PROCEEDINGS OF SCIENCE

Electron/Jet Neural Discrimination based on Nonlinear Independent Components for ATLAS Second-Level Trigger

Eduardo F. Simas Filho[1,2], **José M. Seixas**[*1], **Danilo E. F. Lima**[1], **and Luiz P. Calôba**[1]

[1] *Signal Processing Laboratory, COPPE/Poli, Federal University of Rio de Janeiro, Brazil.*

[2] *Federal Institute of Education, Science and Technology of Bahia, Brazil.*

E-mail: {esimas,seixas,daniloefl,caloba}@lps.ufrj.br

The Large Hadron Collider at CERN is designed to collide proton bunches at 40 MHz in the ATLAS detector. This event rate must be reduced to about 200 Hz for storage by selecting the most interesting events for physics analysis. This will be achieved in ATLAS with a three level trigger system. This work proposes the application of the Nonlinear Independent Component Analysis model in the ATLAS second-level trigger to estimate underlying factors of calorimeter information. As a pre-processing step, the calorimeter Region of Interest data are transformed into concentric energy deposition rings. In order to take advantage of the full detector segmentation, the feature extraction procedure is performed at calorimeter layer level. The independent features estimated from each layer were concatenated into a single feature vector, which was used to feed the input nodes of a neural network classifier for highly-efficient electron identification. The proposed approach achieved 97.0% of electron identification for 7.7% of background noise (QCD Jets) acceptance.

13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research, ACAT2010
February 22-27, 2010
Jaipur, India

[*] Speaker.

PoS(ACAT2010)052

1. Introduction

The Large Hadron Collider (LHC) will reach a previously unexplored energy range and produce massive amounts of information, as it collides two 7 TeV bunches of protons every 25 ns [1]. Despite this very high interaction rate, the relevant physics channels will be very rare. For example, the Higgs signatures are expected at 0.01 to 0.1 Hz depending on its mass [2].

ATLAS is a general purpose detector of LHC [3]. Considering detector segmentation and LHC interaction rate, approximately 60 TB of information will be produced per second. In order to retain the relevant physics and discard, as much as possible, the background noise, a three-level on-line trigger is implemented in ATLAS [2]. The first-level trigger (L1) receives the full event rate (approximately 40 MHz) and needs to produce a decision in less than 2.5 μs, reducing the event rate to 75 kHz. The second-level (L2) analyses only information indicated in the so called Regions of Interest (RoI), which are the parts of the detector, marked by L1, in which relevant signals are found. Finally, the third-level (Event Filter - EF) uses full event information and selection algorithms similar to the ones used in off-line reconstruction. The target output rates and latency for L2 and EF are, respectively, (3 kHz / 40 ms) and (200 Hz / few seconds).

Calorimeters play an important role on ATLAS on-line triggering. ATLAS calorimeters are split in depth into seven layers, four in the electromagnetic section (PS, E1, E2 and E3) and three in the hadronic one (H0, H1 and H2). Calorimeter layers have different granularities and a RoI typically comprises hundreds of calorimeter cells.

Among the interesting trigger objects for ATLAS, electrons are involved in one of the cleanest channels for standard model Higgs boson production [4]. QCD jets fake electron triggers as they may deposit energy in the detector similarly to electrons.

This work proposes an electron neural discriminator for ATLAS L2 trigger based on a combination of topological (ring-like, to be described further) pre-processing of the calorimeter RoI and segmented (layer-level) feature extraction through nonlinear independent component analysis (NLICA) [5]. Calorimeters are designed to be linear detectors, but, in practical designs, slight nonlinearities may arise due to sensor saturation, light attenuation (in scintillating calorimeters) and energy leakage [6]. Thus, here it is investigated whether the NLICA model is able to reveal, from ring-formatted calorimeter data, discriminating features for efficient electron identification. The trigger decision is based on a neural network, which is fed from the estimated independent components.

2. The Proposed Signal Processing Chain

A proposal for electron/jet discrimination at ATLAS L2 trigger is shown in Fig. 1, The feature extraction procedure (comprising both ring construction and NLICA) is performed in a segmented way (at the layer-level). The relevant features obtained from each calorimeter layer are concatenated into a single feature vector used as inputs for the hypothesis testing procedure (which is performed through a neural network classifier). In the following subsections, the individual blocks of the proposed signal processing chain will be detailed.

2.1 Signal Pre-processing

As proposed in [7], here the calorimeter information is formatted into concentric rings through the following procedure (see Fig. 1): 1. At each calorimeter layer, the most energetic cell is considered as the first ring; 2. The next rings are sequentially formed around the first one; 3. Ring signals are obtained by summing the energy of the cells belonging to a given ring. This procedure makes the signal description independent from the impact point, compacts RoI information (which is now described by a total of 100 rings) and preserves the energy deposition profile information. The number of rings estimated for each layer are PS-8, E1-64, E2-8, E3-8, H0-4, H1-4 and H2-4.

2.2 Nonlinear Independent Component Analysis

The nonlinear independent component analysis (NLICA) model considers that an observed signal vector $\mathbf{x} = [x_1, x_2, ..., x_N]^T$ is generated through a nonlinear mapping $F(.) : \mathbb{R}^N \rightarrow \mathbb{R}^N$ of unknown and statistically independent source signals $\mathbf{s} = [s_1, s_2, ..., s_N]^T$:

$$\mathbf{x} = F(\mathbf{s}) \tag{2.1}$$

NLICA algorithms search for an inverse mapping $G : \mathbb{R}^N \rightarrow \mathbb{R}^N$: $\hat{\mathbf{s}} = G(\mathbf{x})$, so that the components of the estimated source vector $\hat{\mathbf{s}}$ are statistically independent. Several models were proposed to solve the NLICA problem, among them, the post-nonlinear (PNL) mixture model has met significant practical applicability [5].

2.2.1 Post-nonlinear ICA model

Post-nonlinear (PNL) mixtures arise whenever, after a linear mixing process, sensors present nonlinear behavior [8]. As illustrated in Fig. 2, the observed signals can be expressed as:

$$\mathbf{x} = F(\mathbf{As}), \tag{2.2}$$

where $F(.) = [f_1(.), f_2(.), ..., f_N(.)]^T$, and the inverse model is: $\hat{s}_i = \sum_{j=1}^{N} b_{ij} g_j(x_j)$.

In this work, as proposed in [8], the inverse nonlinear functions $g_j(.)$ are modeled by a two-layer MLP (multi-layer perceptron) neural network [9]: $g_j(x_j) = \sum_{h=1}^{N_H} \sigma_h \tanh(\omega_h x_j + \gamma_h)$, where N_H is the number of hidden neurons used to estimate the nonlinearity. The parameters b_{ij}, σ_h, ω_h and γ_h are adjusted in the training procedure aiming at independent component estimation. The parameter N_H needs to be chosen guided on prior information on the problem, as it is closely related

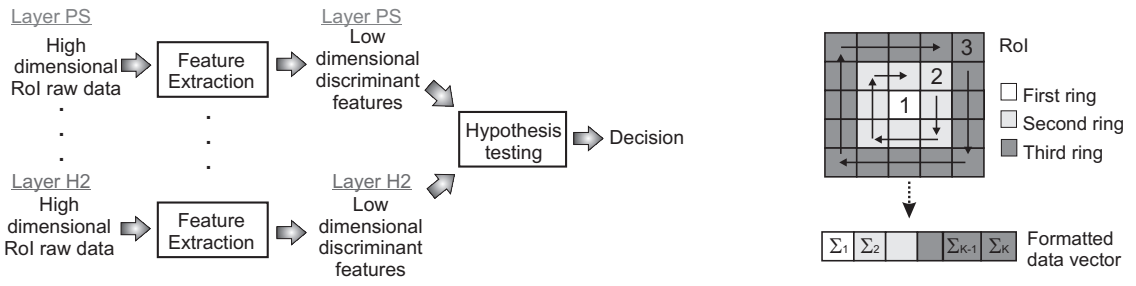

Figure 1: The proposed (segmented) signal processing chain (left) and the ring building procedure (right).

to the allowed degree of nonlinearity in each $g_j(.)$. By increasing N_H, the PNL model is able to search for stronger nonlinearities, on the other hand, using low N_H (Ex. $N_H = 2$ or $N_H = 3$) only mild nonlinear functions $g_j(.)$ can be modeled.

During the training procedure (see Fig. 2), the mutual information $I(\hat{\mathbf{s}})$ is computed and used as statistical independence measure [5]: $I(\hat{s}_1, \hat{s}_2, ..., \hat{s}_N) = \sum_{i=1}^{N} H(\hat{s}_i) - H(\hat{\mathbf{s}})$, where $H(u) = \sum_i p(u_i) \log[p(u_i)]$ is the entropy. By minimizing $I(\hat{\mathbf{s}})$, $\hat{s}_i$ converge to independent components.

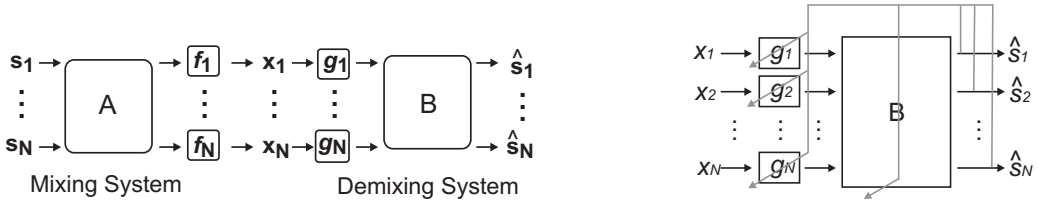

Figure 2: PNL mixing and demixing models (left) and estimation procedure (right).

2.3 Supervised Neural Discrimination

An additional supervised neural network (MLP architecture) was designed for the hypothesis testing phase to perform particle identification. The network comprises one hidden layer and a single output neuron, both using hyperbolic tangent activation functions. The training routine was the traditional error back-propagation algorithm [9]. A target mean square error was used to stop the training procedure. In order to define the optimum number of neurons in the hidden layer (N_{Clas}), exhaustive testing was performed varying N_{Clas} from one to 20.

3. Results

The database used comprised $\approx$ 160,000 electron (with energy between 7 and 80 GeV) and $\approx$ 100,000 QCD dijet events generated through Monte Carlo simulations. This data set was pre-filtered by simulated L1 operation [10]. Approximately 140,000 electron and 13,000 false electron (jet) RoIs reached L2 and were used to design and test the proposed discriminator. The available dataset was split into training (NLICA extraction), validation (neural classifier training stop) and testing (generalization) sets.

In this work, the Receiver Operating Characteristic (ROC) [11] and the SP index were both used for performance comparison. The SP is defined as [7]:

$$SP = \sqrt{\frac{Ef_e + Ef_j}{2} \times \sqrt{(Ef_e \times Ef_j)}}, \tag{3.1}$$

where $Ef_e = P_D$ and $Ef_j = 1 - P_F$ are respectively the detection efficiencies for electron and jet. SP is computed for different threshold levels and the one that provides maximum SP is used.

Figure 3 illustrates the ROC curves obtained from training neural classifiers (9 hidden neurons) using as inputs the ring-signals (Neural Ringer) and the nonlinear independent components estimated through the PNL model (using 2, 3 and 4 neurons to estimate each nonlinear function). It can be observed that PNL feature extraction (using only 2 neurons to estimate each function) produced better discrimination performance when compared to standard Neural Ringer. By increasing

the number of neurons the efficiency decreases. This result indicates that the optimum functions $g_j(.)$ (for nonlinearity estimation purposes) represent mild nonlinearities (slight deviations from full linearity), as was expected for well designed calorimeter signals. The neural classifier outputs obtained after PNL feature extraction (2 neurons) are illustrated in Figure 3. Table 1 provides performance comparison considering the maximum SP index and false alarm probabilities. Approximately 1 percent point of false alarm reduction was achieved through mild PNL estimation when the detection efficiency is kept fixed at 97 %.

Neural Ringer and PNL-based discriminators are compared considering electron efficiency and false electron acceptance in Figure 4, for different values of energy and pseudo-rapidity (η). ATLAS calorimeters present discontinuities around $|\eta| = 1.5$ (known as cracks) where power and signal cables pass through. Thus, at the crack region signatures of particles are not accurately measured (which implies reduced trigger efficiency). Considering different energies, both discriminators present similar performance (with a slight variation between different bins). As a function of η, the PNL feature extraction produced poorer performance around the crack, but compensates for this with higher efficiency out of the crack region. In addition, PNL ICA reduced the fake electron acceptance in almost all bins.

In order to verify whether the L2 processing time requirement is fulfilled, the proposed algorithm was implemented in the actual trigger platform [12] as a sub-routine of the ATLAS baseline electron/jet discriminator design (T2Calo) [4]. Restricted to calorimeter trigger, the full execution requires six different processing tasks (see Table 2): 1. Information required for T2Calo execution is transfered from the temporary memories; 2. T2Calo computes parameters such as the refined RoI center and the event energy; 3. Part of the information required for the Neural Ringer is not requested by T2Calo on step 1, so it is accessed in this stage; 4. In this stage, the ring construction procedure is executed; 5. The following step comprises ring signal normalization;

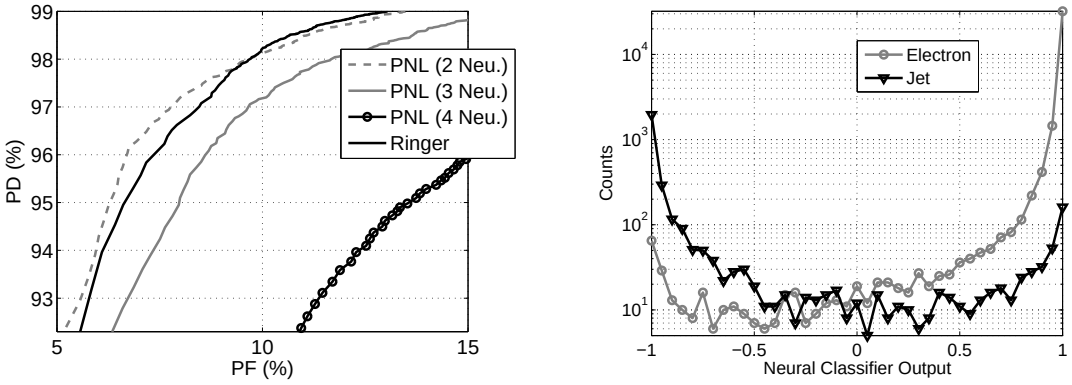

Figure 3: ROC curves (left) and Neural Classifier Outputs (right).

Table 1: Performance comparison between different discriminators.

Discriminator	Neural Ringer	PNL (2 neurons)	PNL (3 neurons)	PNL (4 neurons)
Best SP×100	94.35	94.70	93.70	90.83
P_F for $P_D = 97\%$	8.67±0.20	7.69±0.35	9.67±0.38	17.39±0.40

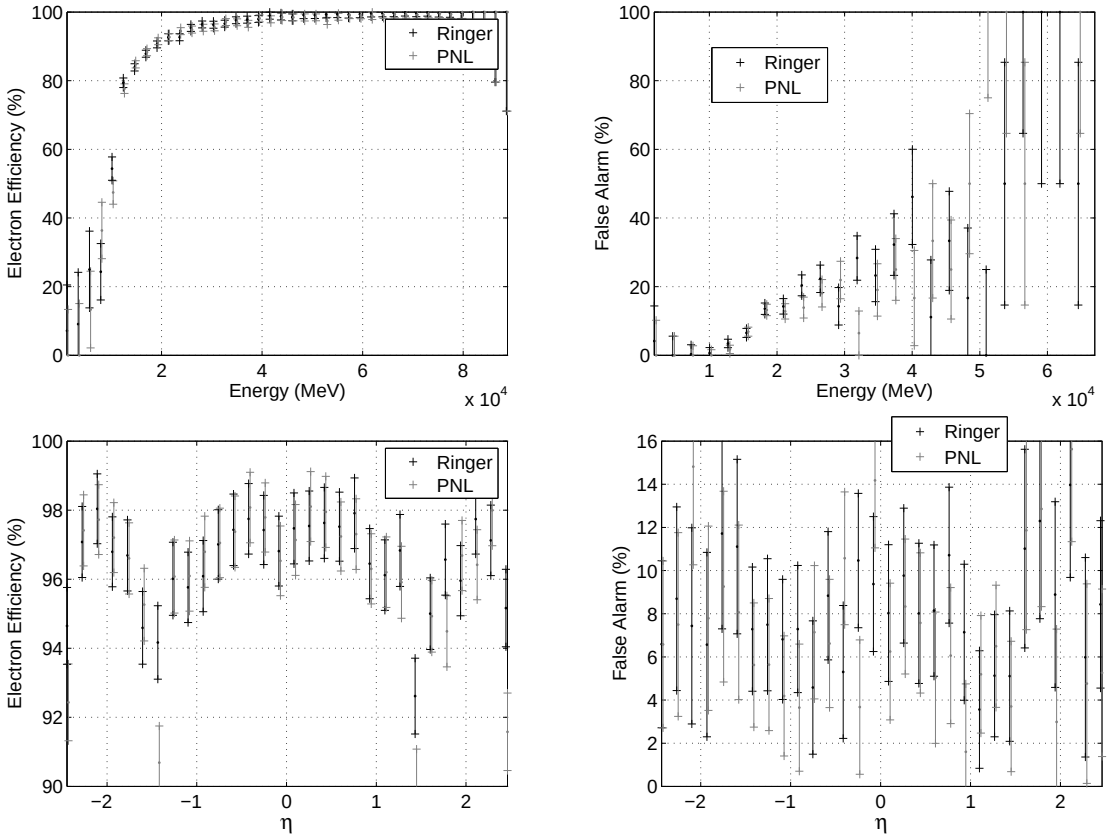

Figure 4: Electron efficiency (left) and false acceptance (right) for different values of energy and η, the error bars were estimated through the inverse of the number of events (1/N).

Table 2: On-line execution times.

Step	**Time** (ms)	% **of Total**
Region Selector - T2Calo	0.4927 ± 0.0787	39.9
Pre-processing - T2Calo	0.1408 ± 0.0148	11.4
Region Selector - Ringer	0.4375 ± 0.0996	35.4
Pre-processing - Ringer	0.0986 ± 0.0165	8.0
Normalization	0.0026 ± 0.0015	0.2
NLICA+Neural Classification	0.0630 ± 0.0094	5.1
Total	1.2352 ± 0.1292	100

6. Finally, NLICA and Neural Classification procedures are executed. The average and standard deviation times obtained for each processing task are shown in Table 2. The total time required for the proposed NLICA calorimeter-based L2 trigger is 1.2352 ± 0.1292 ms (which represents approximatelly 3 % of the design L2 latency). In adition to the calorimeter decision, event acceptance/rejection typically requires other ATLAS subdetector information.

4. Conclusions

A segmented feature extraction procedure based on nonlinear independent component analysis (NLICA) and neural classification operating over ring formatted calorimeter regions of interest was proposed for the ATLAS on-line second-level trigger. Experimental results confirmed that the nonlinearities estimated by the model shall be limited to a slight deviation from full linearity, as the calorimeter response is practically linear. The false alarm was reduced to $(7.69 \pm 0.35)\%$ at a detection probability of 97% and the latency requirements for L2 were shown to be satisfied for the proposed method.

Acknowledgements

The authors would like to express their gratitude to CNPq, FINEP, CAPES, FAPERJ and RENAFAE (Brazil) and CERN (Switzerland) for the support of this work. We also thank the ATLAS Trigger/DAQ collaboration for providing the calorimeter data and for fruitful discussions concerning this work.

References

[1] Evans, L., Bryant, P. "LHC Machine". *Journal of Instrumentation*, 3, S08001 (2008).

[2] C. Padilla/ATLAS Collaboration "The ATLAS Trigger System", *Proceedings of the 16th IEEE-NPSS Real Time Conference*, pp. 326-333, Beijing China (2009).

[3] ATLAS Collaboration "ATLAS Experiment at CERN Large Hadron Collider". *Journal of Instrumentation*, 3, S08003 (2008).

[4] Carlo Schiavi/ATLAS Collaboration "Implementation and Performance of the High Level Trigger Electron and Photon Selection for the ATLAS Experiment at the LHC", *IEEE Transactions on Nuclear Science*, vol. 53, n. 5, pp. 1424-1429 (2006).

[5] Almeida, L. B., *Nonlinear Source Separation*, Morgan and Claypool (2006).

[6] Wigmans, R.: Calorimetry: Energy Measurement in Particle Physics. Clarendon Press (2000).

[7] Anjos, A. et al. "Neural Triggering System Operating on High Resolution Calorimetry Information". *Nuclear Instruments and Methods in Physics Research*, 559, pp.134-138 (2006).

[8] Taleb, A. and Jutten, C. "Source Separation in Post-nonlinear Mixtures". IEEE Trans. on Signal Processing, no. 47, pp. 2807-2820 (1999).

[9] Haykin, S. *Neural Networks and Learning Machines*, Prentice Hall, 3rd Ed. (2008).

[10] Achenbach, R., et al., "First data with the ATLAS Level-1 calorimeter trigger", *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference*, vol. 1, pp. 1851-1858, Dresden, Germany (2008).

[11] Van Trees, H. L. *Detection, Estimation, and Modulation Theory*, Part I. Wiley (2001).

[12] Zhang, J., et al., "ATLAS Data Acquisition", *IEEE-NPSS Real Time Conference*, vol. 1, pp. 240-243, Beijing, China (2009).