Neural Network for BES Trigger

Bin Liu, Yanan Guo, Dixin Zhao
BES TRIGGER GROUP
IHEP, China

May, 1999

1 Motivation

In the experiments of high energy physics, the number of background is hundred thousands times more than that of good events, or even higher when the luminosity increases. The goal of trigger system is to retain good events, suppress background without causing too much dead time and to give a fast event selection(see Ref[1]). After several years of datataking, trigger data at different energies and experiences have been accumulated enough for trigger study. We guess there is a fine relationship between trigger conditions and event types. But this relationship could not be described by a simple formula, except by neural network. The background could be suppressed further by utilizing the relationship.

As the BEPC's luminosity is improved, the event rate and dead time of DAQ will increase a lot. A new method or technology should be used in trigger design. Now BES uses a three-level trigger. The third level trigger is consisted of logic circuits carrying out a combinatorial logic. The fourteen trigger conditions have the same weights in event selection which is difficult to solve the complex issues. If the trigger conditions have the different weights, like in Neural Network, a good event classification is expected to get. It would suppress background, decrease dead time and cut down event rate more efficiently under high BEPC's luminosity.

Neural network has been used widely in the experiments of high energy physics, like finding tracks in level two trigger, offline data analysis and particle identification. There are many commercial or free software packages of Neural network, like Matlab and JETNET, to support applications. Also the unique designs of DSP give a fast executing time to realize neural network algorithm, like hardware multiply/accumulate and pipeline. It would be the first time that a neural network based on DSP is used in BES's third level trigger to select events in real time.

2 Acceptable Dead Time

Because there is always an overlap between good event set and background set, we have to make a compromise between retaining good events and suppressing background in trigger. If conditions are loose to retain good events, then background will pass trigger selection a lot that causes dead time increasing and then some good events lost during dead time. If trigger conditions is strict to suppress background, dead time is decreased, but some good events will also be lost. So trigger table is finally confirmed depending on the requirements of experiments. Now the dead time of DAQ on BES is 10%, that means 10 percent of events will not be taken, which is considered acceptable by BES. Now if neural network is used in event selection in trigger, an estimation and calculation is given like following.

The rate of trigger event is 20Hz, the time for DAQ to take data is 10ms/event, then the dead time of DAQ is:

$$20Hz \times 10ms = 200ms/s = 20\%$$

and if it takes 60μs for neural network to select events, then the dead time of neural network is:

$$20Hz \times 60\mu s = 1.2ms/s = 0.12\%$$

which is negligible compared with DAQ dead time of 20%. If 30% of events is real background and expected to be cut off by neural network, then the dead time is reduced:

$$20Hz \times 30\% \times 10ms = 60ms/s = 6\%$$

reduced to:

$$20Hz \times (1 - 30\%) \times 10ms = 140ms/s = 14\%$$

3 Neural Network

3.1 Data preparation

Because the neural network calculation is performed before event data readout, the only available input variables are the logical conditions used in trigger table.
The events are selected by offline analysis, classified and tagged as four event types and their opposites: hadron or non-hadron, bhabha or non-bhabha, neutral or non-neutral and dimu or non-dimu. The classified events are provided by J/ψ and R group in IHEP.
Four Run data at J/ψ peak are used, Run11990, Run11991 and Run12060, Run12061. Data in Run11990 is provided for statistical analysis, neural network training. The other data are used as test sets. Run11990 and Run11991 were collected during R Scan last year, Run12060 and Run12061 during R Scan this year.
Some other R Scan data are provided for testing too. They are hadron, bhabha and dimu events sets at Center-of-Mass energies 2.6Gev, 3.2Gev, 3.44Gev, 3.55Gev, 4.6Gev and 5.0Gev, and selected by R group in first R Scan last year. Also include Run12711 and Run12712 at Center-of-Mass energy 4.02Gev in R Scan this year.
The difference of this year's data to last year's one is clearer MDC.

3.2 Statistical analysis

In order to confirm the input feature vector of neural network, SPSS(Statistical Package for Social Science) is used to analyse the relationships between trigger conditions and event types, which are listed in Table I.

TableI. Relationship between trigger conditions and event types

TRG	STANDARDIZED CANONICAL DISCRIMINANT FUNCTION COEFFICIENTS			
CONDTN	hadron	bhabha	dimu	neutral
TOF B–B	-0.063	0.515	0.509	0.000
$N_{TOF} \geq 1$	0.000	0.237	-0.339	-0.414
$N_{TOF} \geq 2$	0.367	-0.096	0.289	0.094
RADIAL.	-0.167	0.452	-0.115	-0.078
$N_{VC} \geq 1$	0.377	0.143	0.831	0.228
$N_{ETOF} \geq 1$	0.216	-0.468	-0.119	-0.045
END B-B	-0.059	-0.069	0.000	-0.034
$N_{TRK} \geq 1$	-0.076	0.127	-0.376	1.039
$N_{TRK} \geq 2$	0.350	0.174	0.115	0.035
$N_{TRK} \geq 4$	0.463	-0.415	-0.197	0.000
MU–OR	-0.176	-0.486	0.551	0.109
ESC-E_{tot}	-0.013	0.000	0.000	0.000
E_{tot-l}	-0.011	0.000	0.000	-0.045
E_{tot-h}	0.605	0.382	-0.420	-0.083

The standard canonical discriminant function coefficients represent the weights of relationship between trigger conditions and event types. Zero represents no relationship. The greater the value is, the tighter the relationship is. The minus means that the trigger condition makes a converse contribution to event selection.
From the table, the conclusion is that all trigger conditions have contributions to the four event types, so the input vector of the neural network is consisted of all the fourteen conditions.
SPSS is a linear, statistical classification. Neural network is a nonlinear and memorial one. For most nonlinear tasks, neural network could get better classification than SPSS.
For more details about SPSS and its applications, see Ref[3, 4].
Fig. 1 shows the distribution results of SPSS of four event types and their opposites. For examples, SPSS of hadron describes the distribution of hadron in SPSS analysis. N-hadron means non-hadron.

3.3 Training and structure of neural network

JETNET3.0 package[5] is used for training a neural network. The structure, according to this application, is selected as a feed-forward multi-layer perceptron network with a hidden layer which is sufficient for most classification tasks. The training algorithm is Back-Propogation(BP) learning rule which is not only sufficient but often superior for most tasks in spite of its simplicity. It is also a very stable and the most widely used learning algorithm. Some trigger data in Run11990 are used as a training set and some as a validation set which determine the termination point for the training.
The selection conditions of the software are very loose to let all the good events passed. So the good events are interfused with backgrounds and not very pure. The training samples of hadron and bhabha are relatively good. But the samples of neutral and dimu are very little and insensitive to current trigger conditions, not good for training.
Here, the good event is defined as one of the four event types, hadron, bhabha, neutral and dimu.

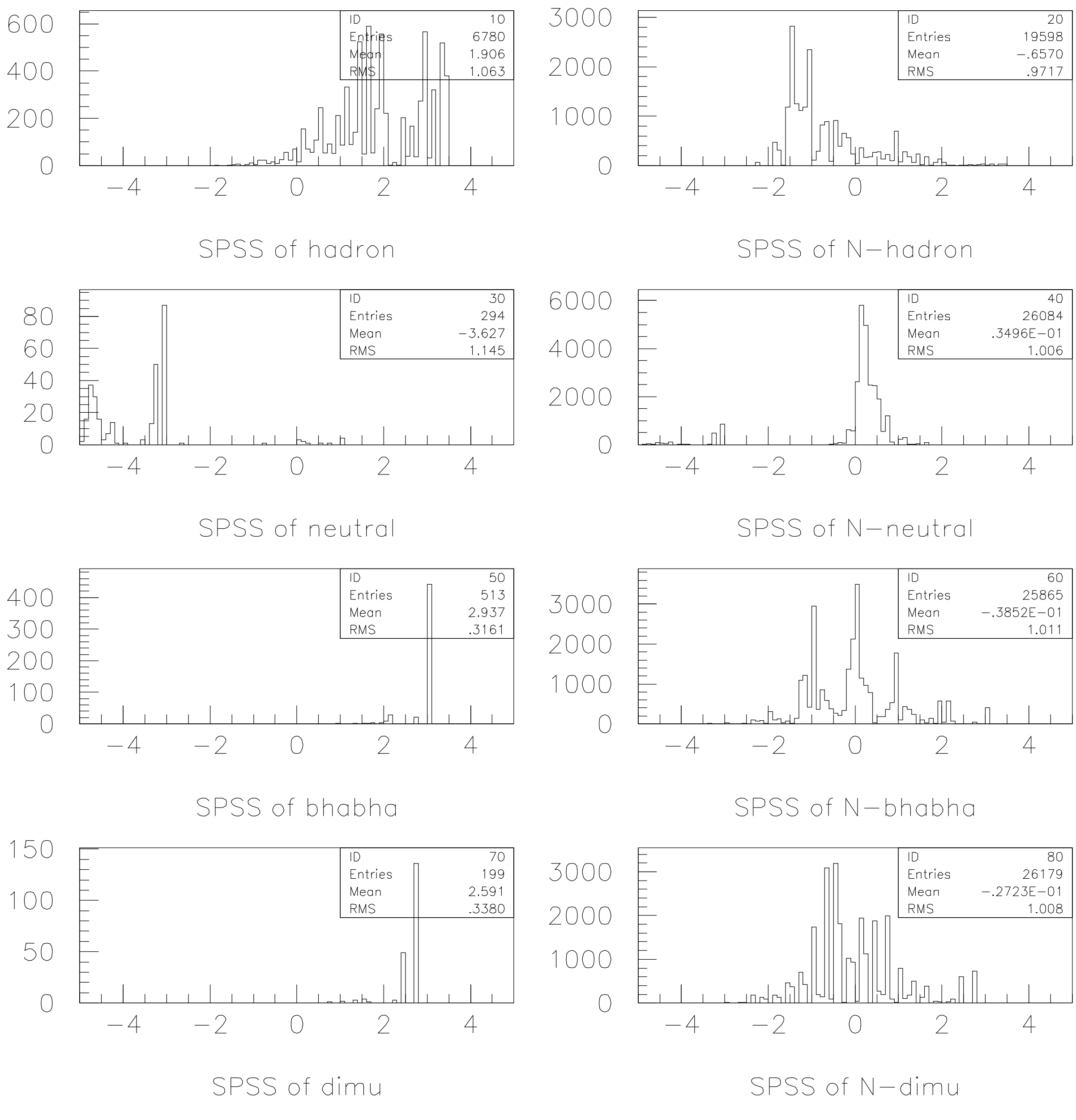

Figure 1: Event Distribution Result of SPSS

The final structure of neural network is: 14 input nodes, 8 hidden nodes and 1 output node. A threshold is set to the network to give a binary output which determine if a event is good or not and will affect the perfomance of the network. (see Fig. 2)

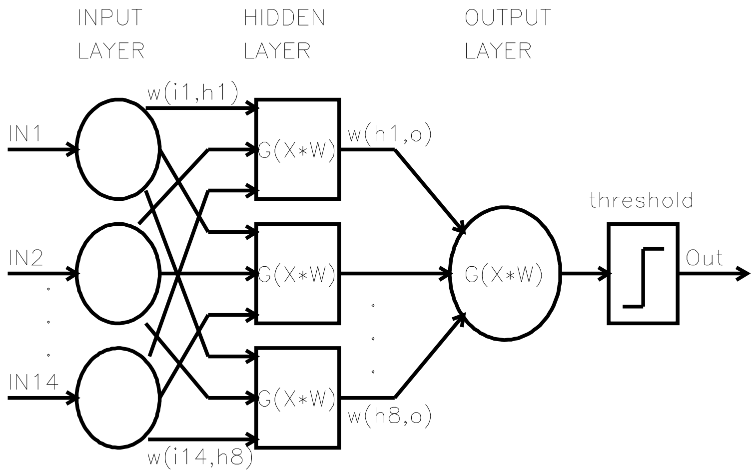

Figure 2: Structure of Neural Network

4 Event Classification

The trigger data in Run11991 and Run12060–12062 is used as a test set to estimate the generalizations of neural network at J/ψ energy. The other test sets are used to estimate the generalizations of neural network at different energies.
The denominators in columns in Table II represent the numbers of four good event types, background selected by offline software and the total event numbers in each run. The numerators represent the numbers of four good event types and background selected by neural network.
In order to improve the generalizations of neural network, in other word, to retain good events as possible, the thresholds should be adjustable for four event types seperately. The generalizations of neural network are showed in Table II and III when threshold of hadron(T_{had}) is set as 64, threshold of bahbha(T_{bb}) as 64, threshold of neutral event(T_{neu}) as 190 and threshold of dimu(T_{du}) as 220.
Here
G_{had} : percent of retained hadrons to total hadrons;
G_{bb} : percent of retained bhabhas to total bhabhas;
G_{neu} : percent of retained neutral events to total neutral events;
G_{du} : percent of retained dimus to total dimus;
S_{bk} : percent of suppressed backgrounds to total backgrounds;
S_{tot} : percent of suppressed events to total events.

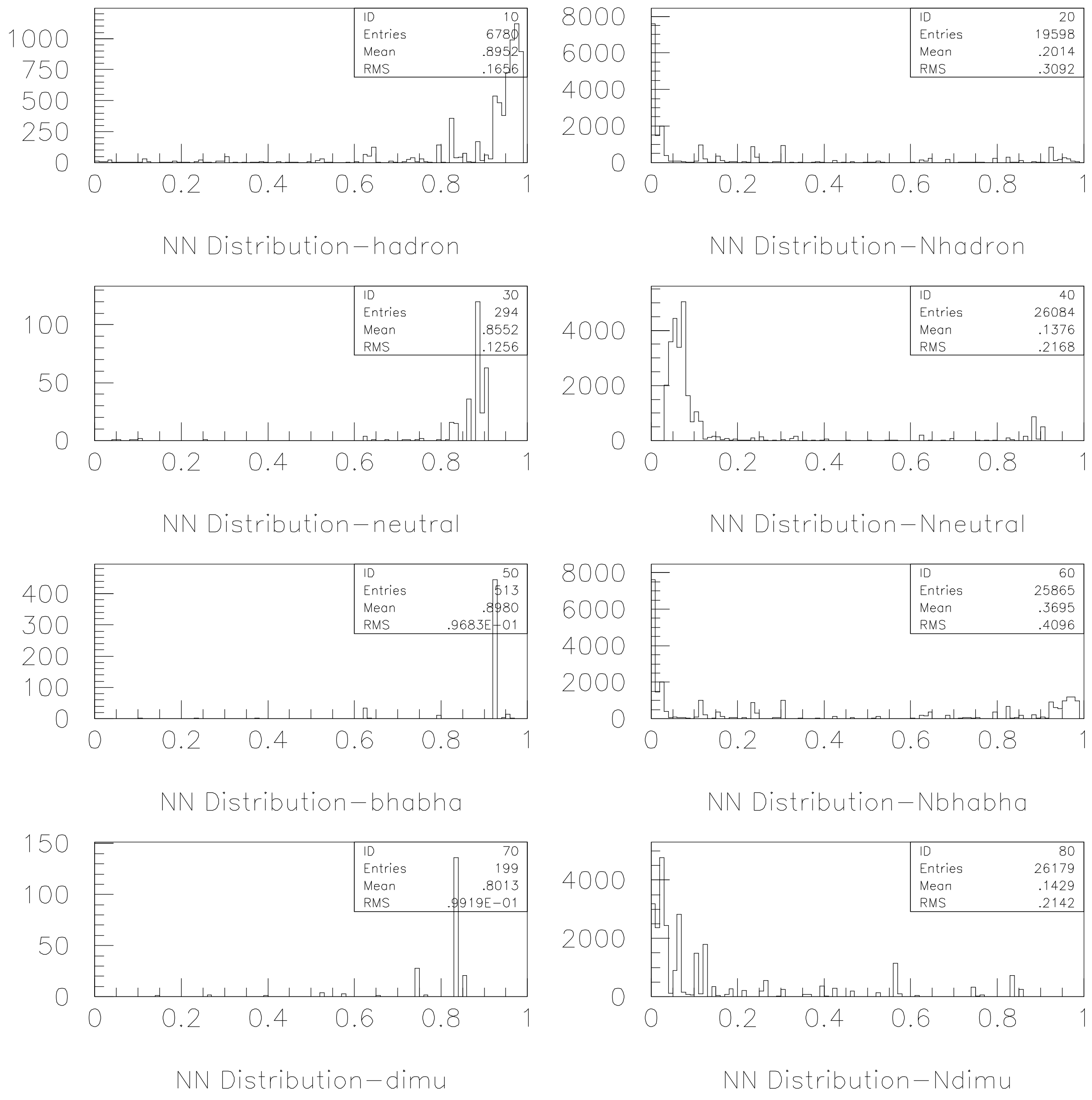

Figure 3: Event Classification of Neural Network

TableII. Performance of Neural Network at J/ψ Energy

RUN	G_{had}(%)	G_{bb}(%)	G_{neu}(%)	G_{du}(%)	S_{bk}(%)	S_{tot}(%)
Run11991	98.5 8888/9026	100 687/687	100 271/271	96.7 236/244	48.7 9634/19771	32.1 9634/29999
Run12060	98.1 9363/9541	99.9 742/743	98.2 217/221	95.0 304/320	47.9 8812/18392	30.2 8812/29217
Run12061	98.1 9200/9381	100 717/717	99.1 212/214	95.3 301/316	47.2 8896/18866	30.2 8896/29494
Run12062	98.0 9478/9675	100 724/724	99.6 231/232	98.4 313/318	48.0 9150/19049	30.5 9150/29998

TableIII. Performance of Neural Network at R Scan Energies

RUN	G_{had}(%)	G_{bb}(%)	G_{neu}(%)	G_{du}(%)	S_{bk}(%)	S_{tot}(%)
2.6Gev	98.0 5532/5649	99.0 28177/28433	– –	81.3 1257/1547	– –	– –
3.2Gev	99.1 1782/1798	99.4 9397/9457	– –	84.4 341/404	– –	– –
3.4Gev	99.7 1645/1651	99.3 11039/11115	– –	80.0 323/404	– –	– –
3.55Gev	99.7 2152/2159	99.7 15425/15478	– –	83.0 469/565	– –	– –
4.6Gev	99.5 966/971	98.9 3221/3257	– –	71.1 167/235	– –	– –
5.0Gev	99.4 1400/1408	99.3 4167/4199	– –	66.1 205/310	– –	– –
4.02Gev Run12711	67.6 433/641	100 206/206	98.9 85/86	92.3 12/13	80.0 23233/29061	77.5 23233/30007
4.02Gev Run12712	72.4 436/602	100 162/162	100 85/85	90.9 10/11	79.9 21523/26942	77.4 21523/27802

Fig. 3 shows the generalizations of neural network of four event types and their opposites. Nhadron means non-hadron, same as Nbhabha, Ndimu and Nneutral.

Some discussions and explanations about the result:

- The generalizations at resonant energies are acceptable, especially at high luminosity. Since at these energies, the cross sections are great and the rate of good event are high, the interference of background is relatively low, so it is easy to select the clear good events through software and the generalizations of neural network are good.
- The generalizations to different data sets at same J/ψ energy and to data sets at different energies are satisfactory.
- At non-resonant energies, the situation is just opposite. It is very difficult to select the clear good events with the same selection conditions at resonant energies and the generalizations of neural network seem not satisfied. But if strict conditions are used, relatively clear good events will be gotten and then the generalizations seem better. The difference could be seen in Table III. The good events at energies from 2.6Gev

to 5.0Gev are selected with relatively strict conditions. The events in Run12711-12 at energy 4.02Gev are with loose conditions. Under strict selection conditions, the generalizations also could be acceptable.

- The rejected hadron events by neural network in Run12711 had been hand scanned, 99%(206/208) of them are found as real background. Thus the conclusion is that the selection conditions seem too loose.
- The training of neutral event and dimu are not complete because of few training samples and their insensitive to current trigger conditions mentioned above. That will also affect the generalizations of neural network to neutral event and dimu.
- The thresholds of the four event types should be adjustable. Reducing the thresholds could improve the good event retaining performances of neural network. But the thresholds should be decided according to the requirements of experiments finally.
- A prescaler will be used to save some events rejected by neural network. These events will be used for further study to see how many and what kinds of events will be lost.

5 Summary

From the result and analysis above, neural network is thought of beneficial in BES trigger system, especially when the luminosity of BEPC increases a lot in future. Its superior characteristic could not be shown under current situation. But it still could be used in on-line decision and off-line analysis.
We should mention here that input feature vector of neural network is the fourteen binary trigger conditions, which could not contain sufficient informations of events and it would limit the generalizations of neural network. With complete event data, neural network classification can be done better in online software.

6 Hardware architecture

Fig. 4 shows the hardware architecture of Neural Network based on DSP expected for BES trigger system.
Here DSP performs the event selection, output the selected event types and give level 3 signal to main trigger controller. CAMAC system reads the output of DSP, sets the thresholds of neural networks and sends test vectors to DSP to check system performance. It also controls the status of DSP.
Table IV and V are Trigger Condition Table for J/ψ and R Scan datataking used currently.

Table IV. Trigger Condition Table for J/ψ

\ TYPE CONDTN\	BHABHA	CHAR	2–MU	CHAR2	NEUTRAL	COSMIC	ESC	BB2
active?	N	Y	N	Y	Y	N	N	N
TOF B–B	–	–	Y	–	–	–	–	Y
$N_{TOF} \geq 1$	–	Y	Y	–	–	–	–	–
$N_{TOF} \geq 2$	–	–	–	Y	–	–	–	Y
RADIAL.	–	–	–	–	Y	–	–	–
$N_{VC} \geq 1$	–	Y	Y	–	–	Y	Y	–
$N_{ETOF} \geq 1$	Y	–	–	–	–	–	–	–
END B-B	–	–	–	–	–	–	Y	–
$N_{TRK} \geq 1$	–	Y	–	–	–	–	–	Y
$N_{TRK} \geq 2$	–	–	–	Y	–	Y	–	–
$N_{TRK} \geq 4$	–	–	–	–	–	–	–	–
MU–OR	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–
ESC-E_{tot}	–	–	–	–	–	–	Y	–
E_{tot-l}	–	Y	–	–	–	–	–	Y
E_{tot-h}	–	–	–	–	Y	–	–	–

Table V. Trigger Condition Table for R-Scan

\ TYPE CONDTN\	BHABHA	CHAR	2–MU	CHAR2	NEUTRAL	Eneutral	ESC	BB2
active?	N	Y	N	Y	Y	N	Y	N
TOF B–B	–	–	Y	–	–	–	–	Y
$N_{TOF} \geq 1$	–	Y	Y	–	–	–	–	–
$N_{TOF} \geq 2$	–	–	–	Y	–	–	–	Y
RADIAL.	–	–	–	–	Y	–	–	–
$N_{VC} \geq 1$	–	Y	Y	Y	–	–	Y	–
E_{radial}	Y	–	–	–	–	Y	–	–
ETOF B-B	–	–	–	–	–	–	Y	–
$N_{TRK} \geq 1$	–	Y	–	–	–	–	–	Y
$N_{TRK} \geq 2$	–	–	–	Y	–	–	–	–
$N_{TRK} \geq 4$	–	–	–	–	–	–	–	–
MU–OR	–	–	–	–	–	–	–	–
E_{trk}	–	–	–	–	–	–	–	–
ESC-E_{tot}	–	–	–	–	–	Y	Y	–
E_{tot-l}	–	Y	–	–	–	–	–	Y
E_{tot-h}	–	–	–	–	Y	–	–	–

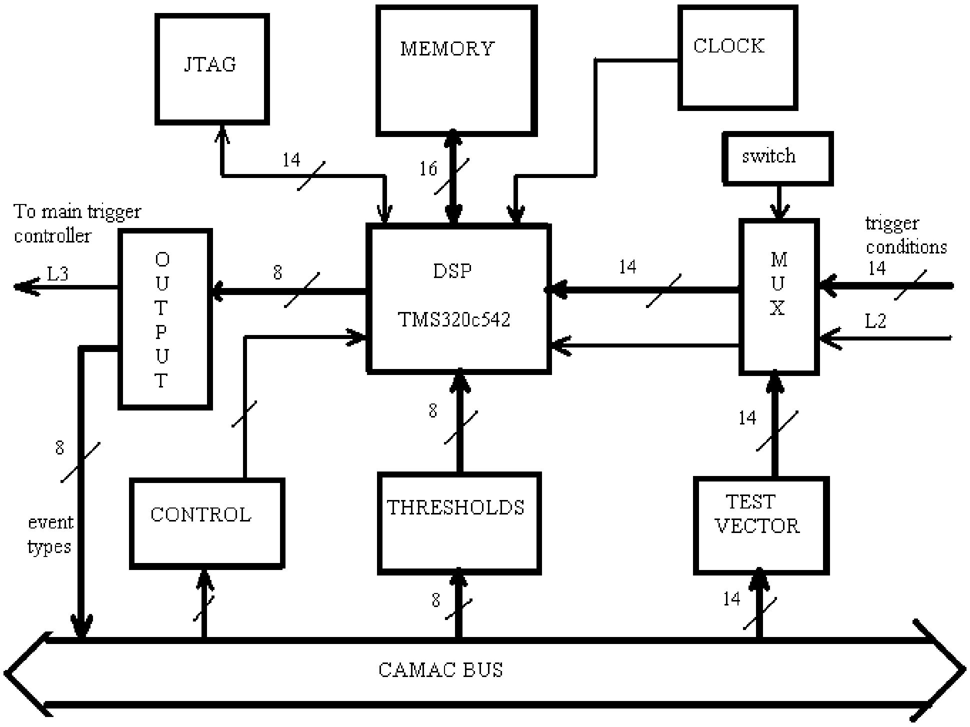

Figure 4: Hardware architecture of NN based on DSP

References

[1] Y.N. Guo et al., The Trigger of Beijing Spectrometer, High Energer Physics and Nuclear Physics, (1990)Vol 15, 01

[2] C.Baldanza et al., Results From A Neural Trigger Based on The MA16 Microprocessor, The 4-th Int. Workshop on Software Engineering, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, April 4-8,1995

[3] Zhijin Jiang et al., The use of the discriminant analysis method for e π μ separation in BES, Nuclear Instruments and Methods in Physics Research A 345(1994) 541-548

[4] SPSS(McGraw-Hill, New York, 1975)

[5] JETNET3.0-A Versatile Artificial Neural Network Package, LU TP 93-29, CERN-TH.7135/94,1993