Ring Mapping For a Neural Electron-Jet Separation Using Multi-layered and Multi-granularity Calorimeters

A. Rabello[1] e J.M. Seixas[1]

[1] *COPPE/EE/UFRJ, {rabello,seixas}@lps.ufrj.br*

CP 68504, Rio de Janeiro, 21945-970 Brazil

Abstract

This work describes a neural network approach to the electron/jet discrimination at the Second Level Trigger of ATLAS. The studies are based on a fast compression algorithm using the energy deposition profile of particles in the calorimeters. A comparison between this approach and the current (classical) one for the same task points out better results. The proposed method achieves 95.67% of efficiency in discriminating electrons with less than 4.26% of jets being misclassified as electrons.

1 Introduction

In 2005, the European Centre for Particle Physics (CERN) will begin to operate the LHC experiment. The LHC will investigate the existence of the Higgs boson and, for that, around 40 million interactions will be generated per second. Around one of the collision points of LHC, a complete set of sophisticated detectors will be placed forming the ATLAS (A Toroidal LHC ApparatuS) laboratory. The volume of data for each interaction is huge and therefore, there is no way of keeping data from all events for off-line analysis. To solve this problem, the ATLAS experiment will count on a very efficient triggering system that will perform real-time separation between events that represent ordinary physics from events that may represent Higgs decays.

In order to be fast enough, the trigger system is divided into 3 levels connected in a cascading structure. In this way, the first level (L1) trigger will have the responsibility of reducing the total event rate of the ATLAS experiment. The second level (L2) is responsible for achieving further reduction of the output rate of the first level and the same is applicable to the third level trigger. If one puts that in numbers, the L1 must reduce the event rate from 40 MHz to around 100 kHz, the L2 shall reduce this rate to around 1 kHz and the third has to output no more than 100 events per second.

To accomplish the task, L1 trigger runs basically on low-programmability hardware, searching for events on a subset of ATLAS detectors. This *search* is performed in a way that the first level is able to detect Regions of Interest (RoIs), indicating to the next triggering levels the regions that were identified as being excited by the sub-products of the current collision.

The L2 uses this information in order to localize its efforts, identifying each object and comparing the set of objects at an event to a set of target signatures that may indicate a Higgs decay. If any signature matches to the newly found objects, the event is accepted, otherwise it is rejected and deleted from the buffers. In case of acceptance of L2, the event is passed to the third level trigger for further depuration.

Upon the identification of RoIs found by L1, L2 will be able to use all ATLAS sub-detectors at full granularity. In practice, however, the identification of some objects is done using only one or two detectors. In fact [4, 3], because of the high event rate (and considering that each event produces in average 5 objects), it's desirable that the second level trigger has the capability to run a smart sequential algorithm that uses the smallest amount of data in order to reject or accept an event. Thus, only the minimum amount of data is transfered into the system from the buffers, making the whole process fast and efficient.

One of the most representative objects of Higgs decays are electrons. Electrons are mainly detected using the calorimeter cells. However, they are commonly faked by jets. Using the full granularity of the calorimeter, L2 algorithms may further reduce L1 event rate about 7 times [4].

Calorimeters The ATLAS laboratory contains 2 different types of calorimeters: electromagnetic (EM) and hadronic (HAD). The EM calorimeter is composed of 4 layers with different granularities. The HAD calorimeter is also composed of 3 layers, but with the same granularity. Besides that, each of these parts can be sub-divided into other two: the barrel and the endcap. The region between the barrel and the endcap contains a gap, for instrumental purposes. At such region, usually discriminating algorithms loose performance[4].

Classical e^-/Jet Discrimination The classical approach that has been followed by ATLAS is based on the extraction of four highly discriminating physics quantities for each object signaled by L1 as an electron: the energy containment on the electromagnetic and hadronic sections of the calorimeter, the object width and the energy spreading on the first electromagnetic layer. Cuts are set empirically in order to optimize the discrimination efficiency using such object features.

Taking into account the high luminosity of ATLAS ($10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$) and also that the background rate of jets is around 25kHz, the electron efficiency obtained with such classical approach is $(94.1 \pm 0.3)\%$ for a background rate of $(3.55 \pm 0.16)kHz$. Such result does **not** consider objects centered on the region that barrel and end-cap superimpose, but considers pile-up effects.

2 The Neural Approach

Data Each RoI represents, physically, a square region of size 0.4 by 0.4 over the $\eta \times \phi$ plane (pseudo-rapidity versus angle) and is composed of many calorimeter layers that superimpose over the ϕ direction. Each layer has a different function in which detection of particles is concerned and has a specific granularity.

The simulation data set used in this work is composed of 600 RoIs corresponding to electrons and 3600 RoIs to jets without any pile-up. The RoIs that correspond to jets were chosen according to first level trigger acceptance (refer to [4] for details).

ANN In this work, we propose that the electron/jet separation on L2 be performed by an artificial neural network (ANN). ANNs are very good pattern recognizers and have many interesting features for this kind of environment [1].

In order to perform the task, we used a fully-connected neural network trained with the back propagation method [2]. Each hidden and output neuron had as activation function a hyperbolic tangent (tanh). Since both training and testing phases become slow when the dimensionality of the data input space is high (as it is the case, due to the fine granularity of the calorimeter system), we compacted the input data (around 1,000 cells) by exploitation of its inherent isotropy. For each RoI, this type of compaction adds-up the values of energy deposited over each calorimeter cell that is around the cell with the greatest value of energy deposited (energy peak), forming concentric rings at each layer. Such values are fed into the nodes of the neural discriminator. Figure 1 shows how this preprocessing was accomplished [4]. Because of saturation, the resulting sums were normalized using the sum of all calorimeter cells in the RoI. This way, energy variations from RoI to RoI are ignored by the process and discrimination will be performed based only on the profile of the energy deposition.

The ring structure maps the 1008-dimensional space onto a 67-dimensional space. Also, for simplicity, we decided to ignore RoIs that cover the gap between the end of barrel and the beginning of the end-cap of the EM calorimeters. Such region is characterized by variable granularity within the same layer, making the discrimination even harder. This leads-us to a 58-dimensional space, in which the discriminator was developed.

One also must define the number of neurons at the hidden layer. In fact, this number is deeply related to the number of discriminating variables that the process under investigation might require. We decided to perform tests with 20, 17, 14 and finally 10 neurons on such layer of neurons.

Results In order to achieve maximum electron-jet discrimination efficiency, the stop condition was defined by the discrimination efficiencies of both classes [2]. This way, [5] proposes the following criteria for neural training: $SP =$ (Sum of Efficiencies) $\times$ (Product of Efficiencies).

Ideally, the final value of SP should be, in this case, 2. In practice, however, this value is reached only on pathological cases. The data sets were divided

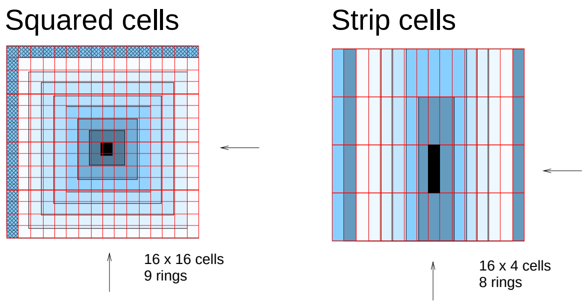

Figure 1: This figure exemplifies ring construction for two type of layers from the calorimeter. The energy peak is marked in black and indicated by arrows.

Learn. Rate	# hidden	effic (electrons)	effic (jets)	SP
0.02	20	91.33%	97.23%	1.6746
0.03	20	96.67%	93.47%	1.7180
0.1	17	95.33%	95.13%	1.7273
0.1	**17**	**95.67%**	**95.90%**	**1.7577**
0.1	17	95.67%	95.13%	1.7364
0.1	**14**	**95.00%**	**95.90%**	**1.7531**
0.1	14	95.33%	95.35%	1.7334
0.1	**10**	**95.67%**	**95.74%**	**1.7531**
0.03	10	95.33%	94.80%	1.7183
0.03	10	96.33%	94.80%	1.7454

Table 1: The performance (test set) of the neural classifier when the number of hidden neurons is varied. The best results among all are marked in bold face.

in two (train and test), in order to guarantee the correctness of the results. Since we had much less electrons than jets, and training is performed by a random walk over the data sets, we replicated the electron training set 6 times before mixing it with the jet set. This avoids the neural network developing any biasing for jet recognition, and producing as consequence, a low electron recognition efficiency.

3 Summary and Outlook

At this work, we proposed a neural discriminator that works using only the data from a subspace of preprocessed calorimeter cells. For training such network we used a special parameter to measure its discrimination efficiency, stopping when such efficiency reached its maximum for both training and test sets.

As shown on Table 1, even using such *subset* of the input data, the efficiency for electron recognition was improved almost 2% with a much lower background rate (considering the initial background rate of L2 to be 25kHz, such efficiency for jet recognition will give us 1.025+-0.014kHz). If we in turn aimed to reach the same background rate of the classical approach, the efficiency for electron recognition achievable by the neural discriminator is about 99.67+-0.33%. This is almost 6% above the classical approach results, in an environment where electrons may mean very rare physics. Other works suggest that adding pile-up effects over the events decreases about 2 percentual points on these results [6]. This study shall be conducted soon.

One has also to consider sequentiality when designing the trigger. Using neural networks this can be easily achieved by developing specialized neural networks that can work on a layer basis. With that, one can keep the load rates of L2 to a minimum while pushing efficiency to its maximum.

The current work is also being implemented on a fast DSP platform using the C programming language. This study will characterize the algorithm execution time, which is very important on the L2 environment.

References

[1] R.K. Bock AND J. Carter AND I.C. Legrand, ATLAS/DAQ Note 11: What can Artificial Neural Networks do for the Global Second Level Trigger, **CERN, March/1994**.

[2] Simon Haykin, Neural Networks: A comprehensive foundation, Prentice Hall, **1999**.

[3] ATLAS Collaboration, ATLAS DAQ, EF, LVL2 and DCS Technical Progress Report, **CERN 1998**.

[4] ATLAS/DAQ Note 2: S. Gonzalez AND T. Hansl-Kozanecka AND M. Wielers, Selection of high-pT electromagnetic clusters by the level-2 trigger of ATLAS, **CERN, February/2000**.

[5] J.M. Seixas AND I.F. Isidro, A Neural Electron/Jet Discriminator Based on a Highly-Segmented Calorimeters, Artificial Intelligence in High-Energy and Nuclear Physics (AIHENP), Creta/Greece, **1999**.

[6] J.M. Seixas AND L.P. Calôba AND M.N. Souza AND A.L. Braga AND A.P. Rodrigues AND H. Gottschalk, Neural Networks Applied to a Second Level Trigger System Based on Calorimeters, Third International Workshop on Software Engineering, **1993**, pp. 353-358, Artificial Intelligence and Expert Systems for High Energy and Nuclear Physics, Oberammergau, Germany.