УДК 539.1.074

$\gamma - \pi^0$ DISCRIMINATION WITH A SHOWER MAXIMUM DETECTOR USING NEURAL NETWORKS FOR THE SOLENOIDAL TRACKER AT RHIC

G.L.Gogiberidze[1], *R.R.Mekhdiyev*[2]

This paper presents a modern approach to discriminate $\gamma - \pi^0$ particles in the RHIC STAR shower maximum detector using neural nets. At the initial energy of 30 GeV the rejection factor, approximately 6, has been obtained for π^0.

The investigation has been performed at the Particle Physics Laboratory, JINR.

Распознавание $\gamma - \pi^0$ в детекторе максимума ливня для эксперимента STAR на коллайдере RHIC

Г.Л.Гогиберидзе, Р.Р.Мехтиев

С использованием методики нейронных сетей в рамках пакета JETNET смоделировано разделение $\gamma - \pi^0$ в детекторе максимума ливня установки STAR. При начальной энергии частицы 30 ГэВ достигнуто шестикратное подавление π^0 по отношению к γ-квантам.

Работа выполнена в Лаборатории сверхвысоких энергий ОИЯИ.

1. Introduction

The application of software neural network for the pattern recognition and triggering tasks, is well known. Here we have used a standard back propagation training algorithm following the JETNET package [1] to study the problem of discrimination between photons and neutral pions in the shower maximum detector (SMD) for the solenoidal tracker at RHIC (STAR Project) [2]. This problem is especially difficult at high energies when two electromagnetic showers initiated by photons from the π^0 decay are overlapped and appear as a single cluster in the apparatus.

This ivestigation was motivated by the needs of the STAR experiment for the γ and γ-jet physics, to determine the kinematics of the primary quark-gluon scattering. On the other hand, a clean sample of π^0-mesons is important for QCD studies.

[1]Permanent address: Institute of Physics, Georgian Academy of Sciences, GE-380077 Tbilisi, Georgia, e-mail: goga@sunse.jinr.dubna.su

[2]Permanent address: Institute of Physics, Azerbaijan Academy of Sciences, AZ-370143 Baku, Azerbaijan, e-mail: mehdiyev@sunse.jinr.dubna.su

It is also interesting to investigate the ability of the neural network to distinguish two particle types, and how this ability depends on the energy of primary particles.

2. Pattern Generation

A sampling calorimeter constructed of lead and plastic scintillator layers, was chosen for electromagnetic end-cup calorimater (EMCE) in STAR Project [2]. The shower maximum detector, a part of EMCE, is placed at a depth of ~ $5X_0$, between two longitudinal segments of EMCE. Its purpose is to measure centroids and lateral profiles of the showers to help electrons and gamma quanta to be separated from π^0.

To generate patterns for training and testing, 24 strips, 1.0 cm wide, placed in radial direction, were included to one tower of the EMCE instead of two active plates of shower maximum detector. The total width of the tower was about 30 cm. Particles were directed to the central part of the tower, which is rather large and covers half of the tower. For practical application we suggest that signal can be read out from all strips of the tower as well as from the strip that covers the halves of the neighbouring towers. The edge effects of the electromagnetic shower development have not been considered here.

In total, 9000 events of each particle type were generated, 3000 — at the energy of 10, 20, and 30 GeV, correspondingly; 2500 of each set of particles were used for training and 500 — for test purposes. These single particle events have been pulled through the detector using a STAR simulation program which is close to the official version.

SMD was placed between two longitudinal segments of EMCE.

The amount of material in front of the SMD was estimated as follows:

— 8.0 mm aluminum walls of time projection chamber (TPC);

— 12.7 mm aluminum front plate;

— five layers of EMCE: 5-mm-thick lead absorber plates and 4-mm-thick scintillator (Polystyrene) tiles;

— 3.2-mm-thick aluminium box of the shower maximum detector.

The maximum value of the magnetic field was 0.5 T.

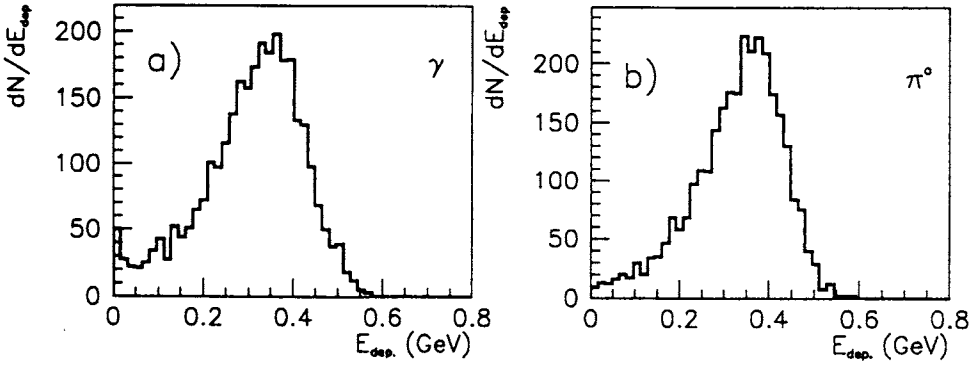

Fig.1. The distribution of the total energy deposition in shower maximum detector for γ (a) and π^0 (b) at 30 GeV

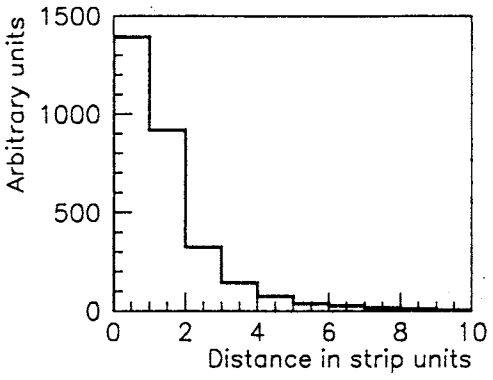

Fig.2. The distribution of the distance from the location of the SMD strip with maximum energy deposition

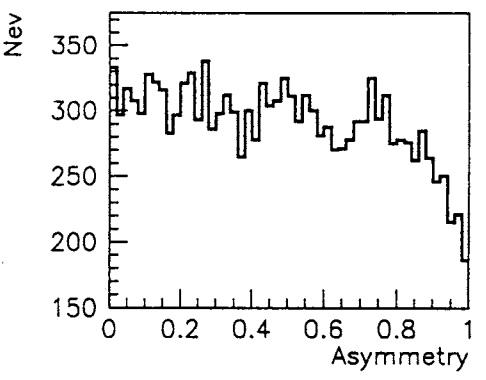

Fig.3. Asymmetry $R = \dfrac{(E1 - E2)}{E1 + E2}$ in the decay energies $E1$ and $E2$ of γ from π^0 decay

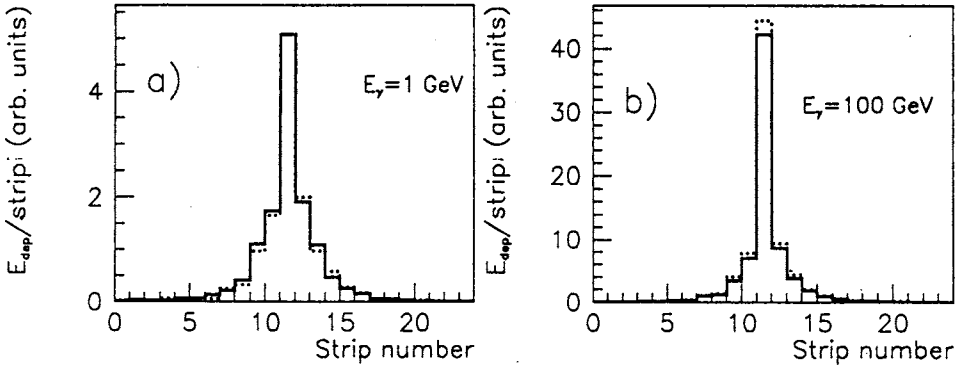

Fig.4. Profiles of e.m. shower in SMD initiated by 1 GeV (a) and 100 GeV (b) γ without (full line) and with (dotted line) magnetic field of 0.5 T

The distribution of the deposited energy in the SMD for incident γ (a) and π^0 (b), correspondingly, is presented in Fig.1. It is clearly seen that even for π^0 there is a non-zero probability that the deposited energy disappears from the first longitudinal segment of EMCE.

Let us enumerate possible cases when it is unachievable to resolve two photons from the π^0 decay:

- One of the photons is missed because it is outside the detector acceptance. Figure 2 shows the distribution of energy deposition depending on the distance from the location of the SMD strip having maximum energy deposition.
- One of the photons has a low energy and cannot be distinguished from the noise. The curve presented in Fig.3 shows asymmetry $R = \dfrac{(E1 - E2)}{E1 + E2}$ in the π^0 decay resulting in two γ-s with energies $E1$ and $E2$.

- One of photons does not give a significant energy deposition in the SMD due to statistic factors (γ does not convert before the SMD).
- Effective distance between two photons is small, compared with the transverse size of the electromagnetic shower.

The last effect strongly depends on the thickness of the matter before the SMD and the magnetic field which turns trajectories of the charged particles and therefore spreads out the electromagnetic shower. Figure 4 shows the averaged profiles of the e.m. shower in the SMD strip layers without (full line) and with (dotted line) magnetic filed of 0.5 T for γ with incident energy of 1 GeV (a) and 100 GeV (b), correspondingly.

Only these particles which give the energy deposition in the SMD not less than 40% on the average, are considered further. About 8% of events were rejected using this cut in the paper.

3. Network Training

The task to separate γ and π^0 is reduced to determining the correspondence between selection of the X_i input values and Y output values whose true meanings are known beforehand.

Our strategy was not to use sophisticated variables as, for example, in [3] but just the energy deposition in strip for input nodes. It is significant, that this method can be used not only for an off-line analysis, but also for a triggering, of course, with necessary hardware.

The architecture of the network employed in the present study is a conventional one, i.e., a network of the feed forward type with two hidden layers.

In this paper we do not intend to determine the best architecture, where the number of the hidden nodes is considered optimal. We have used two layers of hidden nodes. These results were obtained with the 24-16-8-1 architecture. The increase of the hidden nodes number does not improve the situation much, but dramatically rises the training time.

Due to the high number of connected weights in the neural network the principle of mirror symmetry has been used. We have prepared ordinary and «inverted» patterns with an equal probability, where strip number i was replaced by strip number $24 - i + 1$.

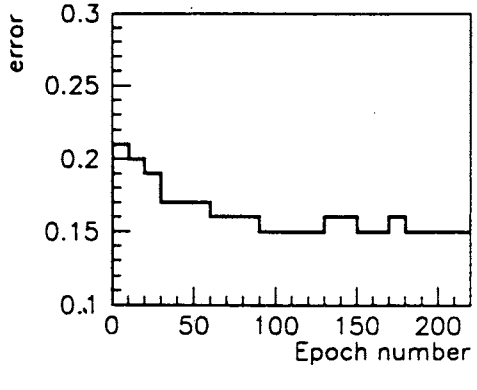

Fig.5. Mean square error E vs. epoch number

This is done to meet the ordinary requirements for neural nets: the number of patterns should be more than the number of the connected weights by 10 times.

The network starts with random weigths, the training vector (normalized energy deposited in strips) feeds into the network through; the output vectors o_i (dimensional in our case) are calculated and compared with target vector t. If the pattern is γ, then t is equal to 1, if not — t is equal to zero. When k events are presented to the network, the weights are updated by minimization of the summed square error E for all patterns k:

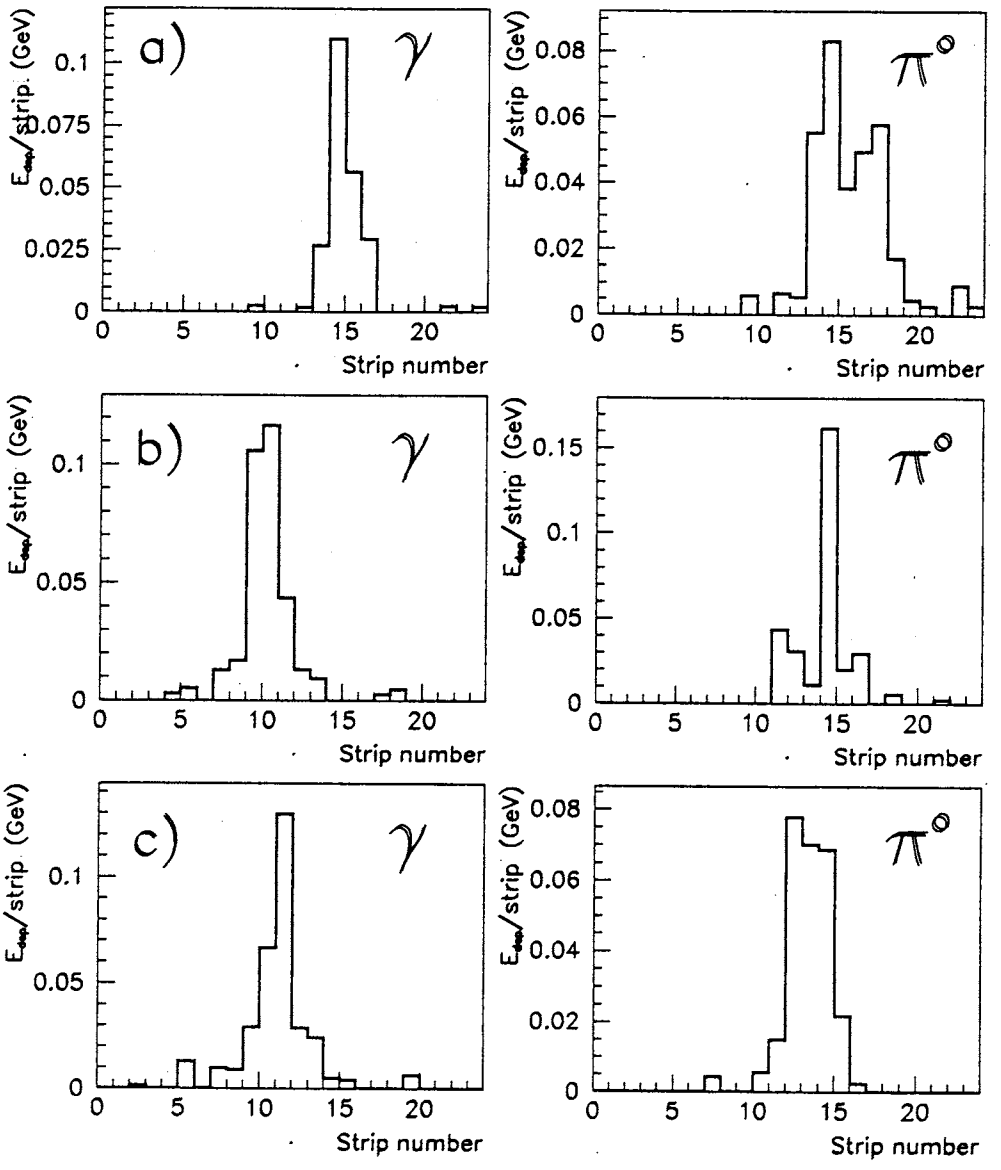

Fig.6. Energy deposition in strips of SMD by: well identified (1), partially identified (b), and poor identified (c) γ (left side) and π^0 (right side)

$$E = \frac{1}{2k} \sum_k (o_k - t_k)^2. \tag{1}$$

Weights were updated using the Langevin learning procedure which includes the Gaussian noise (0.2 of learning rate in our training). This procedure was repeated every k events. Our choice was k = 10.

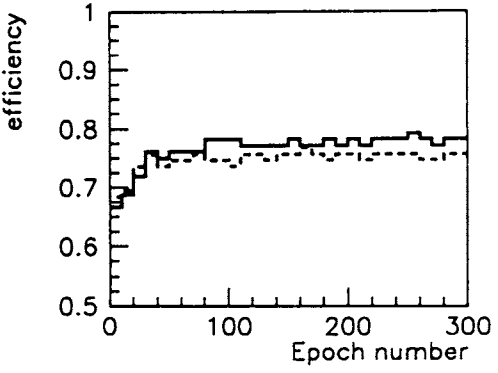

Fig.7. Neural network response (threshold ε = = 0.5) for training (solid line) and testing (dash line) sets in dependence on the epoch number

Presentation of the number of patterns equal to their quantity in the test set according to the adopted terminology is called epoch. The decreasing of a mean square error during the network training is shown in Fig.5. For each epoch the percentage of the correctly classified patterns is computed by means of threshold $\varepsilon = 0.5$, classifying an input as γ, if the output is greater than ε, and as pion — in another case. It must be emphasized that various ε thresholds can be used for practical purposes. The neural net recognizes patterns using shapes of energy deposition in the SMD strips with different results: (a) well identified; (b) partially identified and (c) hardly or misidentified γ and π^0 with the incident energy of 30 GeV (Fig.6).

The increase of network response (threshold $\varepsilon = 0.5$) for training and testing sets in dependence on the epoch number is presented in Fig.7. The difference illustrates adaptation of the neural net to the training set. This can be improved by a significant increase of the training set, but, at present, this way is costly due to CPU time.

4. Results

The set of test events used for checking the neural net response is completely independent of the training event set. Figure 8 presents the distribution of e.m. shower width for the events initiated by photons and pions of 10 GeV (a) and 30 GeV (b) initial energy. Shower width is defined as follows: $W = \frac{1}{2}\sum_k (R_k)^2 \times E_k$ for γ (solid line) and π^0 (dashed line), where R_k is the distance between the center of the shower and strip number k, E_k-energy deposited in the strip. Using this information we can reject γ from π^0 for the 10 GeV initial energy, but for 30 GeV the difference is small. With about 10% of γ loss, the cut on the shower transverse size, can reject only 45% of π^0.

Figure 9 shows the output signal for showers initiated by γ and π^0 for the 10 GeV (a) and 30 GeV (b) initial energy, correspondingly. As is seen, at the initial energy of 30 GeV the neural network can reject about 66% of π^0 with γ losses less than 10%.

To determine a possibility of improving the neural net response using a complementary information for 30 GeV events, we have combined the EMCE and the SMD information. Three additional input variables have been used:

- Maximum energy in the tower of the first longitudinal part of EMCE divided by the total energy deposition in this first part;
- The same for the second longitudinal part;

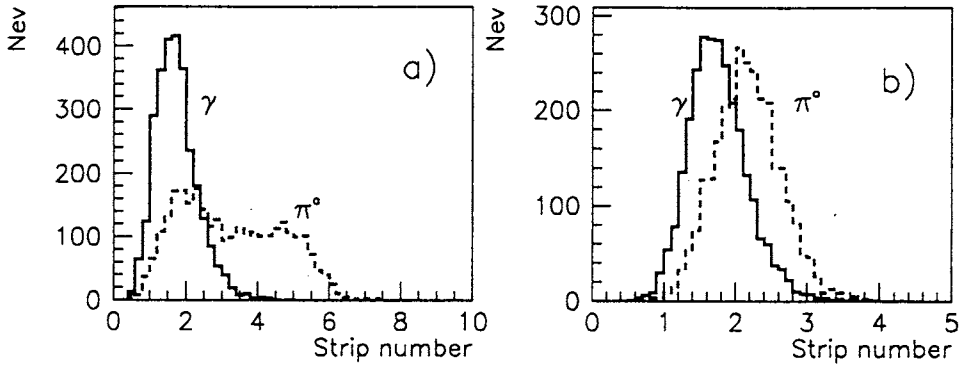

Fig.8. Distribution of e.m. shower width, as defined in the text, for 10 GeV (a) and 30 GeV (b) initial energy

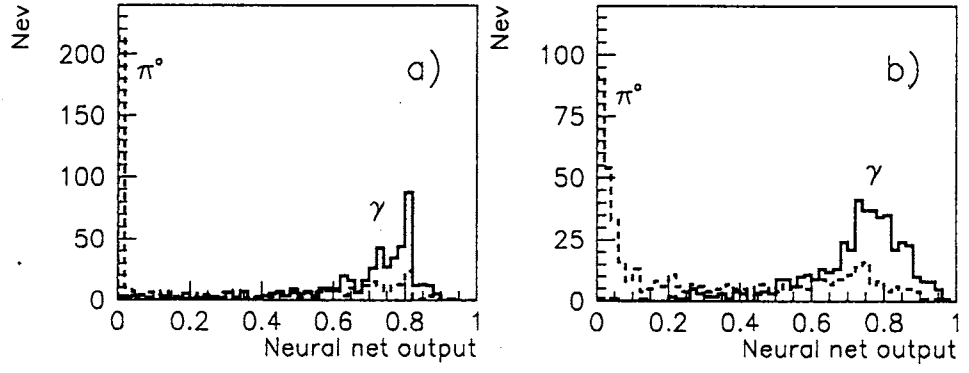

Fig.9. Neural net output for γ (dashed) and π^0 (solid) of 10 GeV (a) and 30 GeV (b) initial energy

- Total energy deposition in the first longitudinal part is divided by the total energy deposition in the EMCE.

We have also rejected the events with a small energy value (10% in average) deposited in the second hidden layer (less than 1% of all events). The 27-20-10-1 architecture and large number of training epoch (5000) have been used. The discrimination results on $\gamma - \pi^0$ made in the neural net at the initial energy of 30 GeV are given in Fig.10.

Let us define the rejection factor as the ratio of the number of signal gone through some conditions — to the number of background events also gone through the same conditions. With losses of γ less than 50%, the rejection factor, approximately 6, for π^0 has been obtained.

In Ref. [3] it is written that with the energy increasing the relative efficiency of neural network algorithm for $\gamma - \pi^0$ separation is better, compared with other algorithms. In our investigation the efficiency of neural network algorithm does not deteriorate with the energy increase.

It should be also noted, that despite the fact that the use of the information from EMCE decreases the mean square error as defined in (1) from 0.14 to 0.12, still it does not

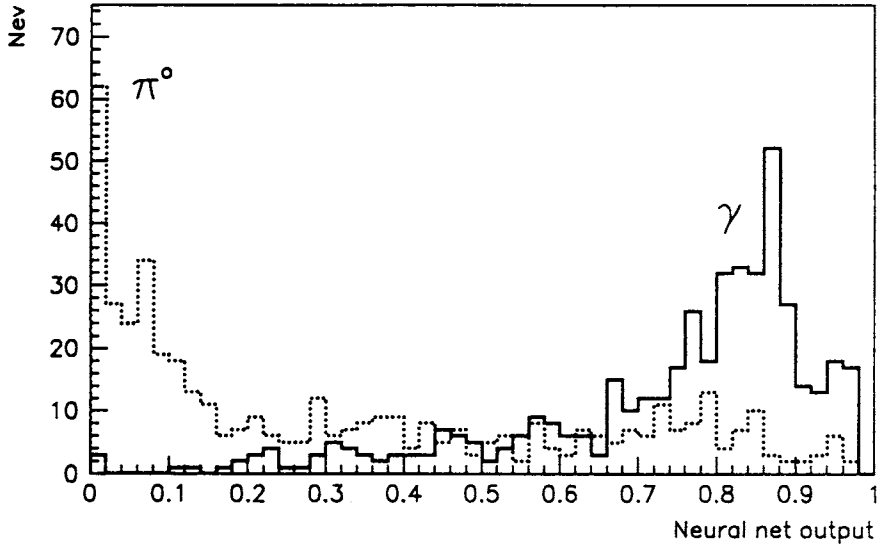

Fig.10. Neural net output for γ (dashed) and π^0 (solid) with additional EMCE information

drastically improve $\gamma - \pi^0$ separation. This can be explained in the spirit of [4]. Let us consider the border patterns, those where the neural net output is close to the border between the classes, e.g., is about 0.5. Adaptation of the neural net on the training set happens due to the number of the border patterns in the training set approximately equal to the number of connected weigths ~ 10^3. Additional training of the neural net on these border patterns (but independent of the initial training set) can improve the situation, but to generate the patterns requires huge CPU time resources.

5. Conclusion

Back propagation neural network is a powerful tool to study γ and π^0 separation in the STAR detector. Training and testing were performed with the Monte Carlo data. Results are presented for «poor» (one layer of strips) SMD choice. Rejection factor against π^0 equal to 6 at 53% γ-efficiency is achieved for the initial particle energies of 30 GeV.

References

1. Lönnblad L., Peterson C., Rögnvaldsson T. — Comput. Phys. Comm., 1992, v.70, p.167.
2. The Electromagnetic Calorimeter for the Solenoidal Tracker at RHIC, PUB-5380, p.4—40.
3. Babbage W., Thompson L. — Nucl. Instr. and Meth., 1993, v.A330, p.482.
4. Peterson C., Rögnvaldsson T. — CERN School of Computing, CERN 92-02, 1991, p.113.

Received on March 25, 1996.