Machine Learning Techniques for selecting Forward Electrons $(2.5 < |\eta| < 3.2)$ with the ATLAS High Level Trigger

Meinrad Schefer
University of Bern, Switzerland
E-mail: meinrad.schefer@cern.ch

Abstract. The ATLAS detector at CERN measures proton proton collisions at the Large Hadron Collider (LHC) which allows us to test the limits of the Standard Model (SM) of particles physics. Forward moving electrons produced at these collisions are promising candidates for finding physics beyond the SM. However, the ATLAS detector is not construed to measure forward leptons with pseudorapidity η of more than 2.5 with high precision. The ATLAS performance for forward leptons can be improved by enhancing the trigger system. This system selects events of interest in order to not overwhelm the data storage with the information of around 1.7 billion collisions per second. First studies using the NeuralRinger algorithm for selecting forward electrons with $2.5 < |\eta| < 3.2$ show promising results. The NeuralRinger using machine learning to analyse detector information to distinguish electromagnetic from hadronic signatures, is being presented. Additionally, its performance on simulated ATLAS Monte Carlo samples in improving the high level trigger for forward electrons will be shown.

1. Introduction

The ATLAS detector at the LHC measures proton-proton collisions from bunch crossings at a rate of 40 MHz [1]. To permanently store all the information taken by ATLAS would fill up the storages quickly and overwhelm them with irrelevant data for analyses. Therefore, a two level trigger system has been introduced to select relevant data out of collision events of interest. The hardware based Level 1 Trigger reduces the rate of 40 MHz of events to 100 kHz which takes it less than 2.5 μs [2]. It defines Regions of Interest (RoIs) for all the events that are passed on to the High Level Trigger (HLT). The software based HLT reduces the rate of events that are being saved in the data storages to around 1.2 kHz which translates to about 1.2 GB/s. It runs reconstruction algorithms on particle candidates identified in the RoIs and classifies them with working points corresponding to their likelihood. Then a filter decides which events are kept and which are discarded according to the composition, amount and likelihoods of the identified particle candidates.

In the last years the LHC has been continuously increasing the amount of collisions per bunch crossing which is mainly achieved by better focusing the bunches. With a higher rate of collisions the production rate of rare and interesting events increases. However, the ATLAS trigger system is more challenged because signals from various collisions are overlapping. To better cope with

this despite a constant bandwidth limitation, the NeuralRinger Algorithm has been implemented for triggering electron candidates in the barrel region more efficiently. In the following chapter its workflow is explained. Thereafter, it will be described how the NeuralRinger can be expanded to more forward regions. Then, the results of first performance studies of the NeuralRinger for forward electrons are shown.

2. The NeuralRinger Algorithm

In 2017 the NeuralRinger Algorithm has been introduced for the HLT in the pseudorapidity region of $|\eta| < 2.5$ to reduce CPU demands [3]. The Neural Ringer is using calorimetric data in the RoIs of electron candidates defined by the Level 1 trigger to build rings and to calculate their energy sums. The ring building process starts with the selection of the highest energetic cell in the first layer of the electromagnetic calorimeter (ECAL) within the RoI which is defined as the first ring. Then, the surrounding eight cells are defined as the second ring which is followed by the next ring consisting of the cells surrounding the previous ring. This is being repeated until a fixed number of rings are built in the first layer. Thereafter, the same procedure is applied for the next layer and continued for all the layers in the ECAL and the hadronic calorimeter. This results in a total number of 100 rings.
Then, for every of the 100 rings the total transverse energy sum of its corresponding cells is calculated. Those 100 energy sums are each normalized by the total transverse energy sum of all the 100 rings. The normalized ring sums are then used as the inputs to a Neural Network (NN) which identifies whether the electron candidate is being kept or discarded. This NN has been trained on large Monte Carlo (MC) generated samples of real and fake electron candidates. Non-prompt electron candidates (fakes) originate from other particles which leave similar energy deposition signatures in the detector. Figure 1 depicts the ring building and the energy sum calculation processes for electrons and fakes.

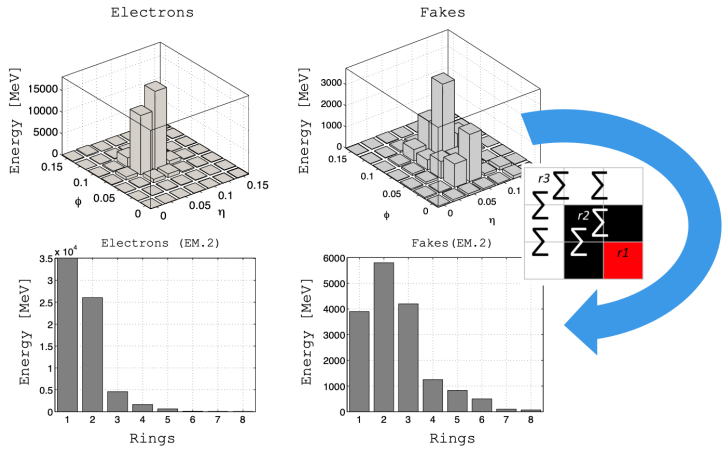


Figure 1: Schematics of the ring building process from the calorimeter cell signatures (top) to the calculation of the energy sums (bottom) for real and fake electron candidates [4]. The energy sums on the bottom show averaged values for the first eight rings and do not specifically show the ones corresponding to the cell energies shown on top.

The NN used by the NeuralRinger algorithm is an ensemble of Multi Layer Perceptrons (MLPs). They are used for different regions in transverse energy (E_{T}) and η. A correction for different pileup scenarios is being applied by replacing the output node transfer function with a linear activation function after all the trainings are completed. Like that the NeuralRinger was not only able to reduce CPU demands but also to reject a significantly higher amount of fake electron candidates without loss of efficiency.

3. Implementation of Rings for Forward Electrons

With the planned upgrades of extending the ATLAS pixel detector for Run 4 up to $\eta = 4$ [5], it becomes more and more important to also improve the electron trigger algorithms in these

forward regions. Many physics studies could benefit from an increased performance of ATLAS within the more forward regions. However, these regions are also more challenging due to the currently reduced tracking information, lesser granularity and more inactive material of the detector.
Extending the NeuralRinger algorithm to more forward regions could be an important step in improving the trigger performance there. First studies on its implementation in the regions of $2.5 < |\eta| < 3.2$ are described here. There the affected calorimeter subdetectors are two layers of Electromagnetic Endcap (EMEC 1 and EMEC 2) and three layers of the Hadronic Endcap (HEC 0, HEC 1&2, HEC 3) calorimeters. Due to the much larger cell size in $\eta \times \phi$ it is not feasible to build the same amount of 100 rings as in the regions of $|\eta| < 2.5$. The cells of both EMEC layers are 0.1×0.1 in $\eta \times \phi$. Therefore, at most four rings can fit in the $2.5 < |\eta| < 3.2$ region which stretches seven cells in η. In the HEC part with cells of 0.2×0.2 in $\eta \times \phi$ only two rings can fit. For these studies it has been decided to build four rings in each of the EMEC layers and two rings in each of the HEC layers with 14 rings in total. Rings that would reach outside $2.5 < |\eta| < 3.2$ are being cut at the borders of the region for these initial studies. Figure 2 shows example rings in the EMEC and HEC layers.

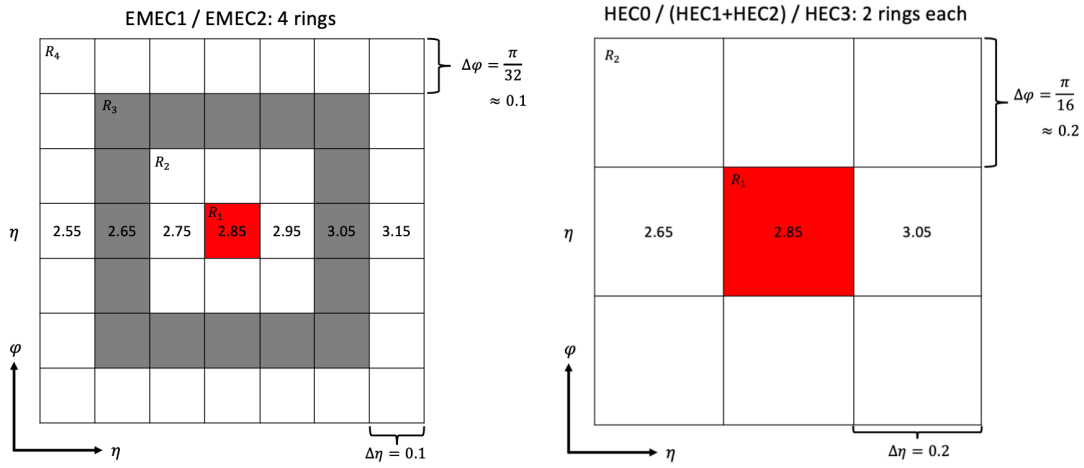


Figure 2: Schematics of the ring building process in the $2.5 < |\eta| < 3.2$ region used in this study. On the left side rings are shown for the EMEC layers and on the right for the HEC layers where for both cases the highest energetic cell (red) is in the middle of the region.

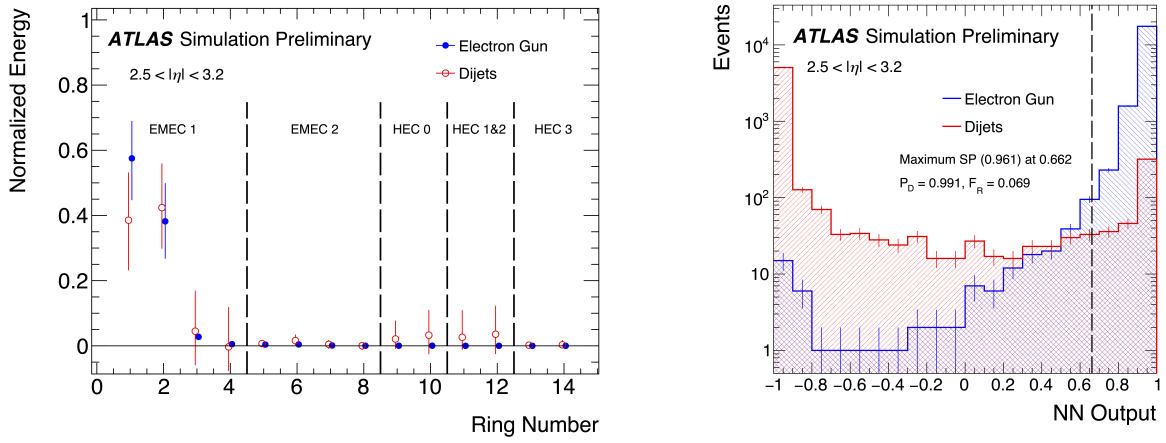
4. Neural Network Training and Results

For these initial studies electron gun samples have been used to provide real electron candidates as signal and MC dijet samples have been used for fake electron candidates as background. In this case, the electron gun provided MC simulated electrons originating in a certain range around the center of the ATLAS detector with transverse energies between 20 and 100 GeV going into regions of $|\eta| > 2.5$. No proton collisions are simulated for the electron gun samples. The candidates in both samples are required to satisfy the cut-based 'loose' identification criteria. For each of the selected candidates 14 forward rings were built and their ring energy sums were calculated and normalized as it has been described in section 2 and section 3. Then, a set of MLPs with one hidden layer has been trained in order to distinguish signal from background candidates. The amount of neurons in the hidden layer has been varied from two to ten and for every of these configurations ten different initializations have been used. This allows to well compare the performances of the models with different amounts of neurons in the hidden

layer. Additionally, for every initialization the input candidates were separated in ten data folds. Trainings have been performed for all combinations of 9 of these folds where the tenth fold has been used for testing. All the trainings were using Mean Square Error (MSE) loss functions and were stopped after 25 successive failures of the Sum Product (SP) index validation improvement. The SP index is calculated by

$$\mathrm{SP} = \sqrt{\sqrt{\mathrm{P_D}(1-\mathrm{F_R})} \cdot \frac{1}{2}(\mathrm{P_D}+1-\mathrm{F_R})}, \tag{1}$$

where $\mathrm{P_D}$ is the probability of detection and $\mathrm{F_R}$ is the fake rate . The $\mathrm{P_D}$ is the probability for a signal candidate to be properly classified by the neural network and the $\mathrm{F_R}$ is the probability for a background candidate to be classified as signal. Figure 3 (a) shows the average normalized energies in each ring for the signal and background candidates and (b) the NN output with the highest achieved SP index.



(a) (b)

Figure 3: A rings mean profile (a) and the NN output with maximum SP index (b) for signal and background forward electron candidates are shown. The profiles in (a) are normalized so that the sum of the 14 rings is equal to 1.0. The given errors depict the RMS width of the individual normalized ring sums. The markers of both samples have been shifted within the same bin for visibility reasons.
The NN output value in (b) ranges from -1 to 1. The given error bars depict the statistical error of each bin. The optimal threshold output value with regards to maximising the SP index is shown as a dashed black line at 0.662.

In Figure 3 (a) it can be seen that the average normalized energy of the first ring is much higher for the signal, whereas the ones for rings nine to twelve are clearly higher for the background. The information from these rings are very important for the NN to distinguish signal from background. Figure 3 (b) shows that the corresponding NN was able to very well distinguish signal from background. The different configurations of MLPs varying the amount of neurons in the hidden layer were generally performing very well. However, the ones with only two neurons in the hidden layer clearly showed a worse performance than the others. From three neurons on, only a very faint tendency of an improved performance with a higher amount of neurons could be observed. The neural network performances were also compared to the ones of the cut-based approach which is currently in use for forward electrons. There are three different cut-based working points in use for different probabilities of detection, called ‘loose’, ‘medium’

and 'tight'. For the comparison the NN models' fake rates were evaluated at a fixed probability of detection of $P_D = 94\%$ which corresponds to the one of the 'medium' cutbased working point for the used signal samples. Using the 'medium' cutbased conditions a fake rate of about 9.1% arises using the background samples. However, the selections provided by the MLPs with more than three neurons in the hidden layer only result in an average fake rate just below 0.6%. This would result in an overall reduction of the forward electron fake rate by more than a factor 15 which would be a major improvement. One has also to consider that these results were obtained by using an electron gun sample which provides a very clean signal signature. Therefore, the expected performance of using the NeuralRinger for triggering forward electrons during proton collisions within ATLAS is slightly worse.

5. Conclusion

Studies on extending the NeuralRinger algorithm to more forward regions of $2.5 < |\eta| < 3.2$ have been presented. For these, a structure of 14 rings fitting the granularity has been designed and implemented using the cells of the electromagnetic and hadronic calorimeters. Those 14 rings have been built for the electron candidates inside electron gun and dijet samples as signal and background. The energy information from the cells in the rings were then fed to a series of multi layer perceptrons with different configurations. They provided discriminants to distinguish signal from background. By comparing with the cut-based approach which is currently in use, a reduction of the amount of falsely identified background candidates by a factor 15 could be observed. This is a strong improvement on the forward electron trigger. However, this was only a preliminary study which has to be followed up by using less clean signal datasets like the ones of simulated collisions with Z bosons decaying into two electrons ($Z \to ee$) and generally using more statistics. Currently $Z \to ee$ and dijet samples are being obtained and reconstructed with the forward rings, which have been newly produced in 2021 and 2022. Additional studies on the usefulness of all the 14 rings and on different ring structures are being prepared. However, the improvements on the forward electron trigger are very much depending on the progress made in the forward L1 trigger where similar studies are performed at the moment. It will be very important to coordinate both efforts in the L1 trigger and in the HLT to successfully improve the forward trigger.

Acknowledgments

These studies were realized with the support of the Swiss National Science Foundation (SNSF Grant Number: 186265).

References

[1] The ATLAS Collaboration 2008 *Journal of Instrumentation* **3** S08003–S08003 URL https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003
[2] Aad G *et al.* (ATLAS) 2020 *JINST* **15** P10004 (*Preprint* 2007.12539)
[3] The ATLAS Collaboration 2020 *The European Physical Journal C* **80** URL https://doi.org/10.1140/epjc/s10052-019-7500-2
[4] Freund W S and on behalf of the ATLAS Collaboration 2020 *Journal of Physics: Conference Series* **1525** 012076 URL https://dx.doi.org/10.1088/1742-6596/1525/1/012076
[5] The ATLAS Collaboration 2017 Technical Design Report for the ATLAS Inner Tracker Pixel Detector Tech. rep. CERN Geneva URL https://cds.cern.ch/record/2285585