

# Sistema *Online* Inteligente para Classificação de Fótons Baseado em Sinais de Calorimetria em um Ambiente com Alta Taxa de Eventos

Luiz E. B. Filho, José M. de Seixas, Juan L. Marin, Edmar E. P. de Souza, Eduardo F. Simas Filho

**Resumo**—Dentre o vasto programa de física de partículas do experimento ATLAS do LHC no CERN, fótons são partículas de interesse em distintos modelos físicos, que podem estar relacionados com fenômenos ainda não verificados experimentalmente. Para identificar fótons um sofisticado sistema sequencial de seleção *online* de eventos (*trigger*) torna-se necessário, uma vez que nas colisões de prótons, até 52 TB/s de informação pode ser produzida, no qual a maior parte da informação é caracterizada como ruído de fundo. Neste trabalho, uma estratégia inteligente de classificação *online* de fótons é proposta. A energia medida nos calorímetros do ATLAS é então formatada em anéis concêntricos em torno da célula com maior energia depositada. Redes Neurais Artificiais são projetadas com a informação dos anéis como entradas, para a discriminação entre fótons e ruído de fundo. Os resultados do método proposto indicaram melhorias com uma redução em até 31 pontos percentuais na aceitação de ruído de fundo, considerando uma eficiência de 83,7% na identificação *online* de fótons.

**Palavras-Chave**—Redes Neurais Convolucionais, Extração de Características, Fótons, Reconhecimento de Padrões.

**Abstract**—Among the vast particle physics program of the ATLAS experiment at LHC/CERN, photons are particles of interest in different physics models, which may be related to phenomena that have not yet been verified experimentally. To identify photons, a sophisticated sequential online event selection trigger system becomes necessary, since in proton collisions, up to 52 TB/s of information can be produced, in which the most of the information is characterized as background noise. In this work, an intelligent photon classification strategy is proposed. The energy measured from ATLAS calorimeters is then shaped into concentric rings around the cell with the highest deposited energy. Convolutional Neural Networks are designed with information from the rings as inputs, to discriminate between photons and background noise. The results of the proposed method indicated improvements with a reduction of up to 31 percentage points in the acceptance of background noise, considering an efficiency of 83.7% in online photon identification.

**Keywords**—Convolutional Neural Networks, Feature Extraction, Photons, Pattern Recognition.

## I. INTRODUÇÃO

Complexos sistemas tecnológicos tem sido construídos para experimentos de física de altas energias, com o objetivo de medir com precisão propriedades fundamentais subatômicas,

Luiz E. B. Filho, LPS-COPPE/Poli, Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, e-mail: balabram@lps.ufrj.br; José M. de Seixas, LPS-COPPE/Poli, Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, e-mail: seixas@lps.ufrj.br; Juan L. Marin, CELET, Instituto Federal da Bahia, Vitória da Conquista-BA, e-mail: juan.marin@ifba.edu.br; Edmar E. P. de Souza, PPGEE, Universidade Federal da Bahia, Salvador-BA, e-mail: edmar.egidio@ufba.br; Eduardo F. Simas Filho, PPGEE, Universidade Federal da Bahia, Salvador-BA, e-mail: eduardo.simas@ufba.br.

e investigar processos físicos ainda não observados experimentalmente. Neste contexto, o LHC (*Large Hadron Collider*) localizado na fronteira entre a Suíça e a França, é um dos maiores aceleradores de partículas da atualidade, no qual acelera feixes de prótons em sentidos opostos ao longo de um túnel circular com cerca de 27 km de circunferência, a uma profundidade média de 100 m. No LHC, as colisões entre os feixes de prótons ocorrem em diferentes pontos de cruzamento de feixes em até 14 TeV<sup>1</sup>, a uma taxa de 40 MHz, o que corresponde a uma colisão a cada 25 ns.

O experimento ATLAS (*A Toroidal LHC Apparatus*) [1], é um dos maiores experimentos de física de partículas já construído, instalado em um dos pontos de colisão de prótons do LHC. Sendo composto por sub-detecores e gera aproximadamente 1,3 MB de informação a cada colisão pelo seu sistema de reconstrução de eventos. Considerando a taxa de colisão nominal do LHC, o ATLAS produz cerca de 52 TB/s, requerendo um sofisticado sistema de seleção *online* (*trigger*) [2], para identificar os eventos relevantes que são então armazenados para análises *offline*. Dentre o vasto programa de física do ATLAS, fótons são partículas que possuem destacado interesse em diversos modelos de física envolvendo sinais raros, assim como processos físicos ainda não verificados experimentalmente. Todavia, o intenso ruído de fundo produzido durante as colisões, torna a identificação de fótons uma tarefa complexa, no qual processos raros como o bóson de Higgs [3], que podem decair em fótons em estados finais, podem ser mascarados pela alta produção de informação ordinária que falseiam os fótons de interesse. Nos últimos anos, técnicas de aprendizado de máquina têm sido usadas em várias aplicações, como detecção de partículas, reconstrução de eventos e filtragem *online* [4], [5]. Isso possibilita aos sistemas de classificação identificar padrões e características nos sinais do sistema de instrumentação, melhorando a distinção entre eventos de interesse e ruído de fundo.

O atual sistema de *trigger* de fótons do ATLAS utiliza variáveis unidimensionais baseadas na deposição de energia, extraídas do sistema de calorimetria, aplicando cortes rígidos sequenciais nestas grandezas, para classificação sinal/ruído. Entretanto, com o aumento progressivo da energia das colisões, bem como da luminosidade, sistemas inteligentes apresentam-se como uma alternativa para melhoria na rejeição de falsos fótons. Nos trabalhos [6], [7] o algoritmo NeuralRin-

<sup>1</sup>O (eV) é unidade definida como a quantidade de energia cinética para ganha por um elétron quando submetido a uma diferença de potencial de 1V no vácuo.

ger foi proposto para identificação *online* de fótons usando dados do segundo período de coleta de dados, consistindo em um método de extração de características que descreve o espalhamento lateral e longitudinal da energia depositada pela partícula que incide no sistema de calorimetria. Estes sinais extraídos alimentam redes neurais artificiais, que mostraram melhores resultados se comparados a estratégia de seleção por cortes rígidos em variáveis unidimensionais. Neste trabalho, extensões do NeuralRinger no que tange diferentes arquiteturas e configurações de Redes Neurais Convolucionais são exploradas para o terceiro período de coleta de dados, acreditando de um aumento na luminosidade, e comparadas com redes neurais rasas de uma única camada.

Este artigo está dividido do seguinte modo: na Seção II o experimento ATLAS, seus subsistemas e o sistema *online* de identificação de fótons é apresentado. Na Seção III o método proposto é descrito. A Seção IV apresenta os resultados e por fim as conclusões são discutidas na Seção V.

## II. O EXPERIMENTO ATLAS

Para caracterizar as partículas que resultam das colisões, o experimento ATLAS conta com um complexo sistema de aquisição de dados, contendo diversos circuitos eletrônicos responsáveis por condicionamento de sinais e transmissão da informação para os barramentos computacionais responsáveis por processar as informações adquiridas durante as colisões e selecionar os eventos de interesse do programa de pesquisa do experimento ATLAS. Uma das características mais importantes das partículas é a sua massa, inferida através de medidas de energia nos calorímetros.

O experimento ATLAS possui formato aproximadamente cilíndrico, com cerca de 40 m ao longo da extensão axial e 15 m na extensão radial [8]. Para descrição da posição da partícula, o ATLAS adota um sistema de coordenadas cilíndrica, sendo as variáveis  $\eta$  e  $\phi$  as principais<sup>2</sup>. O sistema de calorimetria do ATLAS ilustrado na Figura 1 é o responsável por medir a energia das partículas resultantes das colisões. A calorimetria utiliza o conhecimento da interação de diversos tipos de partícula com a matéria para a caracterização da mesma, uma vez que em sistemas de calorimetria o material ativo é conhecido. Os calorímetros mais importantes do experimento ATLAS são os calorímetros eletromagnético e hadrônico, sendo ambos em formato cilíndrico e o primeiro posicionado internamente ao segundo, compartilhando o mesmo eixo axial do experimento ATLAS [1].

O calorímetro eletromagnético, que possui três camadas de diferentes profundidades, mede a energia de partículas de mesma natureza, como elétrons e fótons, possuindo cerca de 200 mil canais de leitura. Já o calorímetro hadrônico, que também possui três camadas distintas, mede a energia de partículas hadrônicas, como prótons e jatos hadrônicos, possuindo aproximadamente 10 mil canais de leitura. Entretanto,

<sup>2</sup>O ATLAS usa um sistema de coordenadas baseado na mão direita, com sua origem no ponto de interação (PI) no centro do detector e o eixo z ao longo do feixe de prótons. O eixo x é direcionado a partir do PI até o centro do anel do LHC e o eixo y é perpendicular a este plano. As coordenadas cilíndricas ( $r$ ,  $\phi$ ) são usadas no plano transversal, com  $\phi$  sendo o ângulo azimutal em torno do eixo z. A pseudo rapidez é definida em termos do ângulo polar  $\theta$  como  $\eta \equiv -\ln[\tan(\theta/2)]$ .

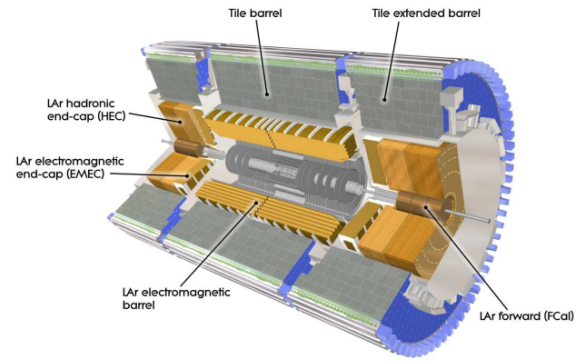


Fig. 1: Sistema de calorímetros do ATLAS. Extraído de [9]

elétrons e fótons mais energéticos podem atingir o calorímetro hadrônico, bem como jatos hadrônicos depositam energia no calorímetro eletromagnético. A caracterização das partículas nos calorímetros do experimento ATLAS se dá através da interação das partículas com as moléculas do material. Essa interação provoca o que é conhecido como chuva de partículas. Devido a sua natureza, elétrons e fótons tendem a desenvolver um chuva de partícula mais fino do que jatos hadrônicos [10].

### A. Sistema Online de Seleção de Eventos

O sistema *online* de seleção de eventos (*trigger*) do ATLAS utiliza níveis sequenciais de filtragem com o objetivo de reduzir gradualmente o ruído de fundo e mantendo a maior parte dos eventos de interesse. O primeiro nível (L1) é implementado em *hardware* dedicado para lidar com uma grande quantidade de eventos e severas restrições de tempo de resposta.

O *trigger* de alto nível (*high-level trigger*, ou HLT) é implementado em software que opera de forma paralela em um *cluster* de computadores. Um *pipeline* de memória armazena os eventos gerados pelas colisões de feixes de prótons enquanto a decisão do sistema de *trigger* não é finalizada. O algoritmo *CutBased* é o método atual empregado na identificação de fótons. Este utiliza uma estratégia onde são aplicados cortes em sequência nas variáveis obtidas pelo sistema de calorimetria, seguindo a sequência  $R_\eta$ ,  $E_{ratio}$ ,  $R_{had}$  [11]. O  $R_{had}$  é a razão de  $E_T$  na primeira camada do calorímetro hadrônico para  $E_T$  do *cluster* EM. O  $R_\eta$  é a razão da soma das energias das células contidas em um retângulo  $3 \times 7$  em  $\eta \times \phi$  (medida em unidades de célula) para a soma das energias das células em um retângulo  $7 \times 7$ , ambos centrados na célula mais energética. Por fim, o  $E_{ratio}$  é a razão da diferença de energia entre o depósito de energia máximo e o depósito de energia em um máximo secundário no *cluster* para a soma dessas energias.

O algoritmo *NeuralRinger* [12] foi inicialmente proposto para a discriminação de elétrons, atuando na etapa rápida do HLT do ATLAS desde 2017. Como elétrons e fótons possuem similaridade no desenvolvimento do chuva em sistemas de calorimetria e considerando o bom desempenho do algoritmo *NeuralRinger* no experimento ATLAS, espera-se que o mesmo

possa ser alcançado ao tentarmos discriminar fótons e jatos. O processo de construção dos anéis utiliza a posição da célula mais energética dada em  $\eta$  e  $\phi$ , sendo esta a célula o primeiro anel. Com as células adjacentes ao primeiro anel, realiza-se a soma de energia, formando assim o segundo anel de energia, conforme ilustrado na Figura 2. Este processo é repetido até atingir o número definido de anéis em cada camada, conforme descrito na Tabela I.

TABELA I: Anéis por camada do sistema de calorimetria do experimento ATLAS

Camadas	PS	EM1	EM2	EM3	H1	H2	H3
Anéis	8	64	8	8	4	4	4

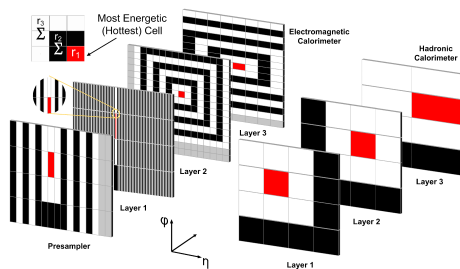


Fig. 2: Ilustração, em camadas, da construção dos anéis no sistema de calorimetria do ATLAS. Extraído de [13]

### III. METODOLOGIA PROPOSTA E DESCRIÇÃO DO CONJUNTO DE DADOS

O conjunto de dados utilizado neste trabalho foi produzido por meio de simulações de Monte Carlo, nos quais compostos por fótons (sinal) selecionados pela produção  $\gamma$ -Jet e o ruído de fundo, compostos por jatos hadrônicos que falseiam a assinatura de fótons, principalmente composta por mésons  $\eta$  e  $\pi^0$ . Uma estratégia de validação cruzada Stratified K-Fold [14] foi aplicada para estimar as incertezas estatísticas dos sistemas de classificação, sendo definido por  $K = 10$ . De modo a contornar o desbalanceamento entre o número de eventos entre as classes, foram atribuídos pesos no ajuste do treinamento para os eventos de cada classe, calculados considerando a quantidade de eventos existentes para a respectiva classe em relação ao total de eventos. Nove de cada dez subconjuntos são alocados para treinamento e um para validação/teste. Para cada configuração de treinamento, foram realizadas 5 inicializações, com o objetivo de diminuir a possibilidade do treinamento a ser escolhido no final do procedimento, tenha sofrido com mínimos locais ou problemas de convergência. Para cada inicialização, 5.000 épocas foi definido como limite. Os treinamentos foram realizados para 72 faixas<sup>3</sup> de  $E_t \times \eta$ , 10 subconjuntos, 5 inicializações,

<sup>3</sup>O sistema de calorimetria do ATLAS tem diferentes respostas em função da energia da partícula e da posição que ela interage no calorímetro, devido as diferenças de granularidades entre as células ao longo do detector. Por este motivo adotou-se a estratégia de treinamentos para faixas específicas de energia e posição de interação da partícula no detector, de modo a projetar sistemas de classificação especialistas, uma vez que os sinais de entrada apresentam diferentes perfis em função destes dois parâmetros.

totalizando em 3.600 treinamentos por arquitetura. Em cada subconjunto somente a melhor inicialização foi selecionada, sendo estimada médias de figuras de avaliação de desempenho e desvio padrão associado (relativo a todas as inicializações usando o mesmo subconjunto). Na Figura 3 são mostradas as quantidades de exemplos das duas classes de interesse para cada partição  $E_t \times \eta$ .

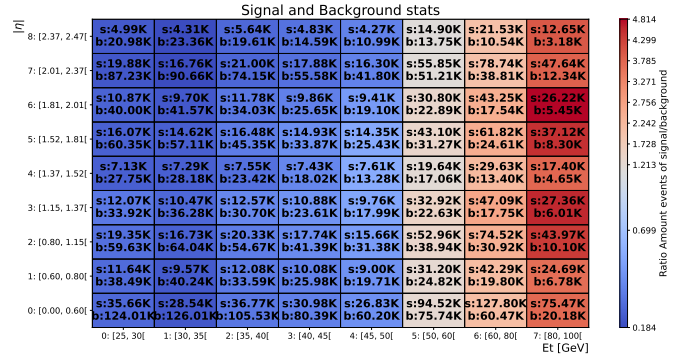


Fig. 3: Quantidade de eventos para cada bin, onde s é classe sinal, fótons, e b é a classe ruído de fundo, jatos. A tonalidade da cor representa a taxa de eventos entre sinal e background.

Para possibilitar a adequada utilização dos anéis  $r_i$  como entradas dos classificadores neurais foi utilizada a normalização conforme descrito na Equação 1:

$$r_i^{norm} = \frac{r_i}{\sum_{j=1}^N r_j} \forall i = 1, 2, \dots, N, \quad (1)$$

onde  $N$  é o número de anéis,  $r$  é valor da energia do anel.

O presente trabalho emprega a probabilidade de detecção ( $P_D$ ), a probabilidade de falso alarme ( $P_F$ ) e o índice soma-produto (SP) como medidas de avaliação de desempenho. O  $P_D$  mede a fração de instâncias relevantes que foram recuperadas. Escrito pela seguinte fórmula:

$$P_D = \frac{VP}{VP + FN}, \quad (2)$$

onde  $VP$ ,  $FN$  indicam as quantidades de verdadeiro positivo, falso negativo, respectivamente.

Por outro lado, o  $P_F$  mede a probabilidade do erro de classificação de ruído de fundo como sinal. Sendo calculado pela seguinte fórmula:

$$P_F = \frac{VN}{VN + FP}, \quad (3)$$

onde  $VN$ ,  $FP$  indicam as quantidades de verdadeiro negativo, falso positivo, respectivamente.

Por fim, o índice SP [15] foi escolhido, pois através dele é possível atingir um resultado que maximize o  $P_D$ , enquanto o  $P_F$  é minimizado, definido como:

$$SP = \sqrt{\sqrt{P_D(1 - P_F)} \times \frac{P_D + (1 - P_F)}{2}}. \quad (4)$$

### A. Projeto e Especificação do Sistema Inteligente

Neste trabalho, foram utilizadas diferentes arquiteturas de RNA, variando entre Rede Completamente Conectada (RCC) e Rede Convolutiva (RC). Os parâmetros de treinamento não foram alterados entre as arquiteturas. O treinamento utiliza o algoritmo ADAM [16], cujos parâmetros encontram-se na tabela II.

TABELA II: Parâmetros do otimizador ADAM.

Parâmetros de Otimização ADAM	
Taxa de Aprendizado	0.001
$\beta_1$	0.9
$\beta_2$	0.999
$\epsilon$	1e-07

O critério de parada do treinamento escolhido foi de 25 falhas na melhoria do SP. A função objetivo aplicada no treinamento foi a entropia cruzada binária [17]. Ao final do treinamento foi recuperado o modelo cujo SP obteve melhor desempenho. Também foi atribuído um ponto de operação (limiar de decisão da rede neural) tendo como referência a métrica  $P_d$  do método *Cut-based*. Durante o treinamento, a camada de saída é composta por um neurônio linear acrescido de um neurônio com apenas a função de ativação sigmoide. Dessa maneira, a saída do modelo continua sendo binária. Neste trabalho foram utilizadas duas arquiteturas de redes neurais, conforme descritas à seguir:

- **Rede Completamente Conectada (RCC):** Este modelo (referido como v0,) possui apenas duas camadas. Para a primeira camada diferentes variações de neurônios foram testadas, como: 2, 5, 8, 10, 15, 50 e 100. A função de ativação para estes foi a ReLU e a inicialização escolhida foi a Xavier [18]. A camada de saída permanece igual a descrição acima.
- **Rede Convolutiva (RC):** Neste caso, houve maior variação em relação ao número de camadas e quantidades de neurônios. A Tabela III possui o detalhamento em relação as camadas de entradas, lembrando que a a camada de saída sempre mantém-se inalterada. Não foi utilizada uma função de ativação na rede convolutiva e o *padding* é sempre válido. A ReLU é usada como função de ativação para as camadas RCC e após ela entra a camada de saída.

TABELA III: Tabela de nomenclaturas das diferentes versões usadas para a arquitetura da rede CNN.

Nomenclatura	Camada 1	Camada 2	Camada 3	Camada 4	Camada 5 (RCC)	Camada 6 (RCC)
v4_0	20 filtros stride = 2	-	-	-	50 neurônios	-
v4_1	20 filtros stride = 1	-	-	-	50 neurônios	-
v5_0	20 filtros kernel = 4 stride = 2	Dropout	-	-	50 neurônios	-
v5_1	20 filtros kernel = 4 stride = 1	Dropout	-	-	50 neurônios	-
v6_0	20 filtros kernel=5 stride=1	Dropout	40 filtros kernel=4 stride=1	Dropout	50 neurônios	100 neurônios
v6_1	80 filtros kernel=5 stride=1	Dropout	160 filtros kernel=4 stride=1	Dropout	100 neurônios	200 neurônios

O crescente aumento da luminosidade, em busca de uma maior quantidade de eventos gerados, ocasiona uma maior

ocupação do detector, levando ao aumento de empilhamento de sinais (*pileup*)<sup>4</sup>. Podendo afetar a reconstrução de fótons, assim será proposto um fator de correção em função do número médio de interações por cruzamento de pacotes ( $\langle \mu \rangle$ ) [19]. Este ajuste é linear e feito utilizando o método do  $\chi^2$ , calculado por:

$$\chi^2 = \frac{(y - f(x))^2}{e_y^2 + (0,5 \times (e_{xl} + e_{xh})f'(x))^2}, \quad (5)$$

onde  $f(x)$  é a função a ser ajustada, neste caso uma função afim;  $e_y$  é o erro inferior (superior) das ordenadas, caso  $f(x)$  esteja abaixo (acima) de  $y$ , e  $e_{xl}$  ( $e_{xh}$ ) é o erro inferior (superior) nas abscissas. Este ajuste foi aplicado em todos os modelos após o treinamento. O ajuste é realizado com a remoção da função de ativação da camada de saída (*tanh*) de maneira a evitar a propagação da não-linearidade.

## IV. RESULTADOS

Conforme apresentado na Tabela IV, comparando-se a arquitetura Rede Completamente Conectada (RCC), a que apresentou o melhor desempenho foi a 100 neurônios, com o maior valor médio de SP de 87,64% e menor valor médio de  $P_F$  com 12,12%. Enquanto que a comparação da Rede Convolutiva (RC) pela Tabela V, a v6\_1 obteve o melhor desempenho médio.

TABELA IV: Desempenho Médio para todas as regiões da arquitetura RCC

Neurônios	SP (%)	$P_F$ (%)	$P_D$ (%)
2	85,17 ± 2,56	14,19 ± 2,71	84,54 ± 2,92
5	85,69 ± 2,47	13,60 ± 2,66	84,99 ± 2,81
8	85,95 ± 2,31	13,36 ± 2,46	85,27 ± 2,74
10	86,05 ± 2,25	13,27 ± 2,40	85,39 ± 2,65
15	86,20 ± 2,17	13,16 ± 2,35	85,57 ± 2,57
20	86,31 ± 2,10	13,09 ± 2,22	85,72 ± 2,50
50	86,59 ± 1,95	12,83 ± 2,07	86,02 ± 2,36
100	<b>86,78 ± 1,85</b>	<b>12,67 ± 1,96</b>	<b>86,24 ± 2,26</b>

TABELA V: Desempenho Médio para todas as regiões da arquitetura RC

Nomenclatura	SP (%)	$P_F$ (%)	$P_D$ (%)
v4_0	86,93 ± 1,75	12,56 ± 1,92	86,43 ± 2,19
v4_1	87,04 ± 1,69	12,47 ± 1,79	86,56 ± 2,18
v5_0	86,95 ± 1,71	12,50 ± 1,81	86,42 ± 2,12
v5_1	87,06 ± 1,68	12,48 ± 1,87	86,61 ± 2,11
v6_0	87,47 ± 1,58	12,22 ± 1,73	87,17 ± 2,01
v6_1	<b>87,64 ± 1,49</b>	<b>12,12 ± 1,68</b>	<b>87,41 ± 1,90</b>

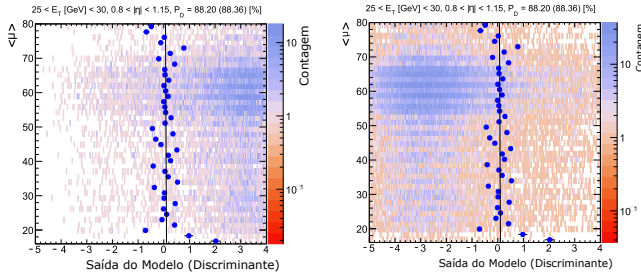
Após aplicar a restrição de  $P_D$  de acordo com a referência e escolhendo o melhor modelo de cada arquitetura, é possível verificar, na Tabela VI que todas as arquiteturas obtiveram uma melhora significativa em comparação ao método de referência, obtendo aumento significativo do SP (mais de 15 pontos percentuais) e diminuição expressiva do  $P_F$  (mais de 30 pontos percentuais).

<sup>4</sup>Número médio de interações entre prótons por cruzamento de feixe.



TABELA VI: Desempenho Após Ajuste de  $P_D$  em Função da Referência

Método	SP (%)	$P_F$ (%)	$P_D$ (%)
<i>Cut-Based</i>	70,66	41,28	83,70
RCC	$86,89 \pm 0,07$	$9,86 \pm 0,15$	$83,70 \pm 0,01$
RC	<b><math>87,03 \pm 0,07</math></b>	<b><math>9,58 \pm 0,14</math></b>	<b><math>83,70 \pm 0,01</math></b>



(a) Sinal

(b) Ruído de fundo

Fig. 4: Ajuste de empilhamento para a distribuição de sinal (a), e ruído de fundo (b).

Uma segunda etapa do desenvolvimento do classificador consistiu em usar um limiar de decisão como função do empilhamento de sinais ( $\mu$ ). Esse ajuste trata-se na substituição da função sigmoide utilizada no treinamento por uma função linear, sendo o resultado do ajuste exibido na Figura 4. Este permitiu constatar que a classificação não está sendo impactada pelo crescente aumento do empilhamento. Uma vez que o coeficiente angular é praticamente nulo, e os pontos seguem aproximadamente uma reta vertical. A Figura 4a mostra que a maioria dos fótons estão localizados à direita da reta, enquanto a Figura 4b mostra que boa parte dos jatos hadrônicos encontram-se a esquerda da linha de decisão, mostrando o potencial de classificação dos modelos neurais. Por usar uma função linear na camada de saída, a saída do modelo neural não foi limitada entre 0 e 1.

## V. CONCLUSÕES

O experimento ATLAS, um dos maiores no complexo de aceleração do LHC, é responsável por medir e caracterizar os subprodutos advindos das colisões. Com a alta taxa de colisão, os fenômenos de interesse são intensamente contaminados por ruído, dificultando sua correta identificação. Nesse contexto, o presente trabalho propôs um conjunto de redes neurais para a seleção de fótons, cujos fenômenos associados são fortemente contaminados por jatos hadrônicos. Neste trabalho, os modelos neurais foram capazes de diminuir a taxa de aceitação de falsos fótons quando comparado com o método de seleção tradicional do experimento (*Cut-Based*), mantendo-se a mesma eficiência na detecção de fótons reais. Isso garante os eventos de interesse continuem sendo detectados com menor contaminação de ruído de fundo. Para isto, foram estudadas diferentes arquiteturas de classificadores e as redes convolucionais foram capazes de aumentar o índice  $P_F$  em  $\sim 0.3$  ponto percentual em relação à RCC e mais de 31 pontos percentuais se comparado ao classificador *Cut-Based*.

## AGRADECIMENTOS

Os autores agradecem ao CNPq, à FAPERJ, à FAPESB e à RENAFAE pelo apoio financeiro. Este estudo foi financiado em parte pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código Financeiro 001 e Programa CAPES-COFECUB.

## REFERÊNCIAS

- [1] ATLAS Collaboration, “The ATLAS experiment at the CERN Large Hadron Collider,” *Journal of Instrumentation*, vol. 3, no. 08, pp. S08 003–S08 003, aug 2008. [Online]. Available: <https://doi.org/10.1088/1748-0221/3/08/S08003>
- [2] G. E. Orellana and ATLAS Collaboration, “Projected ATLAS Electron and Photon Trigger Performance in Run 3,” CERN, Geneva, Tech. Rep., 2021. [Online]. Available: <https://cds.cern.ch/record/2730417>
- [3] ATLAS Collaboration, “Search for higgs boson pair production in the two bottom quarks plus two photons final state in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the atlas detector,” *Phys. Rev. D*, vol. 106, p. 052001, Sep 2022. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevD.106.052001>
- [4] A. P. Morais, A. Onofre, F. F. Freitas, J. Gonçalves, R. Pasechnik, and R. Santos, “Deep Learning Searches for Vector-Like Leptons at the LHC and Electron/Muon Colliders,” *Eur. Phys. J. C*, vol. 83, no. 3, p. 232, 2023, 26 pages, 11 figures, 10 tables, Published version. [Online]. Available: <https://cds.cern.ch/record/2853391>
- [5] S. Spencer, T. Armstrong, J. Watson, S. Mangano, Y. Renier, and G. Cotter, “Deep learning with photosensor timing information as a background rejection method for the cherenkov telescope array,” *Astroparticle Physics*, vol. 129, p. 102579, 2021.
- [6] J. Marin, E. Simas, E. Souza, B. Peralva, J. Pinto, M. Araújo, and J. Seixas, “Conjunto de redes neurais artificiais para detecção online de fótons de altas energias,” in *Anais do 15 Congresso Brasileiro de Inteligência Computacional*, C. J. A. B. Filho, H. V. Siqueira, D. D. Ferreira, D. W. Bertol, and R. C. L. ao de Oliveira, Eds. Joinville, SC: SBIC, 2021, pp. 1–8.
- [7] J. L. Marin, E. F. de Simas Filho, B. S. Peralva, E. E. P. de Souza, and J. M. de Seixas, “Uso de aprendizado de máquina para detecção online de fótons de altas energias,” in *Anais do XVI Congresso Brasileiro de Inteligência Computacional (CBIC’2023)*, E. Simas, D. D. Ferreira, and L. R. Oliveira, Eds. Salvador, BA: SBIC, 2023, pp. 1–8.
- [8] L. Evans and P. Bryant, “LHC machine,” *Journal of Instrumentation*, vol. 3, no. 08, pp. S08 001–S08 001, aug 2008.
- [9] J. Pequenaço, “Computer Generated image of the ATLAS calorimeter,” 2008. [Online]. Available: <https://cds.cern.ch/record/1095927>
- [10] R. Wigmans, *Calorimetry: Energy Measurement in Particle Physics*, ser. International series of monographs on physics. Clarendon Press, 2017.
- [11] M. Aaboud *et al.*, “Performance of the ATLAS Trigger System in 2015,” *Eur. Phys. J. C*, vol. 77, no. 5, p. 317, 2017.
- [12] J. M. Seixas, L. P. Calôba, M. N. Souza, A. L. Braga, and A. P. Rodrigues, “Neural second-level trigger system based on calorimetry,” 1995. [Online]. Available: <https://cds.cern.ch/record/278995>
- [13] G. Aad *et al.*, “Performance of electron and photon triggers in ATLAS during LHC Run 2,” *Eur. Phys. J. C*, vol. 80, no. 1, p. 47, 2020.
- [14] S. Raschka, “Model evaluation, model selection, and algorithm selection in machine learning,” 2020.
- [15] E. F. Simas Filho, J. M. de Seixas, and L. P. Calôba, “Modified post-nonlinear ica model for online neural discrimination,” *Neurocomputing*, vol. 73, no. 16, pp. 2820–2828, 2010, 10th Brazilian Symposium on Neural Networks (SBRN2008). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0925231210002511>
- [16] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2017.
- [17] A. Mao, M. Mohri, and Y. Zhong, “Cross-entropy loss functions: Theoretical analysis and applications,” 2023.
- [18] Y. Bengio and X. Glorot, “Understanding the difficulty of training deep feed forward neural networks,” *International Conference on Artificial Intelligence and Statistics*, pp. 249–256, 01 2010.
- [19] G. Aad *et al.*, “Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data,” *JINST*, vol. 14, no. 12, p. P12006, 2019.