

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

Luiz Eduardo B. Filho¹, José M. de Seixas¹, Edmar Egidio P. de Souza², Eduardo F. S. Filho²,
Juan L. Marin³, Bernardo S. M. Peralva⁴

¹Laboratório de Processamento de Sinais, COPPE/POLI, UFRJ

²Laboratório de Sistemas Digitais, PPGEEC/UFBA

³Instituto Federal da Bahia, Vitória da Conquista-BA, IFBA

⁴Universidade do Estado do Rio de Janeiro, IPRJ.

Emails: {balabram, seixas}@lps.ufrj.br, {edmar.egidio, eduardo.simas}@ufba.br,
juan.marin@ifba.edu.br, bernardo@iprj.uerj.br

Dentre o vasto programa de física do experimento ATLAS do LHC, fótons estão envolvidos em distintos modelos físicos, que podem estar relacionados com interações ainda não observadas experimentalmente. Para identificar fótons um sistema sequencial de seleção *online* de eventos (*trigger*) torna-se necessário, uma vez que durante as colisões de prótons, ≈ 52 TB/s de informação pode ser produzida, na qual a maior parte da informação produzida trata-se de ruído de fundo. Neste trabalho, uma estratégia inteligente de identificação *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. Diferentes arquiteturas de redes neurais artificiais foram projetadas com a informação dos anéis como entradas, para discriminar 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.

Redes Neurais Convolucionais, Fótons, Reconhecimento de Padrões.

Resumo—Among the vast particle physics program of the ATLAS experiment at LHC, photons are involved in different physics models, which may be related to phenomena that have not yet been verified experimentally. To identify photons, a 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 online intelligent photon identification strategy is proposed. The energy measured from ATLAS calorimeters is then shaped into concentric rings around the cell with the highest deposited energy. Convolution 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.

Convolutional Neural Networks, 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, 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 o maior acelerador 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 100m. No LHC, as colisões entre os feixes de prótons ocorrem em diferentes pontos de cruzamento de feixes em até 14 TeV¹, a uma taxa de 40 MHz, o que corresponde a uma colisão a cada 25ns.

O experimento ATLAS (*A Toroidal LHC ApparatuS*) [1], é um dos experimentos do LHC, instalado em um dos pontos de colisão de prótons. Sendo composto por sub-detecores capazes de medir o subproduto das diferentes partículas geradas nas colisões. Em geral, aproximadamente 1,3 MB de informação é produzida 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, que torna proibitivo o arquivamento em mídia permanente de toda a informação produzida nas colisões. Deste modo, é requerido um sofisticado sistema de seleção sequencial *online* (*trigger*) [2], para identificar os eventos relevantes que são então armazenados para análises *offline* e descartar o ruído de fundo produzido nos distintos canais físicos de interesse. Este sistema de seleção em *software* basicamente é dividido em duas etapas, em que a primeira é dotada de algoritmos rápidos (porém menos precisos) para identificação de candidatos a fótons (removendo previamente uma parcela de ruído de fundo). Com diminuição da banda passante, algoritmos mais precisos (que utilizam mais

¹O (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.

informação dos calorímetros) são executados processando a decisão da etapa anterior para uma melhor refinamento dos eventos a serem salvos em mídia permanente.

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 na área de física de altas energias, como detecção de partículas, reconstrução de eventos e filtragem *online* [4], [5]. Tais aplicações tem possibilitado 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.

Nos trabalhos [6], [7] o algoritmo NeuralRinger foi proposto para identificação *online* de fótons, 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 calorimetria². Estes sinais extraídos alimentam redes neurais artificiais com uma única camada oculta, que mostraram melhores resultados se comparados a estratégia de seleção por cortes rígidos em variáveis unidimensionais (*CutBased*).

A etapa rápida do sistema de *trigger* de fótons do ATLAS até o fim de 2024 utilizava variáveis unidimensionais baseadas na deposição de energia, extraídas do sistema de calorimetria, aplicando cortes rígidos sequenciais nestas grandezas discriminantes, para classificação sinal/ruído. Estas grandezas (*Shower Shapes*) descrevem basicamente o grau de espalhamento lateral do chuveiro de partículas, bem como a proporção de deposição de energia entre as camadas do calorímetro. Entretanto, com o aumento progressivo da energia das colisões, bem como da luminosidade³, sistemas inteligentes apresentam-se como uma alternativa para melhoria na rejeição de falsos fótons. O NeuralRinger em sua proposta original, utilizando redes neurais do tipo MLP (*Multilayer Perceptron*) foi adotado como sistema padrão de identificação *online* de fótons, para a operação de tomada de dados em 2025. Esta inovação foi realizada após um extenso período de validação, com comparações de desempenho entre o projeto realizado com dados simulados, e dados experimentais de algumas rodadas realizadas em 2024, no qual proporcionou uma redução importante na taxa de falsos positivos e liberação de banda passante no sistema de classificação de fótons. Todavia, o método pode ser aprimorado, por meio da avaliação do seu

desempenho com estratégias de aprendizagem de máquina mais sofisticadas. Neste trabalho, extensões do NeuralRinger são propostas com intuito de melhorar o desempenho na redução de ruído de fundo, investigando diferentes arquiteturas e configurações com modelos sofisticados de Redes Neurais Convolucionais, nas quais foram comparadas com redes neurais rasas de uma única camada. Uma análise exploratória dos parâmetros internos das redes convolucionais projetadas foi realizada, de modo a compreender o impacto no desempenho em relação a quantidade de filtros, camadas e número de neurônios incluídos na camada densa.

Este artigo está dividido do seguinte modo: na Seção II o experimento ATLAS, seus subsistemas e o sistema *online* de identificação de fótons são apresentados. 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 η e ϕ as principais⁴. 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 na identificação de fótons 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, projetado para medir energia de partículas como elétrons e fótons, possuindo cerca de 200 mil canais de leitura. Já o calorímetro hadrônico, que também possui três camadas distintas, é projetado para medir

²A largura do chuveiro de partículas é um dos principais atributos utilizados para identificação de partículas de natureza eletromagnética em relação a partículas de natureza hadrônica.

³Medida do número de colisões por centímetro quadrado produzida a cada segundo.

⁴O 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 , ϕ) são usadas no plano transversal, com ϕ sendo o ângulo azimutal em torno do eixo z. A pseudo rapidez é definida em termos do ângulo polar θ como $\eta \equiv -\ln [\tan(\theta/2)]$.

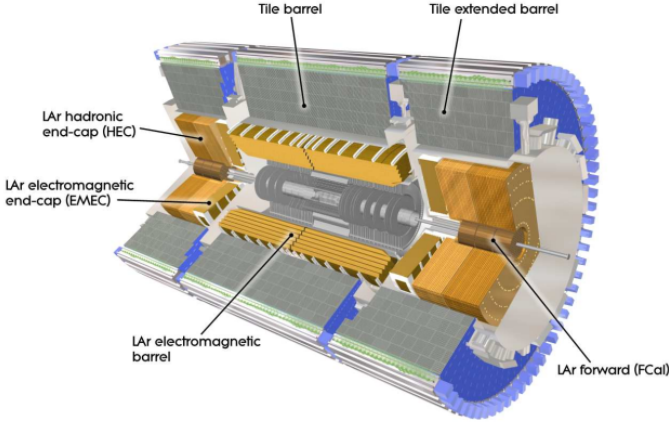


Figura 1: Sistema de calorímetros do ATLAS. Extraído de [9]

a energia de partículas hadrônicas, como prótons e jatos hadrônicos, possuindo aproximadamente 10 mil canais de leitura. Entretanto, elétrons e fótons mais energéticos podem atingir o calorímetro hadrônico, bem como a subestrutura de 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], e tal característica é útil para sua identificação.

A. Sistema Online de Seleção de Eventos

O sistema *online* de seleção de eventos (*trigger*) do ATLAS utiliza níveis sequenciais de seleção de eventos 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*, tradicionalmente utilizado em calorimetria de altas energias, utiliza uma estratégia onde são aplicados cortes em sequência nas variáveis obtidas pelo sistema de calorimetria, seguindo a sequência R_η , E_{ratio} , R_{had} [11]. O R_{had} é a razão de energia depositada na primeira camada em uma janela de células (*cluster*) do calorímetro hadrônico em relação a energia depositada em um *cluster* no calorímetro eletromagnético. O R_η é a razão da soma das energias das células contidas em um retângulo 3×7 em $\eta \times \phi$ para a soma das energias das células em um retângulo 7×7 , 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* de células do calorímetro, para a soma dessas energias.

O algoritmo *NeuralRinger* [12] foi inicialmente proposto para a discriminação de elétrons, atuando como sistema padrão na etapa rápida do HLT do ATLAS desde 2017. O processo de construção dos anéis utiliza a posição da célula mais energética dada em η e ϕ , 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

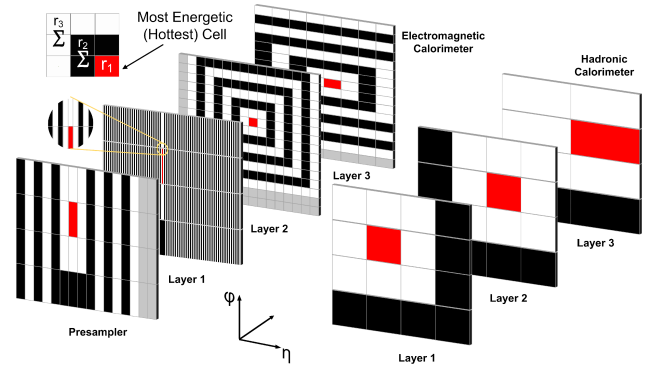


Figura 2: Ilustração, em camadas, da construção dos anéis no sistema de calorimetria do ATLAS. Extraído de [13]

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 γ -Jet e o *ruído de fundo*, compostos por jatos hadrônicos que falseiam a assinatura de fótons, principalmente composta por mésons eta e π^0 . Uma estratégia de validação cruzada Stratified K-Fold [14] foi aplicada para estimar as incertezas estatísticas dos sistemas de classificação, sendo definido de modo heurístico com 10 subconjuntos ($K = 10$).

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 do algoritmo de otimização. Para cada inicialização, 5.000 épocas foi definido como limite.

Os treinamentos foram realizados para 72 faixas⁵ de $E_T \times \eta^6$, 10 subconjuntos, 5 inicializações, 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 $E_t \times \eta$.

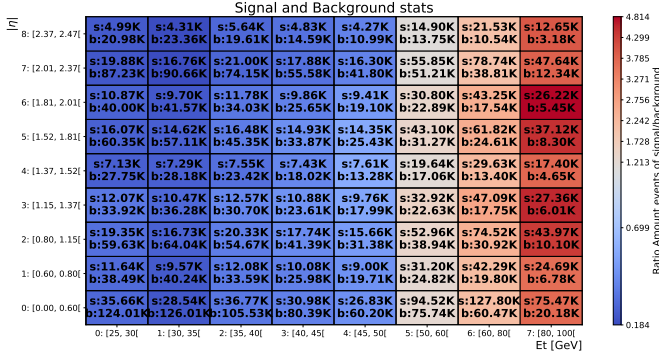


Figura 3: Quantidade de eventos para cada bin, onde s é classe sinal, fótons, e b é a classe ruído de fundo, jatos. A tonalidade da cor representa a taxa de eventos entre sinal e ruído de fundo.

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 somaproduto (SP) como medidas de avaliação de desempenho. O P_D mede a fração de exemplos da classe de sinal que foram corretamente identificados, conforme equação 2:

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

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 equação 3:

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

⁵O 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.

⁶ E_T é a energia transversa da partícula.

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 pela equação 4:

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

A. Projeto e Especificação do Sistema Inteligente

Neste trabalho, foram utilizadas diferentes arquiteturas de RNA, variando entre Rede Completamente Conectada (RCC) e Rede Convolutacional (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
β_1	0.9
β_2	0.999
ϵ	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 medida 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 (MLP - *MultiLayer Perceptron*). Para a primeira camada diferentes variações de neurônios foram avaliadas de modo exploratório, 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].
- **Rede Convolutacional (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. Não foi utilizada uma função de ativação na rede convolutacional e o *padding* é foi mantido válido para todas as arquiteturas avaliadas. A ReLu é usada como função de ativação para as camadas RCC.

O crescente aumento da luminosidade, em busca de uma maior quantidade de eventos a serem gerados, ocasiona uma maior ocupação do detector, levando ao aumento de empilhamento de sinais (*pileup*)⁷. Considerando que o empilhamento

⁷Número médio de interações entre prótons por cruzamento de feixe.

Tabela III: Tabela de nomenclaturas das diferentes versões usadas para a arquitetura da RC.

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

pode perturbar a reconstrução de fótons, um fator de correção em função do número médio de interações por cruzamento de pacotes ($\langle\mu\rangle$) [19] é proposto para ajustar o patamar de decisão da rede neural em função de ($\langle\mu\rangle$), mantendo a mesma taxa de P_D . Este ajuste é linear e feito utilizando o método do χ^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 abscisas. 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 Convolucional (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 $\pm$ 2,56	14,19 $\pm$ 2,71	84,54 $\pm$ 2,92
5	85,69 $\pm$ 2,47	13,60 $\pm$ 2,66	84,99 $\pm$ 2,81
8	85,95 $\pm$ 2,31	13,36 $\pm$ 2,46	85,27 $\pm$ 2,74
10	86,05 $\pm$ 2,25	13,27 $\pm$ 2,40	85,39 $\pm$ 2,65
15	86,20 $\pm$ 2,17	13,16 $\pm$ 2,35	85,57 $\pm$ 2,57
20	86,31 $\pm$ 2,10	13,09 $\pm$ 2,22	85,72 $\pm$ 2,50
50	86,59 $\pm$ 1,95	12,83 $\pm$ 2,07	86,02 $\pm$ 2,36
100	86,78 $\pm$ 1,85	12,67 $\pm$ 1,96	86,24 $\pm$ 2,26

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 do índice SP (acima de 15 pontos

Tabela V: Desempenho Médio para todas as regiões da arquitetura RC.

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

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	87,03 $\pm$ 0,07	9,58 $\pm$ 0,14	83,70 $\pm$ 0,01

percentuais) e diminuição da taxa de P_F (redução de 30 pontos percentuais).

Uma segunda etapa do desenvolvimento do classificador consistiu em definir um limiar de decisão como função do empilhamento de sinais $\langle\mu\rangle$. Esse ajuste foi realizado a partir da substituição da função sigmoide utilizada durante o treinamento por uma função linear. Este permitiu avaliar o impacto da tomada de decisão das redes neurais em função do nível empilhamento. Os pontos azuis indicados nas Figuras 4 e 5 indicam o valor do patamar de decisão da rede neural, para manter uma eficiência de detecção de fótons em 88,20%.

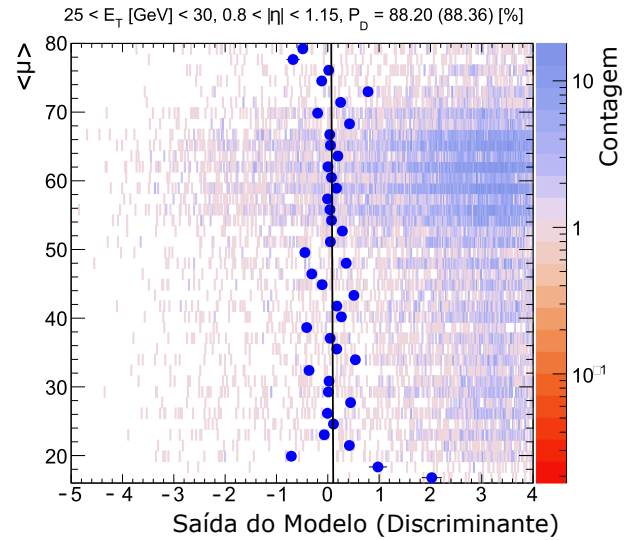


Figura 4: Ajuste do patamar de decisão em função do empilhamento para a classe de sinal.

A Figura 4 mostra que a maior parcela dos fótons estão localizados à direita da reta, enquanto a Figura 5 mostra que a maior parte dos jatos hadrônicos encontram-se a esquerda

da linha de decisão. Por usar uma função linear na camada de saída, a saída do modelo neural não foi limitada entre 0 e 1. Para a faixa de empilhamento entre ≈ 20 e 30, os patamares apresentam-se mais à direita, indicando que a maior parte dos fótons dessa faixa ficam mais próximos do limite de saturação de decisão das redes neurais, indicando forte probabilidade de acerto.

Para as demais faixas de empilhamento, os pontos azuis apresentam variações em torno da aproximação da reta linear, o que torna inconclusivo a hipótese de tais variações terem sido provocadas pelo efeito do empilhamento da amostra, ou se a variação é resultado da flutuação estatística presente no conjunto de dados, uma vez que a distribuição das amostras em função do empilhamento não apresenta uma distribuição normal, e sim com a presença de um pico na região de $\langle \mu \rangle \approx 60$ (muito mais amostras concentradas em poucas faixas de empilhamento).

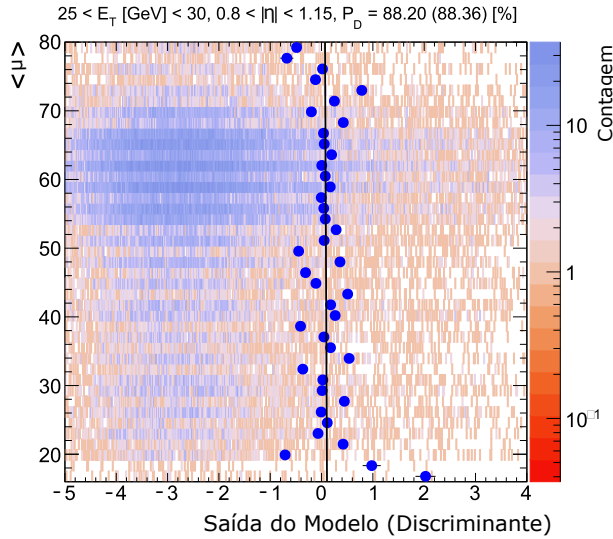


Figura 5: Ajuste do patamar de decisão em função do empilhamento para a classe ruído de fundo (jatos hadrônicos).

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 alimentados por sinais formatados em anéis, 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 reduzir a taxa P_F em ~ 0.3 ponto percentual em relação à RCC e mais de 31 pontos percentuais se comparado ao discriminante *Cut-Based*. Como trabalhos futuros, uma avaliação de desempenho para verificar se o tempo de processamento das redes convolucionais propostas são adequadas frente às restrições de execução do *trigger* é requerida. Um estudo dedicado para entender o impacto do empilhamento de sinais nas amostras deve ser conduzido, considerando o reponderamento das amostras em função do empilhamento serão avaliadas para compreender melhor os resultados do aprendizado dos modelos inteligentes. Uma avaliação em relação ao número mínimo de anéis que podem ser utilizados (incluindo uma análise de relevância), e que atendam as eficiências exigidas também será investigada.

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. Pequeno, “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.