RPCs for the ARGO YBJ experiment, Nucl. Instr. Meth. A443 (2000) 342–350.
[11] R.S. Miller, S. Westerhoff, Conceptual design of a next-generation all-sky γ-ray telescope operating at TeV energies, Astropart. Phys. 11 (1999) 379–393.
[12] C. Aguirre, et al., Study of hadronic component in air showers at Mt. Chacaltaya, Nucl. Phys. B (Proc. Suppl.) 75A (1999) 186–190.
[13] T. Yuda, et al., Proc. Intern. Symp. EHECR, 1996, p. 175.
[14] C. Bacci, et al., Results from the ARGO-YBJ test experiment, Astropart. Phys., in press.
[15] L.W. Jones, D.M. Dei, Proc. 26th ICRC, vol. 2, 1999, pp. 524–526.
[16] M.A.K. Glashmacher, et al., The cosmic ray energy spectrum between 10^{14} and 10^{16} eV, Astropart. Phys. 10 (1999) 291–302.
[17] M. Aglietta, et al., Search for 100 TeV γ-ray emission from the galactic disk, Astrophys. J. 397 (1992) 148–152.
[18] M. Aglietta, et al., Comparison of the electron and muon data in extensive air showers with the expectations from a cosmic ray composition and hadron interaction model, Nuov. Cim. 112B (1997) 139–147.
[19] S. Hayakawa, Cosmic Ray Physics, Wiley, New York, 1969.
[20] F. Aversa, et al., Identification of cosmic ray electrons and positrons by neural networks, Astropart. Phys. 5 (1996) 111–117.
[21] S. Chattopadhyay, Z. Ahammed, Y.P. Viyogi, Application of neural network for photon–hadron discrimination in a preshower in high-energy heavy-ion collisions, Nucl. Instr. Meth. A421 (1999) 558–566.
[22] R.A. Ong, Very high-energy gamma-ray astronomy, Phys. Rep. 305 (1998) 93–202.
[23] C.M. Hoffman, et al., Gamma-ray at high energies, Rev. Mod. Phys. 71 (1999) 897–936.
[24] M. Abbrescia, et al., ARGO proposal, 1996, unpublished.
[25] C. Bacci, et al., Addendum to the ARGO proposal, 1998, unpublished.
[26] D. Heck, et al., CORSIKA: a Monte Carlo code to simulate extensive air showers, Forschungszentrum Karlsruhe Wissenhaltliche Berichte – FZKA 6019, 1998.
[27] W.R. Nelson, et al., SLAC Report, 265 (1985).
[28] A.F. Bielajew, et al., SLAC Report, 6499 (1994).
[29] K. Werner, Strings, pomerons and the venus model of hadronic interactions at ultrarelativistic energies, Phys. Rep. 232 (1993) 87–299.
[30] H. Fesefeldt, Report PITHA 85-02, RWTH Aachen, 1985.
[31] J. Knapp, et al., Comparison of hadronic interaction models used in air showers simulations and of their influence on shower development and observables, Forschungszentrum Karlsruhe Wissenhaltliche Berichte – FZKA 5828 (1996).
[32] J.W. Cronin, Gamma-ray astronomy by the air shower technique: performance and perspectives, Il Nuovo Cimento 19C (1996) 847–863.
[33] A.M. Hillas, et al., The spectrum of TeV gamma rays from the Crab nebula, Astrophys. J. 503 (1998) 744–759.
[34] K. Asakimori, et al., Cosmic-ray proton and helium spectra: results from the JACEE Experiment, Astrophys. J. 502 (1998) 278–283.
[35] F. Aharonian, et al., The status of the HEGRA experiment at La Palma, Nucl. Phys. B (Proc. Suppl.) 48 (1996) 491–493.
[36] M.C. Chantell, et al., A hybrid version of the Whipple observatory's air Cherenkov imaging camera for use in moonlight, Astropart. Phys. 6 (1996) 205–214.