


# Filtragem online de elétrons com redes neurais convolucionais a partir de dados de calorimetria

Lucas Barra de Aguiar Nunes

LPS<sup>1</sup>

UFRJ<sup>2</sup> - COPPE/Poli


Rio de Janeiro, RJ, Brasil 

lucas.nunes@lps.ufrj.br

Natanael N. Moura Junior

LPS<sup>1</sup>

UFRJ<sup>2</sup> - COPPE/Poli


Rio de Janeiro, RJ, Brasil 

natmourajr@lps.ufrj.br

José Manoel de Seixas

LPS<sup>1</sup>

UFRJ<sup>2</sup> - COPPE/Poli

Rio de Janeiro, RJ, Brasil 

seixas@lps.ufrj.br

1 - Laboratório de Processamento de Sinais (LPS)

2 - Universidade Federal do Rio de Janeiro (UFRJ)

**Resumo**—Na etapa atual de aquisição de dados no LHC (Large Hadron Collider), no interior do Experimento ATLAS, o volume de dados atinge 120TB/s como consequência das colisões de prótons, que ocorrem a cada 25ns. Entretanto, os eventos de interesse são extremamente raros de tal modo que um sistema de filtragem online é implementado para armazenar apenas os eventos de física de interesse do experimento. Nesse contexto, é proposto o *Ringer Inception*, um comitê de redes neurais convolucionais para a seleção de elétrons no experimento baseada na análise de anéis concêntricos de energia de um chuveiro de partículas, a partir de dados de calorimetria de altas energias. Essa arquitetura se inspira na rede *Inception* aliada ao conhecimento especialista de física de altas energias para aumentar a eficiência na detecção de eventos *boosted* (elétrons de alta energia com pouca distância entre si) raros. Exibindo resiliência ao empilhamento, o modelo mostrou-se capaz de detectar elétrons vindos desses decaimentos apresentando um ganho de 14,13 pp na eficiência geral, atingindo 98,52%, sendo 3 vezes mais eficiente em alguns casos enquanto a probabilidade de falso alarme passa a ser menor que 0,85%.

**Index Terms**—Machine Learning, Redes Neurais Convolucionais, Física de Altas Energias, Trigger Online

## I. INTRODUÇÃO

O ATLAS [1] é um dos principais experimentos do Grande Colisor de Hádrons (LHC) [2], localizado no *Centro Europeu para a pesquisa Nuclear* (CERN) [3], na fronteira entre a Suíça e a França. Um feixe de prótons de cerca de 13,6 TeV no centro de massa colide no interior de sua estrutura subterrânea de mais de 20 metros de altura que combina um detector de traço, um calorímetro hadrônico, um calorímetro eletromagnético e um espectrômetro de múons para realizar a identificação dos decaimentos de partículas de interesse. O experimento teve como primeiro objetivo a detecção do Bóson de Higgs, identificado no ano de 2012 [4] com a teoria premiada com o Nobel do ano seguinte [5]. Após a descoberta, os esforços se direcionam a estudar as propriedades do Higgs e aprofundar a busca por candidatos a matéria escura a partir de teorias de supersimetria e ALPs (*Axiom-like particles*) [6] [7].

O experimento gera  $\approx 120TB/s$  com uma colisão de prótons, em média, a cada 25ns [1]. Os eventos, todavia, são extremamente raros, tornando necessária a implementação

de um sistema de *trigger* para identificação deles, o qual é implementado em dois níveis sequenciais de processamento. O calorímetro do ATLAS, foco desse estudo, é dividido em 3 partes (*Pre-Sampler*, Eletromagnético e Hadrônico), totalizando 7 camadas projetadas para captar distintos comportamentos físicos. Essas camadas geram uma região de interesse, no segundo nível de *trigger*, com cerca de 16951 canais (Tabela I). Com uma taxa de eventos elevada que permite um tempo de processamento de alguns microssegundos, o sistema precisa equilibrar velocidade de processamento e precisão para que não haja perda de dados durante a identificação de um chuveiro de partículas (nome dado ao decaimento em cascata de uma partícula resultante do decorrer dela pelo sensor).

A utilização do *Ringer* para sistemas de *trigger* é um trabalho que data mais de 20 anos da publicação desse texto. Em 1995, foi proposta uma rede neural para um sistema de dois níveis de *trigger* que utiliza anéis calculados sob dados de calorimetria com uma normalização sequencial [8]. O desenvolvimento desse trabalho nos anos seguintes resultou na adoção do *Ringer* baseado em MLP como o método padrão de *trigger* do ATLAS no seu segundo nível (nível *Fast Calo*) em 2017 na sua versão v8 chamado de *Neural Ringer* [9]. A proposta na época substituiu o modelo anterior baseado em cortes rígidos sobre variáveis de chuveiro por reduzir expressivamente a taxa de falso positivo. Na sequência dos desenvolvimentos, técnicas de aprendizado profundo foram propostas para obter uma melhora no desempenho do *trigger* [10]. Em especial, foi analisada uma queda de desempenho dos modelos em dados de eventos onde o momento da partícula é muito maior que sua massa (eventos *boosted*) e propõem-se estratégias de solução que recuperam a detecção porém aumentam falsos positivos.

Nesse contexto, o foco desse trabalho é a recuperação da detecção em dados *boosted* sem que haja aumento de falsos positivos. Com inspiração no conhecimento especialista, aborda-se o uso de anéis concêntricos de energia obtidos pela calorimetria do ATLAS como dados de entrada para um modelo baseado em CNN's [11]. Por meio de combinações dos tamanhos dos filtros de convolução, procurou-se construir na topologia do modelo características de análise empregadas em

Tabela I: Tabela do número de anéis por camada do calorímetro

Camada	Número de anéis	Total de células
PS	8	225
EM1	64	16129
EM2	8	225
EM3	8	225
HAD1	4	49
HAD2	4	49
HAD3	4	49
<b>Total</b>	<b>100</b>	<b>16951</b>

variáveis físicas discriminantes para identificação de partículas em física de altas energias, especialmente em eventos *boosted*. Os modelos foram desenvolvidos de forma a fornecer diferentes níveis de pureza na classificação dos eventos (*Tight*, *Medium*, *Loose* e *VeryLoose*, em ordem decrescente), permitindo uma escolha de compromisso entre número e qualidade de amostras. Além disso, a assimetria espacial do calorímetro e a variação do comportamento físico em função da energia do evento levaram ao uso de um comitê de redes no qual cada membro é treinado a classificar eventos com energia e posição específica [12]. O problema também exige uma restrição na qual o desempenho do sistema de *trigger* seja invariante à taxa de empilhamento das colisões, implementado por um limiar de decisão em função dessa variável. Nomeada de *Ringer Inception*, a técnica proposta foi capaz de manter uma taxa de detecção de elétrons similar a versão anterior v8, porém com a capacidade de detectar eventos mais raros com menos falsos positivos.

Esse artigo é organizado em 5 seções. A seção 2 descreve o experimento ATLAS, o atual sistema de *trigger* e dificuldades de detecção. A seção 3 descreve a solução proposta e o método aplicado. A seção 4 descreve os resultados encontrados. A seção 5, conclui esse trabalho.

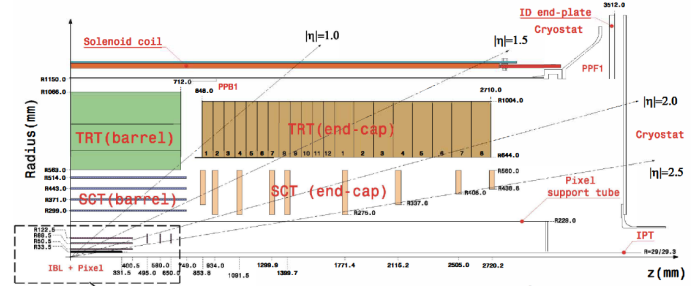
## II. O SISTEMA DE *trigger* DO EXPERIMENTO ATLAS

### A. Detector

O detector empregado no experimento ATLAS consiste numa associação de diversos detectores menores. O foco será no seu calorímetro, instrumento de medição base para o desenvolvimento desse trabalho.

O calorímetro do ATLAS possui um formato cilíndrico composto por um conjunto de células que captam a energia de parte do chuveiro que, quando combinadas, montam uma imagem da deposição espacial da energia do chuveiro. O sistema de coordenadas das células de cada camada do sensor é dada em função de  $\phi$ , ângulo azimutal ao redor do eixo do cilindro, e  $\eta$ , variável que quantifica o quão distante o evento está de um ângulo de 90 graus do feixe (Figura 1). O tamanho físico das células depende da posição das mesmas no sistema de coordenadas. Aquelas mais próximas de um ângulo de 90 graus do feixe (menor  $\eta$ ) tendem a ser menores que as demais, pois a probabilidade de ocorrência de eventos raros nessa região é maior. Esse fato cria regiões no calorímetro

de granularidade distintas em função de  $\eta$ . O calorímetro é simétrico em relação a  $\phi$ .


 Figura 1: Seção transversal do detector ATLAS com o valor de  $|\eta|$  em evidência [1]

Além dessas diferenças gerais, o calorímetro é dividido em 3 grandes camadas. O *Pre-Sampler* (PS) é a camada mais próxima do ponto de colisão e foi desenvolvida para captar chuveiros de partículas que iniciam perto do ponto de colisão e atingem o calorímetro em níveis mais avançados de decaimento. O calorímetro eletromagnético (EM) é constituído de 3 subcamadas (EM1, EM2, EM3) e foi projetado para captar chuveiros formados de partículas que interagem com o campo eletromagnético. Os elétrons, em especial, possuem chuveiros mais concentrados em relação a chuveiros hadrônicos (Figura 2), consequentemente, as camadas EM possuem granularidade maior para captar essas características. Na parte mais externa, o calorímetro hadrônico possui como característica absorver chuveiros vindos de decaimentos hadrônicos e também possui 3 subcamadas (HAD1, HAD2, HAD3). Como esses tipos de chuveiros possuem uma largura maior que os eletromagnéticos, as células possuem uma granularidade menor em relação as outras camadas.

Com uma composição de sensores que interagem com características físicas diferentes o método *Ringer Inception* se insere unindo a informação desses dados para realizar a tomada de decisão.

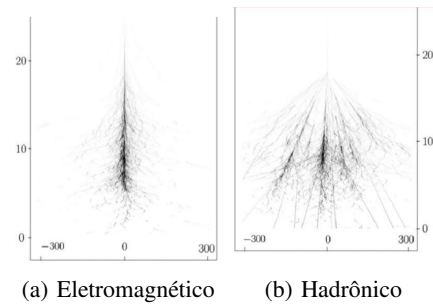


Figura 2: Imagem 2D do desenvolvimento de dois chuveiros de características distintas. O chuveiro hadrônico é mais esparsos em relação ao eletromagnético.

### B. Sistema de trigger

O sistema de *trigger* do ATLAS visa filtrar de forma *online* os eventos que ocorrem dentro do experimento, armazenando apenas aqueles que são de interesse para o estudo da física. Essa seletividade cria uma necessidade de precisão no sistema para que não se perca eventos raros de interesse.

O sistema é dividido em cadeias de processamento para cada tipo de evento a ser detectado (elétrons, fótons, etc). Cada uma possui uma série de etapas sequenciais, divididas em níveis de hardware e software, nas quais um evento precisa ser aprovado para ser membro da dada cadeia. No caso das cadeias de elétrons, foco desse trabalho, o processamento se dá, em ordem, pelas seguintes etapas: L1 (Implementada em hardware), *Fast Calo*, *Fast Electron*, *Precision Calo* e *Precision Electron* (Implementadas em software). Elas são projetadas de forma a reduzir gradualmente o número de eventos pela rejeição de ruído para que a camada seguinte tenha mais tempo computacional para utilização de algoritmos mais complexos que a anterior a fim de melhorar a identificação.

Após o processamento online, os dados são salvos em disco e recebem uma última classificação, considerada como verdade final, por um algoritmo offline de máxima verossimilhança que possui acesso a todos os milhares de canais e variáveis calculadas do detector com a maior precisão disponível.

O Ringer, portanto, se insere na etapa *Fast Calo* e age sobre dados rapidamente construídos de calorimetria com foco em rejeição de ruído.

### C. Eventos boosted

Em 2017, com amostras de simulação de Monte Carlo, foi detectado que o *Neural Ringer* em operação na época apresentava ineficiências em um grupo específico de evento físico raro no qual o momento da partícula é muito maior que a massa (eventos *boosted*). Os dados da amostra são do decaimento de *boosted* dibósons vindos de eventos de fusão de glúons. [13]. Nesse tipo de evento, devido ao momento elevado, dois elétrons vindo do decaimento de um bóson são emitidos com um ângulo pequeno, consequentemente, ambos atingem a mesma região de interesse do algoritmo, resultando em erros de identificação. Esse efeito é evidenciado na figura 3. A variável  $\Delta R$  ( $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ ) é uma medida de quão afastadas estão duas partículas de um decaimento. A estratégia Ringer, apresentou uma eficiência menor que o algoritmo de comparação para baixo  $\Delta R$ . Como esses eventos ocorrem em altas energias, a eficiência em  $E_T$  elevado decai até o ponto no qual a energia é tão alta que os dois decaimentos se alinham e o modelo recupera eficiência (Figura 4).

Em [10], o autor propõe uma nova versão (v12) que reduz a região de interesse de análise como método de recuperação de eficiência para esse tipo de evento. Como a característica do evento é a proximidade do desenvolvimento dos dois elétrons, o modelo utiliza a parte mais central dos anéis de cada camada para retirar a interferência do elétron mais externo. A estratégia conseguiu recuperar eficiência, porém houve o aumento de falsos positivos. O *Ringer Inception* é proposto como forma

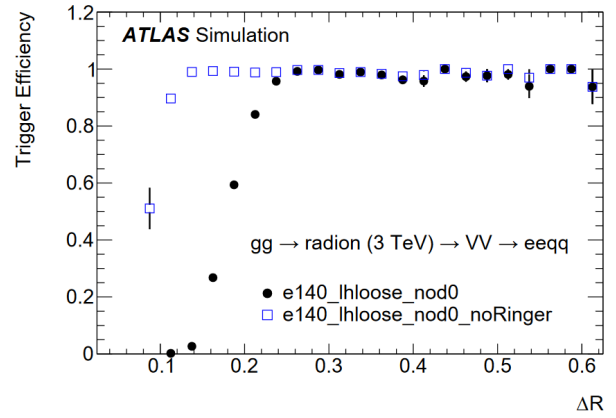


Figura 3: Taxa de aprovação do *trigger* em função de  $\Delta R$  para elétrons simulados de fusão de glúons com critério *loose* e energia  $\geq 140$  GeV.

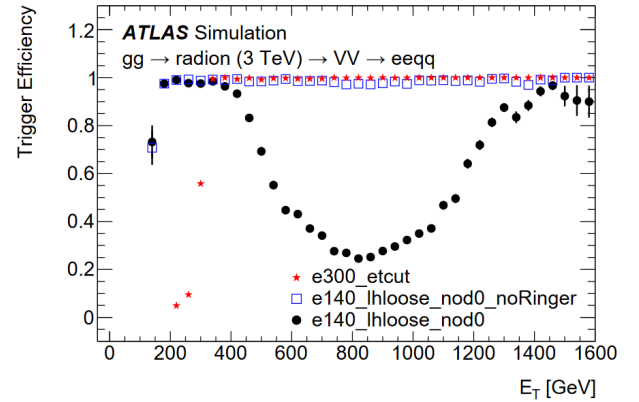


Figura 4: Taxa de aprovação do *trigger* em função de  $E_T R$  para elétrons simulados de fusão de glúons com critério *loose* e energia  $\geq 140$  GeV.

de recuperar a eficiência de detecção sem um aumento na taxa de falsos positivos.

## III. DESENVOLVIMENTO DO MODELO

### A. Anéis Concêntricos

A ação no nível *fast* torna necessária a construção de variáveis que sejam de cálculo rápido e de dimensão baixa pra evitar o aumento da complexidade dos modelos [14] e atingir as restrições de tempo de processamento. Esse fato dificulta a utilização de todas as células do calorímetro como uma imagem tridimensional do chuveiro na entrada do modelo. Dessa forma, é desejável obter uma representação do evento que mantenha características da imagem original sem excesso de pré-processamento para que se possa aproveitar a capacidade das redes neurais de encontrar uma representação latente eficiente para o problema. Como resultado, foi empregado o uso de anéis concêntricos de energia [8] para a caracterização do chuveiro a partir dos dados de calorimetria disponíveis nessa etapa.

No nível L1, são estimados e fornecidos para o *fast* os pontos centrais do evento em cada camada do calorímetro e a região de interesse (RoI) ao redor desse ponto. A partir dessas informações, na etapa seguinte, são calculados os anéis de forma que cada camada do detector seja uma seção transversal do cilindro que compreende o chuveiro. Esse método permite resumir em um vetor de 100 variáveis (Tabela I) cerca de 99% da energia do decaimento [8] de um evento que possuía, anteriormente, 16951 variáveis (1 variável por célula do calorímetro).

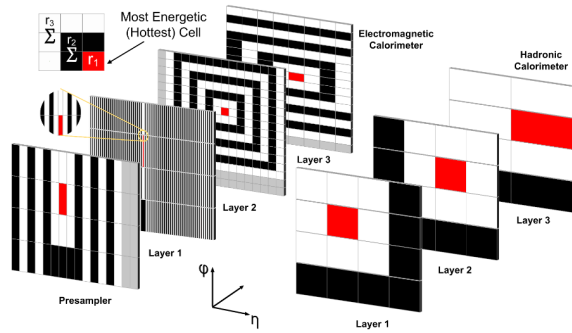


Figura 5: Imagem representando os anéis montados em cada camada, a célula vermelha simboliza a célula central da RoI de cada camada do calorímetro. [13]

### B. Variáveis de chuveiro

Variáveis de chuveiro (*shower shapes*) foram desenvolvidas por especialistas em calorimetria do experimento a partir do conhecimento físico para oferecer capacidade discriminante na identificação de eventos em poucas variáveis.

Na tabela II estão descritas as principais variáveis de chuveiro para identificação de elétrons. Esses valores concentram seus esforços em entender como a energia de um determinado evento é depositada ao longo do calorímetro. Algumas variáveis como  $E_{ratio}$ ,  $R_\eta$  e  $R_\phi$  quantificam a dispersão de energia do chuveiro dentro de uma mesma camada. Se o chuveiro é estreito seus valores tendem a 1, se o chuveiro é largo, tendem a zero. Variáveis como  $f_1$ ,  $f_3$  e  $R_{had}$  calculam a proporção de energia do evento depositada ao longo da profundidade do calorímetro de forma que se quantifique o nível de interação do chuveiro com diferentes camadas. No geral, todas as variáveis descritas trabalham com uma combinação de sub-regiões retangulares da região de interesse do evento. Essas sub-regiões podem ser quadrados de 3x3 ou 7x7 células ou retângulos de 3x7 ou 7x3 ( $\Delta\eta \times \Delta\phi$ ).

Essas variáveis se destacam, principalmente, na identificação de eventos *boosted*, pois são os dados de entrada do algoritmo de cortes rígidos que apresentou bom desempenho em relação ao *Ringer* na figura 4.

### C. Ringer Inception

#### 1) Arquitetura da rede

A arquitetura *Inception* [16] [17] foi proposta em 2014 e uma das primeiras a introduzir o conceito de subcamadas

Tabela II: Variáveis de chuveiro para discriminação de elétrons (Coordenadas em números de células em  $\eta \times \phi$ ) [15]

Variável	Fórmula de cálculo
$R_\eta$	$\frac{E_i^{3 \times 7}}{E_i^{7 \times 7}} \frac{EM2}{EM2}$
$R_\phi$	$\frac{E_i^{3 \times 3}}{E_i^{7 \times 7}} \frac{EM2}{EM2}$
$E_{ratio}$	$\frac{E_{EM1}^{max} - E_{EM1}^{2^{nd} max}}{E_{EM1}^{max} + E_{EM1}^{2^{nd} max}}$
$f_1$	$\frac{E_{EM1}^{total}}{E_{EM}^{total}}$
$f_3$	$\frac{E_{EM3}^{total}}{E_{EM}^{total}}$
$\omega_{\eta^2}$	$\sqrt{\frac{\sum E_i \times \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \times \eta_i}{\sum E_i}\right)^2}$
$\omega_{stot}$	$\sqrt{\frac{\sum (E_i - E_i^{max})^2}{\sum E_i}}$
$R_{had}$	$\frac{E_T^{HAD1}}{E_T^{EM}}  \eta  < 0.8 \cup  \eta  > 1.37$
$R_{had1}$	$\frac{E_T^{HAD1}}{E_T^{EM}} 0.8 <  \eta  < 1.37$

[18] onde um bloco da topologia consiste na combinação de múltiplas camadas com filtros convolucionais de tamanhos distintos. A combinação de sub-regiões também é uma propriedade incrustada na formulação das variáveis de chuveiro. Por meio da junção de áreas de tamanhos distintos do calorímetro, essas variáveis captam as principais características físicas do evento. Como o tamanho dessas regiões já foi mapeado nas formulações, é possível construir um filtro de convolução que aproxime essas regiões no processamento da rede.

O *Ringer Inception*, portanto, surge como um modelo que emprega uma arquitetura de combinação de filtros de tamanho distintos como a *Inception* original, porém, com o tamanho dessas regiões definido pelo conhecimento especialista da física. Essa construção de modelo é feita na intenção de manter a elevada taxa de rejeição de ruído que o *Ringer* oferece em comparação ao antigo modelo de cortes rígidos [12] e extrair propriedades das variáveis físicas que se mostraram eficientes em eventos *boosted* (Figuras 3 e 4).

Na figura 6 é mostrado um exemplo de como a convolução vetorial sobre o vetor de anéis ocorre no espaço do calorímetro. Ao tomar por referência a região central do anel, é possível notar que a área quadrada de 3x3 pode ser obtida por uma convolução vetorial de tamanho 2. Dessa forma, para obter as regiões de 3x3 e 7x7 citadas anteriormente, basta combinar filtros de tamanho 2 e 4.

Além do desenvolvimento lateral em uma camada, um chuveiro também se desenvolve na profundidade do calorímetro, atravessando as diversas camadas do detector. Se todos os anéis fossem mantidos como um vetor único, eventualmente o filtro de convolução se encontraria no espaço entre duas camadas (Figura 7). Esse comportamento inclui uma quebra do padrão de processamento do modelo pois, enquanto um filtro está dentro de uma mesma camada, ao deslizar, ele se direciona para anéis mais externos. Na região de transição das camadas, entretanto, esse comportamento é distorcido pois os anéis mais a direita serão os anéis centrais da camada seguinte

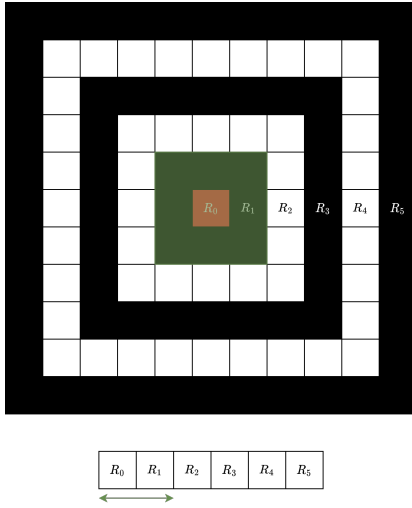


Figura 6: Imagem ilustrativa do comportamento no espaço do calorímetro de uma convolução vetorial em cima dos anéis para uma área quadrada de 3 células e um filtro convolucional de 2 anéis.

e não mais os anéis mais externos da camada anterior. Dessa forma, o modelo foi construído para analisar cada camada do calorímetro como uma entrada distinta e combiná-las nas camadas de Perceptrons após as convoluções (Figura 8b). Tal comportamento, também ocorre nas variáveis físicas que primeiro realizam agregações no espaço lateral dos chuveiros com as sub-regiões e depois combinam esses dados com informações de outras camadas.

Assim como o modelo Inception original, o *Ringer Inception* também implementa  $\text{relu}$  como a função de ativação em todas as camadas exceto a última que utiliza a função sigmoideal.

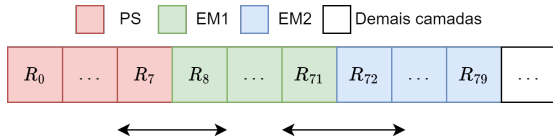
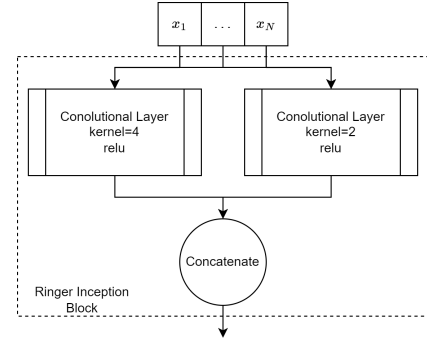


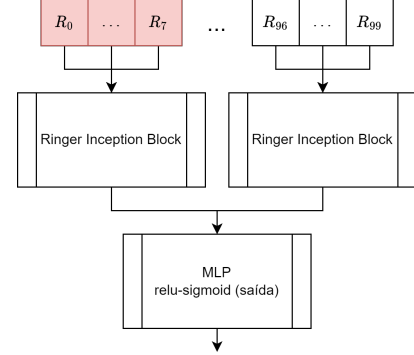
Figura 7: Imagem representando filtros (setas) combinando camadas distintas com anéis externos a esquerda e anéis centrais a direita na região de transição entre camadas, uma distorção do padrão de análise dentro de uma camada.

## 2) Construção do ensemble

O *Ringer Inception* é construído como um *ensemble* de redes da arquitetura descrita anteriormente de forma igual para todos os membros, seguindo a estratégia de *NeuralRinger's* anteriores [10] [12]. Devido a diferenças na granularidade do calorímetro em função de  $|\eta|$  e a variação da interação do chuveiro com o calorímetro em função da energia transversa ( $E_T$ ), o comitê é formado por diversos membros especializados em uma região do calorímetro e de energia conforme a tabela III, totalizando um *ensemble* de 25 membros. Essa



(a) Bloco *Inception* desenvolvido para o *Ringer Inception*



(b) Arquitetura do modelo

Figura 8: Arquitetura do *Ringer Inception*

Tabela III: Regiões do ensemble em intervalos de  $|\eta|$  e  $E_T$  para os membros do comitê de redes neurais. Para  $E_T$  os intervalos são fechados a esquerda e abertos a direita. O mesmo ocorre com  $|\eta|$  com exceção do último intervalo que é fechado dos dois lados.

Variável	Bins	Total de Bins
$ \eta $	(0 0,8 1,37 1,54 2,37 2,5)	5
$E_T$ GeV	(15 20 30 40 50 $\infty$ )	5

divisão visa um ganho de performance ao treinar uma rede para cada região do equipamento com propriedades distintas.

## D. Método de treinamento

### 1) Treinamento da rede

Os dados de treinamento consistem na agregação de dois conjuntos de dados de colisão identificados pelo algoritmo offline do ATLAS: elétrons vindos de decaimentos  $Z \rightarrow ee$  para sinal e jatos hadrônicos com energia  $\geq 17$  GeV para ruído. O offline identifica com diferentes graus de pureza (*Tight*, *Medium*, *Loose* e *VeryLoose*) se um evento é pertencente a uma determinada classe. Para o sinal, escolheu-se eventos de elétrons aprovados no nível *medium* como uma forma de compromisso entre qualidade das amostras e número de eventos. Para o ruído, foram considerados os jatos que não conseguiram aprovação como elétrons no critério mais brando (*VeryLoose*). Como o decaimento descrito e outros processos

físicos importantes geram pares de elétrons, os elétrons (jatos) do treinamento foram aprovados (reprovados) no método de *Tag and Probe* [19]. Esse método consiste na identificação de uma partícula *Tag*, com características claras da partícula desejada, e a seleção de outra (*Probe*), com características físicas possíveis para um par da *Tag*. Dessa forma, alia-se qualidade dos dados para treinamento e se prioriza a identificação de pares de elétrons para que os eventos sejam de maior usabilidade para a física.

O treinamento se deu por um método de KFold [14] com 10 divisões e 5 inicializações por *fold* a fim de evitar resultados devido a má inicialização dos pesos da rede [11]. Cada modelo foi treinado com um número indefinido de iterações com a parada determinada pelo fim da otimização do índice SP [10] sob o conjunto de validação ao final de cada iteração. A melhor inicialização é selecionada como o modelo com maior índice SP entre os dados de validação. Dentre as melhores inicializações de cada *fold* o melhor modelo é selecionado como o de maior SP sob todo o conjunto de dados.

## 2) Ajuste do limiar de decisão

Uma das requisições do experimento é a baixa variância do sistema à mudança nas taxas de empilhamento. Dessa forma, a partir do estimador de valor esperado de empilhamento ( $\langle \mu \rangle$ ) do *trigger*, é feito um ajuste do limiar de decisão que age sobre o discriminante da rede (combinação da saída da penúltima camada com os pesos da última) em função de  $\langle \mu \rangle$ . Como o sistema de *trigger* possui 4 graus de pureza na classificação, cada membro do comitê deve possuir um limiar para cada grau. Uma reta linear em função de  $\langle \mu \rangle$ , portanto, é ajustada sobre o discriminante do conjunto de elétrons de cada grau do conjunto de dados. Esse ajuste é feito para manter a mesma taxa de aprovação de eventos que o algoritmo anterior da etapa *fast* para uma dada pureza, a fim de evitar problemas na alteração de taxa de eventos para as camadas seguintes.

## IV. RESULTADOS

### A. Resultados iniciais

Para avaliar o desempenho do modelo, o conjunto de dados de teste foi construído a partir dos dados de colisão de 2017 obtidos no ATLAS. O sinal compõe todos os elétrons que foram aprovados em uma cadeia de elétron com energia  $E_T \geq 15 GeV$  e o ruído como todos os dados de jatos hadrônicos com energia  $E_T \geq 17 GeV$ . Nas figuras a seguir é avaliada a eficiência do *trigger* no nível *FastCalo* com diferentes versões do *ringer* desenvolvidas para dados de colisão gerais (Tabela V). O eixo y representa a taxa de aprovação de dados válidos para uma determinada cadeia do sistema de *trigger*. Em dados de elétrons é desejável que o valor seja próximo a 1 e em ruído que seja próximo de zero. Nos resultados a seguir, foi avaliada apenas uma topologia do *Ringer Inception* que consiste em um bloco *Inception* com todas camadas convolucionais com 2 neurônios e uma MLP de duas camadas: uma intermediária com 5 neurônios e uma de saída. Para os dados de treinamento as relações sinal e ruído podem ser vistas na tabela IV.

Tabela IV: Contagem de amostras de sinal e ruído para cada região do ensemble de treinamento

$E_T [GeV],  \eta $	Amostras (Sinal, Ruído)
[15, 20), [0, 0.8)	232819, 187639
[15, 20), [0.8, 1.37)	140971, 143657
[15, 20), [1.37, 1.54)	50948, 30037
[15, 20), [1.54, 2.37)	227315, 205792
[15, 20), [2.37, 2.5]	19765, 15512
[20, 30), [0, 0.8)	1499273, 316581
[20, 30), [0.8, 1.37)	896586, 227709
[20, 30), [1.37, 1.54)	283367, 47550
[20, 30), [1.54, 2.37)	1131989, 307253
[20, 30), [2.37, 2.5]	83992, 36522
[30, 40), [0, 0.8)	4188536, 223869
[30, 40), [0.8, 1.37)	2539126, 156992
[30, 40), [1.37, 1.54)	644952, 35206
[30, 40), [1.54, 2.37)	2376984, 196711
[30, 40), [2.37, 2.5]	164364, 24980
[40, 50), [0, 0.8)	4627778, 162421
[40, 50), [0.8, 1.37)	2837728, 112229
[40, 50), [1.37, 1.54)	470828, 25449
[40, 50), [1.54, 2.37)	2429064, 131668
[40, 50), [2.37, 2.5]	187727, 16100
[50, $\infty$ ), [0, 0.8)	1699541, 3769656
[50, $\infty$ ), [0.8, 1.37)	1059911, 2560472
[50, $\infty$ ), [1.37, 1.54)	170675, 560953
[50, $\infty$ ), [1.54, 2.37)	901054, 1318418
[50, $\infty$ ), [2.37, 2.5]	89782, 93785
Total	28955075, 10907161

Tabela V: Versões do Ringer

Versão	Descrição
v8	Modelo MLP utilizado em 2017 [12]
v10	Modelo CNN anterior para referência de comparação [10]
vInception	Modelo CNN proposto nesse trabalho

Em energias próximas ao limite da cadeia, os modelos apresentam certa ineficiência, entretanto, esse comportamento é esperado. Como ocorrem incertezas na estimação da energia limite da cadeia na etapa anterior há corte prematuro dos candidatos por consequência. Em altas energias ( $E_T \geq 100 GeV$ ), todavia, a eficiência das estratégias convolucionais decai, principalmente para cadeias com critérios justos (Figura 9) enquanto o modelo MLP mantém uma detecção próxima de 1. Esse comportamento é indesejado pois regiões de alta energia são as de maior interesse para o estudo de processos físicos raros. O modelo apresenta, no entanto, resposta aproximadamente constante de aprovação em função do valor esperado de empilhamento (Figura 10), um comportamento desejado.

### B. Bin extra de energia

Como forma de mitigar esse comportamento foi implementado o mesmo método utilizado em [10]. Adiciona-se um novo *bin* de energia a partir de 100 GeV na divisão dos membros do *ensemble* (Tabela VI). Os pesos das redes dos intervalos de  $E_T \in [50, 100) GeV$  e  $E_T \in [100, \infty) GeV$ , porém, são os mesmos que os do bin anterior ( $E_T \in [50, \infty) GeV$ ) e apenas o limiar de decisão é reajustado para atender os dados dentro

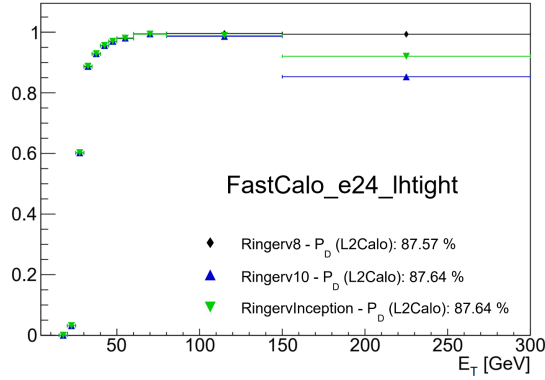


Figura 9: Taxa de aprovação da cadeia de elétrons de energia  $\geq 24$  GeV com critério *tight* em função de  $E_T$  para sinal (Quanto maior melhor). A porcentagem na legenda representa a aprovação geral em todo o conjunto.

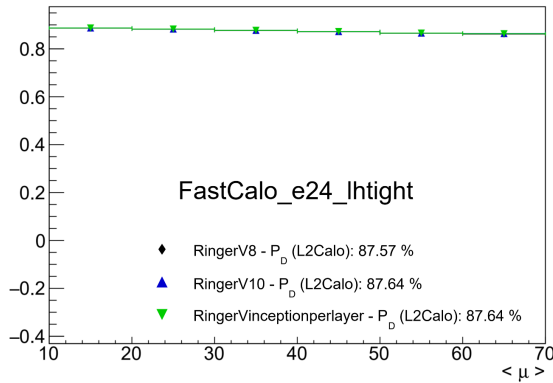


Figura 10: Taxa de aprovação da cadeia de elétrons de energia  $\geq 24$  GeV com critério *tight* em função de  $\langle \mu \rangle$  para sinal (Quanto maior melhor). A porcentagem na legenda representa a aprovação geral em todo o conjunto.

desses novos intervalos. Com esse ajuste, é possível notar a recuperação da eficiência em altas energias e os critérios passam a atingir uma aprovação máxima (Figura 11). Além disso, foi possível reduzir o nível geral de falso positivo em relação a versão v8 embora em algumas regiões de energia, ela ainda apresente um valor menor de falso positivos (Figura 12).

### C. Resultados para eventos boosted

Para avaliar os resultados nos eventos raros descritos na seção II-C, a resposta do modelo foi comparada com a

Tabela VI: Novas regiões do ensemble

Variável	Bins	Total de Bins
$ \eta $	(0 0,8 1,37 1,54 2,37 2,5)	5
$E_T$ GeV	(15 20 30 40 50 100 $\infty$ )	5

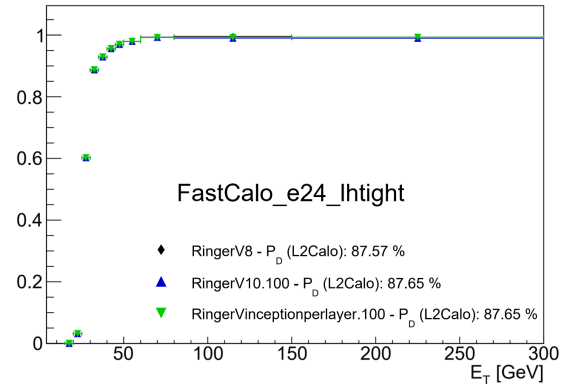


Figura 11: Taxa de aprovação da cadeia de elétrons de energia  $\geq 24$  GeV com critério *tight* para sinal (Quanto maior melhor). Os modelos terminados em ".100" possuem o bin extra de energia. A porcentagem na legenda representa a aprovação geral em todo o conjunto.

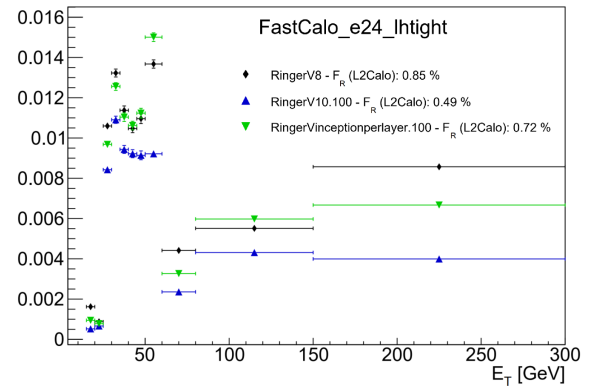


Figura 12: Taxa de aprovação da cadeia de elétrons de energia  $\geq 24$  GeV com critério *tight* para ruído (Quanto menor melhor). Os modelos terminados em ".100" possuem o bin extra de energia. A porcentagem na legenda representa a aprovação geral em todo o conjunto.

referência da época, Ringer v8. Os resultados na figura 13 mostram que a estratégia *Ringer Inception* apresentou ganhos gerais de detecção de cerca de 14,13 pp e, no caso dos menores valores de  $\Delta R$ , a região de maior perda de eficiência, um ganho até 3 vezes maior. O modelo sem o *bin* extra de energia já apresentava ganhos que se mostraram mais expressivos com um novo limiar para altas energias. Essa recuperação ocorreu aliada a uma diminuição de falsos positivos como mostrado pela figura 12. Além disso, devido a perda de performance, a versão v8 apresentou uma variância de desempenho com o valor esperado de empilhamento  $\langle \mu \rangle$  que foi estabilizada com a nova abordagem (Figura 14).

## V. CONCLUSÃO

Um método de identificação online de elétrons baseado em redes neurais convolucionais foi desenvolvido nesse trabalho

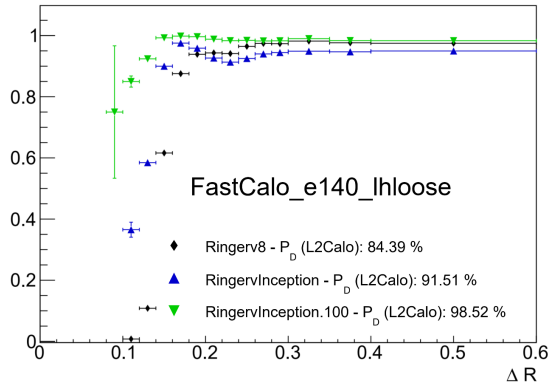


Figura 13: Taxa de aprovação do *trigger* em função de  $\Delta R$  para elétrons simulados de fusão de glúons com critério *loose* e energia  $\geq 140$  GeV. A porcentagem na legenda representa a aprovação geral em todo o conjunto.

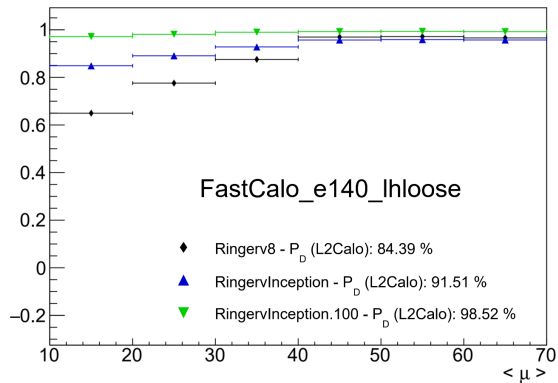


Figura 14: Taxa de aprovação do *trigger* em função de  $\langle \mu \rangle$  para elétrons simulados de fusão de glúons com critério *loose* e energia  $\geq 140$  GeV. A porcentagem na legenda representa a aprovação geral em todo o conjunto.

com foco na melhoria da detecção de eventos raros sem que se prejudicasse o funcionamento base do sistema. O desenvolvimento se deu com foco na construção de uma arquitetura que, por meio da organização das camadas e hiperparâmetros do modelo, incorpora características conhecidamente discriminantes na física. Essa proposta trouxe como resultado uma melhora geral de 16% na detecção de eventos *boosted* aliada a uma diminuição geral de falsos positivos. O modelo ainda se mostrou mais invariante à taxa de empilhamento nesse conjunto de dados. Os resultados expostos, foram avaliados com uma topologia de rede foi avaliada o que permite possíveis avanços futuros.

## VI. AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Os autores desse

trabalho também gostariam de agradecer às instituições de fomento nacionais CNPq, FAPERJ, RENAE e também ao CERN e à colaboração ATLAS pelo suporte e incentivo nesse desenvolvimento e aos futuros.

## REFERÊNCIAS

- [1] A. Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3,” May 2023, arXiv:2305.16623 [hep-ex, physics:physics]. [Online]. Available: <http://arxiv.org/abs/2305.16623>
- [2] E. Keil, “The CERN Large Hadron Collider LHC,” in *INFN Eloisatron Project: 34th Workshop: Hadron Colliders at the Highest Energy and Luminosity*, Dec. 1996, pp. 69–83.
- [3] “About | CERN.” [Online]. Available: <https://home.cern/about>
- [4] T. A. Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Physics Letters B*, vol. 716, no. 1, pp. 1–29, Sep. 2012, arXiv:1207.7214 [hep-ex]. [Online]. Available: <http://arxiv.org/abs/1207.7214>
- [5] “The Nobel Prize in Physics 2013.” [Online]. Available: <https://www.nobelprize.org/prizes/physics/2013/summary/>
- [6] M. Saimpert, “Searches for dark matter with the ATLAS detector,” 2023, place: Geneva. [Online]. Available: <https://cds.cern.ch/record/2839915>
- [7] R. Costa Batalha Pedro, “Searches for dark matter with the ATLAS detector,” 2022, place: Geneva. [Online]. Available: <https://cds.cern.ch/record/2826042>
- [8] J. Seixas, L. Caloba, M. Souza, A. Braga, and A. Rodrigues, “Neural second-level trigger system based on calorimetry,” *Computer Physics Communications*, vol. 95, no. 2-3, pp. 143–157, Jun. 1996. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/0010465596000124>
- [9] J. V. d. F. Pinto and o. b. o. A. Colaboration, “An Ensemble of Neural Networks for Online Filtering Implemented in the ATLAS Trigger System,” *Journal of Physics: Conference Series*, vol. 1162, no. 1, p. 012039, Jan. 2019, publisher: IOP Publishing. [Online]. Available: <https://dx.doi.org/10.1088/1742-6596/1162/1/012039>
- [10] J. V. D. F. Pinto, “Filtragem Online Segmentada Baseada em Redes Neurais operando na informação de um calorímetro de altas energias de fina granularidade.” Tese de Doutorado, Universidade Federal do Rio de Janeiro - UFRJ, Rio de Janeiro, Apr. 2022. [Online]. Available: <http://www.pee.ufrj.br/index.php/pt/producao-academica/teses-de-doutorado/2022/2016033453--221/file>
- [11] S. S. Haykin and S. S. Haykin, *Neural networks and learning machines*, 3rd ed. New York: Prentice Hall, 2009, oCLC: ocn237325326.
- [12] W. S. Freund, “Identificação de elétrons baseada em um calorímetro de altas energias finamente segmentado,” Tese de Doutorado, Universidade Federal do Rio de Janeiro - UFRJ, Rio de Janeiro, Mar. 2018. [Online]. Available: <http://www.pee.ufrj.br/index.php/pt/producao-academica/teses-de-doutorado/tese-1/2016033266--108/file>
- [13] A. Collaboration, “Performance of electron and photon triggers in ATLAS during LHC Run 2,” *The European Physical Journal C*, vol. 80, no. 1, p. 47, Jan. 2020, arXiv:1909.00761 [hep-ex]. [Online]. Available: <http://arxiv.org/abs/1909.00761>
- [14] C. M. Bishop, *Pattern recognition and machine learning*, ser. Information science and statistics. New York: Springer, 2006.
- [15] A. Collaboration, “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data,” *Journal of Instrumentation*, vol. 14, no. 12, p. P12006, Dec. 2019. [Online]. Available: <https://dx.doi.org/10.1088/1748-0221/14/12/P12006>
- [16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going Deeper with Convolutions,” Sep. 2014, arXiv:1409.4842 [cs]. [Online]. Available: <http://arxiv.org/abs/1409.4842>
- [17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” Dec. 2015, arXiv:1512.00567 [cs]. [Online]. Available: <http://arxiv.org/abs/1512.00567>
- [18] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020. [Online]. Available: <https://link.springer.com/10.1007/s10462-020-09825-6>
- [19] “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data,” Jun. 2016.