

Informação Especialista para um Sistema de Filtragem *Online* com *Ensemble* de Redes Neurais em um Ambiente com Alta Taxa de Eventos

Edmar E. P. de Souza¹, Juan L. Marin¹, Eduardo F. S. Filho¹, Paulo C. M. A. Farias¹,
João V. F. Pinto², Micael V. de Araújo², José M. de Seixas²

¹Laboratório de Sistemas Digitais, PPGEE/UFBA

²Laboratório de Processamento de Sinais, COPPE/POLI, UFRJ

Emails: {edmar.egidio, juan.lieber, eduardo.simas, paulo.farias}@ufba.br,
{jodafons,micael.verissimo, seixas}@lps.ufrj.br

Resumo—O experimento ATLAS é projetado para observar os constituintes fundamentais em colisões de partículas no Grande Colisor de Hádrons (LHC). Na tomada de dados a partir de 2021, espera-se aumentar a energia de colisão e a taxa de produção de informações, o que requer melhorias no sistema de filtragem *online*. O NeuralRinger é o algoritmo padrão de identificação *online* de elétrons no ATLAS (etapa rápida) e usa informações do calorímetro (medição de energia) organizadas em anéis concêntricos para alimentar um conjunto de redes neurais. Uma das principais características utilizadas para a discriminação de elétrons e jatos hadrônicos é o perfil lateral do chuveiro de partículas de energia decrescente medido no calorímetro, pois o chuveiro desenvolvido por partículas eletromagnéticas tende a ser menos largo lateralmente se comparado a chuveiros hadrônicos. Neste trabalho, é investigada a utilização do perfil lateral do chuveiro combinado com os anéis. O método proposto pondera os anéis gerados pelo NeuralRinger pela energia e distância até a célula mais energética. Os resultados indicaram uma redução do falso alarme de 1,37 pontos percentuais quando comparado ao desempenho do NeuralRinger em dados experimentais o que é relevante, haja vista a alta taxa de eventos gerados no ATLAS.

Palavras-chave - Redes Neurais, NeuralRinger, Física de Altas Energias, Identificação de elétrons

I. INTRODUÇÃO

Experimentos modernos de física de altas energias (*High Energy Physics* - HEP) requerem notável infraestrutura tecnológica para observação de partículas e interações com rara ocorrência [1] [2]. Em geral, durante a tomada de dados a maior parte da informação coletada já é conhecida (ruído de fundo), exigindo sistemas eficientes de filtragem *online* para preservar o evento de interesse (sinal) e rejeitar o ruído de fundo [3]. A filtragem de eventos em HEP é uma tarefa desafiadora porque uma enorme quantidade de dados é gerada [4] e o ruído de fundo pode apresentar assinatura característica semelhante ao sinal. Portanto, a seleção de eventos em HEP é estimulante para a utilização de técnicas de aprendizado de máquina [5], que tem ganhado destaque em aplicações variadas, como rejeição de ruído de fundo [6] em experimentos de astropartículas e estudos de dinâmica de feixes de partículas [7] em aceleradores.

Estes aspectos podem ser encontrados no experimento ATLAS (*A Toroidal LHC Apparatus*) [8] do acelerador de partículas LHC (*Large Hadron Collider*) [9] [10], instalado no CERN (Organização Europeia para Pesquisa Nuclear), Suíça. O ATLAS é o maior detector de partículas de propósito geral do LHC. É composto por diversos subsistemas de detecção, dentre eles o calorímetro com fina segmentação, com mais de 100.000 sensores responsáveis por medir a energia depositada pelo subproduto das colisões [11].

Elétrons são partículas de interesse destacado pelo ATLAS pois podem estar relacionadas a eventos físicos raros ainda não observados experimentalmente [12]. Um dos principais eventos de interesse são assinaturas relacionadas com o bóson Z, que produz decaimento de dois elétrons ($Z \rightarrow e^+e^-$), em que há forte interesse pelo experimento na detecção desse par combinado de assinaturas. Entretanto, jatos de natureza hadrônica produzidos nas colisões podem apresentar perfil de deposição de energia semelhante aos elétrons, o que dificulta o desempenho de classificação de assinaturas de interesse. Para lidar com essas restrições o ATLAS utiliza um sistema *online* de filtragem (*trigger*) dividido em duas etapas sequenciais. O primeiro nível de *trigger* (L1) é implementado em *hardware* e um segundo nível de filtragem, o *trigger* de alto nível - HLT (*High Level Trigger*) implementado em *software*.

Desde 2017 o ATLAS utiliza como algoritmo de identificação de elétrons no primeiro estágio do HLT (etapa rápida), um sistema de classificação baseado em um *ensemble* de redes neurais artificiais (*NeuralRinger*) [13] [14]. O *ensemble* foi alimentado por anéis topológicos que medem a energia em torno da célula mais energética [15]. O *NeuralRinger* foi desenvolvido com classificadores treinados em intervalos específicos de energia e diferentes regiões do detector. Esta estratégia permitiu que os classificadores neurais possam explorar as diferentes respostas do calorímetro dependendo da faixa de energia da partícula, bem como as distintas granularidades dos sensores em cada região do sistema de calorimetria [8].

No próximo período de tomada de dados agendado para 2021-2024, o LHC irá operar com 14 TeV que é a máxima

energia em colisões na qual foi projetado. O LHC realiza colisões de prótons em uma taxa de 40 MHz e considerando que um evento medido pelo ATLAS tem $\approx 1,5$ MBytes, poderá ser gerado até 60 TBytes/s de informação a ser processada pelo experimento [16]. Espera-se ainda um aumento no número médio de interações por cruzamento de feixes, que pode dificultar a detecção da física de interesse com o possível empilhamento de sinais medidos no sistema de calorimetria (*pile-up*). Este aumento da ocupação do calorímetro distorce o padrão de deposição de energia das partículas, degradando a eficiência dos métodos de estimação de energia. Fato que compromete as medidas de calorimetria utilizadas pelo HLT. Este problema impacta o desempenho dos algoritmos de filtragem *online*, o que requer investigações para tornar os algoritmos de classificação do sistema de filtragem resilientes ao aumento do empilhamento de sinais [16].

Visando melhorias na filtragem *online*, o *NeuralRinger* vem sendo otimizado para atendimento às exigências do experimento. Uma das principais características adotadas para distinção de elétrons e jatos hadrônicos é o perfil lateral do chuveiro de partículas desenvolvido no calorímetro [17]. No trabalho [18], o perfil lateral do chuveiro foi usado em um mapeamento multidimensional de anéis topológicos em uma investigação primária para condições previstas para o LHC. A inclusão desta informação especializada em calorimetria (largura do chuveiro), permitiu a melhoria na detecção de elétrons e redução do ruído de fundo [18].

Este trabalho apresenta contribuições para a melhoria na detecção *online* de elétrons no ATLAS. O método proposto pondera o perfil dos anéis pela energia e distância do anel à região de interesse, incluindo a informação especialista da largura do chuveiro de partículas no sinais medidos. Os coeficientes exponenciais que ponderam a energia e a distância são ajustados iterativamente ao longo do treinamento de cada rede neural do *ensemble*. Este trabalho está dividido do seguinte modo: na seção 2 detalhes do experimento ATLAS e o sistema de filtragem online são apresentados. Na seção 3, a metodologia proposta é discutida. Os resultados encontrados com a metodologia proposta e comparações com o algoritmo de filtragem de referência são apresentados na seção 4. Na seção 5 as conclusões e trabalhos futuros são discutidos.

II. O EXPERIMENTO ATLAS

O esquemático do experimento ATLAS¹ é apresentado na Fig. 1. Com aproximadamente 45m de comprimento e 25m de altura o ATLAS, é composto por um detector de traço na região mais interna em torno do ponto de interação, rodeado por calorímetros e câmara de múons. O detector de traço está imerso em um campo magnético de $\approx 2T$ produzido por

¹O ATLAS usa um sistema de coordenadas baseado na regra da mão direita, com sua origem no ponto de interação (PI) no centro do detector e no eixo z ao longo do feixe. O eixo x aponta do PI para o centro do anel do LHC e o eixo y é perpendicular a este plano. Coordenadas cilíndricas (r, ϕ) são usadas no plano transversal, sendo ϕ o ângulo azimutal em torno do feixe. A pseudo-rapidez é definida em termos do ângulo polar θ (ângulo entre x e y) como $\eta = -\ln[\tan(\theta/2)]$. A distância angular ΔR é definida como $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.

um solenóide supercondutor e fornece reconstrução precisa de traços de partículas carregadas em uma faixa de pseudo-rapidez [8].

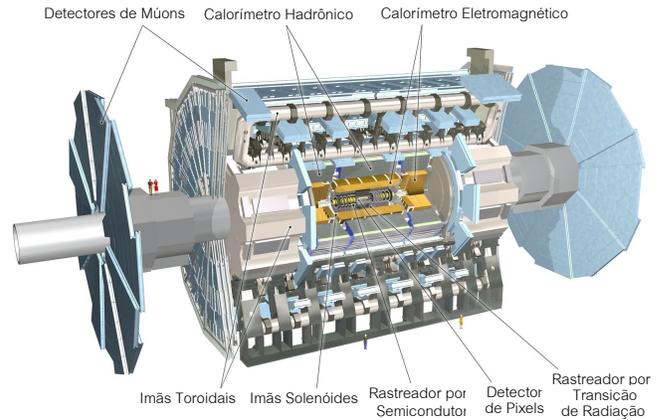


Figura 1: O experimento ATLAS e seus principais subsistemas de detecção. Adaptado de [8].

O sistema de calorimetria conforme indicado na Fig. 2 possui seções eletromagnéticas (EM) e hadrônicas (HAD). Ele é projetado para fornecer uma cobertura total em ϕ e cobre a faixa de pseudo-rapidez $|\eta| < 4,9$, com granularidade fina sobre a região do detector interno. O calorímetro EM é um calorímetro de amostragem de chumbo/argônio líquido (*Liquid-Argon - LAr*) com geometria de acordeão [8].

O barril (região central do detector) é dividido em duas metades ($-1,475 < \eta < 0$ e $0 < \eta < 1,475$) e dois componentes da extremidade ($1,375 < |\eta| < 3,2$). A região de transição entre o barril e as tampas das extremidades ($1,37 < |\eta| < 1,52$) contém material inativo adicional (*crack*). Na região dedicada às medições precisas ($|\eta| < 2,5$) (excluindo o *crack*), o calorímetro EM é segmentado em três camadas longitudinais (EM1, EM2 e EM3) na direção da profundidade do chuveiro. Três camadas de calorímetro hadrônico circundam o calorímetro EM (HAD0, HAD1 e HAD2) utilizadas para medir energia de partículas com natureza hadrônica. Na frente do calorímetro acordeão, uma camada chamada pré-amostrador (PS) fina, cobrindo o intervalo de pseudo-rapidez $|\eta| < 1,8$, é usada para corrigir a perda de energia das partículas antes de chegarem ao calorímetro.

III. Trigger DE ELÉTRONS NO DETECTOR ATLAS

Um sistema de *trigger*, descrito na Fig. 3, projetado em dois níveis sequenciais é usado para filtrar candidatos a eventos de interesse no ATLAS [13]. O *trigger* de primeiro nível (L1) é implementado em *hardware* dedicado para reduzir a taxa de eventos da taxa de cruzamento de feixe de 40 MHz para menos de 100 kHz. Esta etapa é limitada em $2,5 \mu s$ para decidir quais eventos serão mantidos. O L1 também define as Regiões de Interesse (RoIs) que têm aglomerados de células do calorímetro (*clusters*) com alta E_T ou traços de múon nas câmaras de múon.

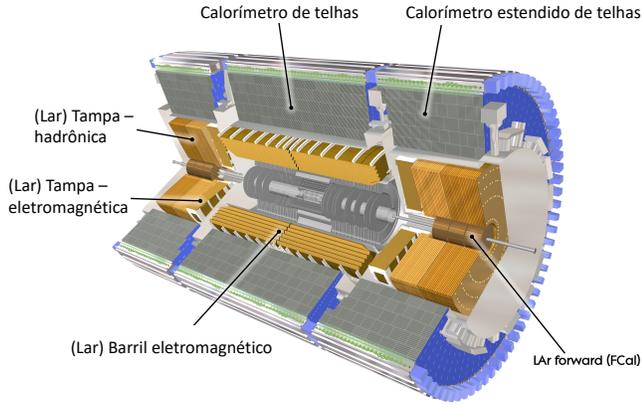


Figura 2: Esquema detalhado dos calorímetros do ATLAS. Adaptado de [8].

Os eventos aceitos no L1 são processados pelo HLT, com base em algoritmos computacionais que reduzem o número de eventos a serem gravados em disco, a uma taxa média de cerca de 1,5 kHz. A seleção de eventos candidatos pelo HLT é realizada a cada etapa, de forma que se houver reprovação de um determinado evento em uma das etapas de *trigger*, as etapas subsequentes não serão executadas para este evento.

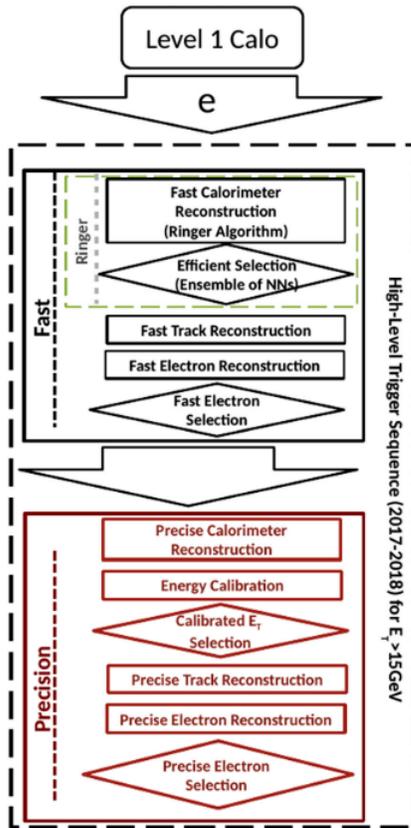


Figura 3: Diagrama simplificado do sistema de *trigger* empregado no ATLAS para elétrons na Run2 [13].

O HLT é subdividido nas etapas rápida e precisa. Na etapa rápida as RoIs selecionadas pelo L1 são analisadas. O algoritmo *NeuralRinger* usado como classificador na etapa rápida, realiza um mapeamento topológico em anéis concêntricos em torno da RoI (ver Fig. 4). O *NeuralRinger*, explora a geometria cônica do formato do chuveiro de partículas utilizando as células do calorímetro, para descrever a quantidade de energia depositada em cada anel concêntrico [19]. Os sinais em anéis alimentam um *ensemble* de redes neurais artificiais, para classificação do evento em sinal/ruído de fundo. O *NeuralRinger* permite a compactação da informação em 100 variáveis discriminantes (ver Tabela I), considerando uma quantidade de ≈ 1000 sensores presentes na RoI.

Camadas	PS	EM1	EM2	EM3	HAD0	HAD1	HAD2
Anéis	8	64	8	8	4	4	4

Tabela I: Número de anéis concêntricos gerados em cada camada do calorímetro.

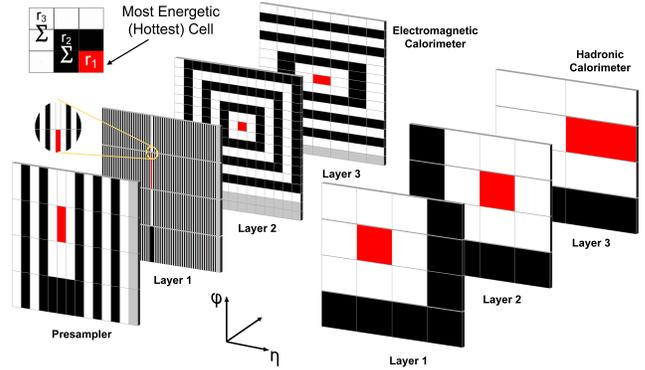


Figura 4: Representação dos anéis concêntricos gerados nas camadas do calorímetro do ATLAS pelo algoritmo anelador [13].

No calorímetro EM, os anéis são centralizados em torno da célula mais energética em cada camada, enquanto no calorímetro hadrônico a posição da célula mais energética na segunda camada do calorímetro EM é usada como eixo [20]. O *NeuralRinger* apresentou redução do falso alarme em até 3 vezes, na detecção de elétrons com energia acima de 15 GeV em relação ao algoritmo de seleção de eventos baseados em cortes rígidos em variáveis que descrevem o chuveiro de partículas (*CutBased*), que operou no HLT até 2017.

A. Trabalhos Relacionados

A proposta do mapeamento do perfil de deposição de energia em anéis concêntricos como sinais de entrada em redes neurais foi detalhada no trabalho [18]. A descrição do perfil de deposição de energia em anéis, alimentando uma rede neural completamente conectada encontrou 97% de eficiência para classificação de elétrons e 10,5% de erro de classificação de jatos hadrônicos em eventos sem *pile-up*. Também em [18], a proposta do uso da largura do chuveiro como variável

discriminante, para discriminação elétron/jato foi avaliada. A largura do chuvaire (variável R_p) unidimensional [21], pode ser definida conforme Eq. 1:

$$R_p = \frac{\sum_{i=1}^C E_i^\alpha r_i^\beta}{\sum_{i=1}^C E_i^\alpha}, \quad (1)$$

onde, E_i representa a energia depositada em uma célula individual i , r_i é a distância euclidiana entre o centro desta célula e a célula de energia máxima depositada, α e β são parâmetros que pode ser ajustados e C o número total de células. Usando a variável R_p unidimensional como discriminante com corte linear, foi encontrado 97% de eficiência de classificação de elétrons, e 9% de jatos incorretamente classificados como elétrons. Estes resultados foram encontrados usando uma busca por otimização dos parâmetros α e β , para minimizar a norma do erro de classificação de jatos em taxas específicas de eficiência de detecção. O mapeamento R_p multidimensional foi apresentado, usando a energia dos anéis topológicos ($RpRing$), conforme descrito na Eq. 2:

$$R_p : X \rightarrow Y, \\ (E_{ri}, rd_i) = \frac{E_{ri}^\alpha rd_i^\beta}{\sum_{i=1}^R E_i^\alpha}, \quad (2)$$

em que, rd_i representa o raio de cada anel medido em relação ao primeiro anel. Este método ofereceu a vantagem de combinar a compactação que a estrutura em anéis permite com o mapeamento discriminante R_p . Ainda em [18], uma rede neural foi adaptada no intuito de ser capaz de otimizar, durante o treinamento, os valores de α e β aplicando retropropagação do erro até uma camada inicial, cujos neurônios foram responsáveis pelo mapeamento R_p . Utilizou-se como entrada os valores de energia e distância euclidiana de cada uma das células do espaço de entrada.

Neste trabalho é apresentada uma extensão da investigação realizada em [18]. Um sistema integrado foi desenvolvido usando o mapeamento R_p multidimensional e um classificador neural. Os parâmetros α e β foram incluídos em uma camada de entrada e integrados ao procedimento de treinamento do classificador. Utilizou-se como entrada os valores de energia dos anéis gerados pelo algoritmo anelador. Deste modo, os parâmetros que ponderam a energia do anel e sua distância em relação ao primeiro anel foram ajustados durante o treinamento dos classificadores. A seção IV detalha a metodologia proposta.

IV. MÉTODO PROPOSTO

O fluxo de processamento proposto para uso do mapeamento R_p nos anéis é indicado na Fig. 5. A energia dos anéis gerados nas células do calorímetro é preprocessada, conforme descrito na Eq. 2. Os parâmetros α e β , são ajustados iterativamente ao longo do treinamento da rede neural, usando *backpropagation*, conforme indicado na Fig. 6.

Para realizar o treinamento do sistema de classificação proposto, foram utilizados conjuntos de dados experimentais, coletados em colisões do LHC em 2017. O conjunto de dados

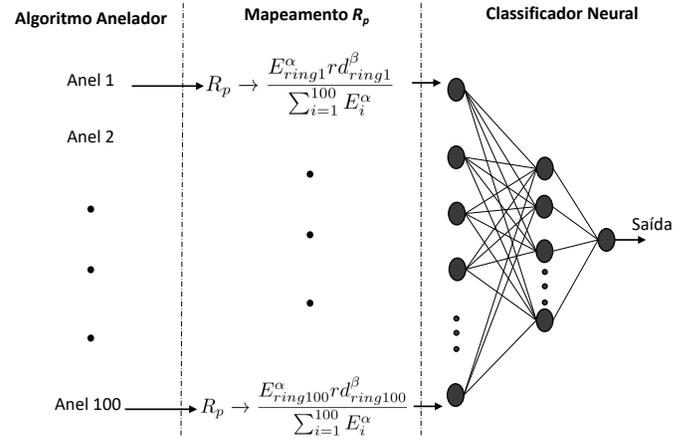


Figura 5: Fluxo de processamento do sistema de classificação desenvolvido. Os anéis concêntricos gerados pelo algoritmo anelador são processados pelo mapeamento R_p e alimentam um classificador neural para discriminação elétron/jato.

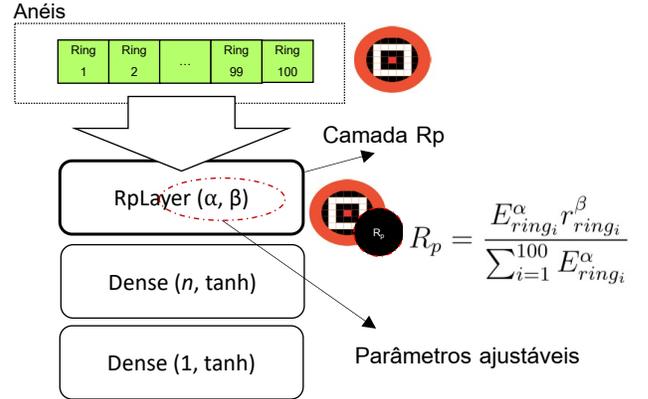


Figura 6: Diagrama da arquitetura da rede neural. Uma camada de entrada com parâmetros α e β ajustáveis durante o treinamento é usada para realização do mapeamento R_p multidimensional.

reservado para sinal é composto de assinaturas provenientes do decaimento do bóson Z ($Z \rightarrow e^+e^-$). Como os dados são de natureza experimental, critérios de decisão dos algoritmos de hipótese do ambiente *offline* são adotados para etiquetar assinaturas de sinal e ruído de fundo (usando modelos de física teórica).

Os conjuntos de dados selecionados para a classe de elétrons foram aqueles sinais aprovados como elétrons pelo método *Tag and Probe* (procedimento que utiliza aspectos de física teórica e medições em calorimetria para associar o par de elétrons que decaem do bóson Z) [22] usando o critério *medium*. O critério *medium* é um ponto de operação do algoritmo de classificação do *offline*, que permite desempenho equilibrado entre detecção de sinal (P_D) e contaminação por ruído de fundo (P_F). O conjunto de dados de ruído de fundo foi composto por eventos

reprovados pelos algoritmos de reconstrução do *offline* usando o critério *vloose*. O critério *vloose* é um ponto de operação dos algoritmos de classificação do *offline* que permite elevada coleta de elétrons porém aumento na aceitação de ruído de fundo.

Os conjuntos de dados são separados por energia transversa (E_T) e faixa em η do sistema de calorimetria. Para cada faixa, descrita na Tab. II, classificadores são treinados, formando um *ensemble* de 25 redes neurais. As Fig. 7 e 8, apresentam a distribuição em E_T e η respectivamente de todos os anéis presentes no conjunto de dados utilizado.

Tabela II: Regiões em η e faixas de energia empregadas para o ajuste dos modelos neurais.

E_T (GeV)	Intervalos				
	[15, 20]	[20, 30]	[30, 40]	[40, 50]	[50, ∞ [
η	[0;0,80]	[0,80;1,37]	[1,37;1,54]	[1,54;2,37]	[2,37;2,50]

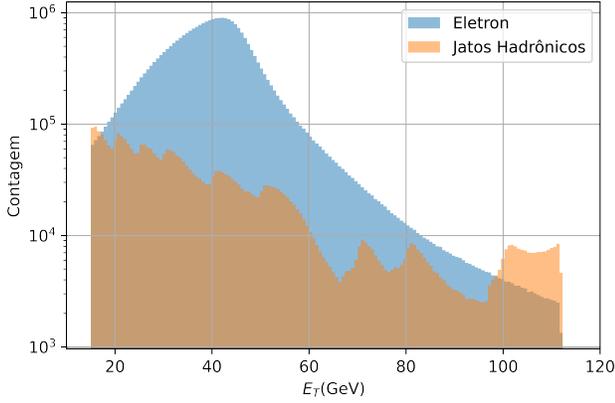


Figura 7: Distribuição em E_T dos conjuntos de dados utilizados para treinamento do *ensemble* de redes neurais.

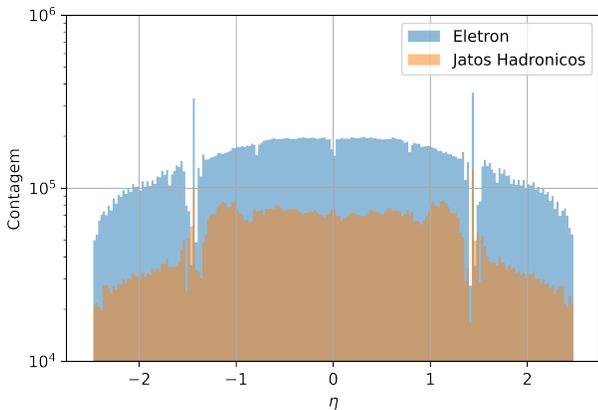


Figura 8: Distribuição pela coordenada η do ATLAS dos conjuntos de dados utilizados para treinamento do *ensemble*.

A quantidade de neurônios na camada oculta foi escolhida, variando o número de neurônios entre 2 e 10. Para análise da flutuação estatística dos conjuntos de dados e seu impacto no treinamento dos modelos, adotou-se um método de validação cruzada (*Stratified k-fold*) que realiza 10 sorteios, para separação dos conjuntos para as fases de treinamento e teste com as proporções de 90% e 10%, respectivamente. Para cada sorteio realizado, 10 inicializações de treinamento são efetuadas para analisar a possível ocorrência de mínimos locais. Dois critérios de parada do treinamento foram adotados. O treinamento dos modelos é interrompido quando 25 sucessivas falhas de melhoria do índice soma-produto (SP) (conforme Eq. 3) ocorre, ou quando o número máximo de 5000 épocas de treinamento é alcançado. O índice SP é um indicador de equilíbrio entre a probabilidade de detecção de sinal (P_D) e a taxa de falso alarme (P_F) [23] [24].

$$SP = \sqrt{\sqrt{P_D(1 - P_F)} \times \frac{P_D + (1 - P_F)}{2}}, \quad (3)$$

em que P_D , é a medida percentual de eventos da classe de sinal, classificados como sinal. A probabilidade de falso alarme (P_F) é a medida percentual de eventos de ruído de fundo classificados como sinal.

Após o treinamento, os classificadores com maior índice SP são selecionados para realização do ajuste do limiar de decisão, em função do número médio de interações por cruzamento de feixes ($\langle\mu\rangle$). Este procedimento visa verificar a influência do *pile-up* na taxa de falso alarme, para um mesmo ponto de operação de probabilidade de detecção. Também tem o propósito de manter a eficiência de detecção (P_D) de sinais constante, independente da taxa de empilhamento. Manter a taxa de P_D constante mesmo com as variações da taxa de empilhamento é muito importante para não deteriorar a eficiência geral do HLT (saída da etapa precisa). Para esta correção, a função de ativação tangente hiperbólica do neurônio de saída é retirada, e substituída por uma função linear, em que os coeficientes a e b (ver Eq. 4) são ajustados para manter a probabilidade de detecção semelhante a referência (taxa de P_D do *NeuralRinger*) em função de $\langle\mu\rangle$.

$$\delta(\langle\mu\rangle) = a \langle\mu\rangle + b, \quad (4)$$

em que a e b são os coeficientes da curva.

Após o treinamento, o desempenho obtido (usando os subconjuntos sorteados para teste) pelo método RpRing foi comparado com o *NeuralRinger* e com o *CutBased*². Foram destacados na seção V resultados obtidos na faixa de energia [30;40] GeV e η [0;0,8]. O sumário das principais configurações do treinamento dos modelos é apresentado na Tabela III.

V. RESULTADOS

Inicialmente avaliou-se a quantidade de neurônios na camada oculta, e o seu impacto nas medidas de avaliação de

²Método de classificação baseado em cortes lineares em variáveis de formato do chuveiro de partículas, que operou no HLT do ATLAS até 2017.

Tabela III: Resumo dos principais parâmetros utilizados no treinamento dos classificadores.

Algoritmo de treinamento	ADAM
Taxa de Aprendizado	0.010
Alvo de Elétron/Ruído de Fundo	+1/-1
Critério de Parada	25 sucessivas falhas de melhorias em SP
Número de Épocas	max. 5000
Número de Subconjuntos	10 (9 para treino, 1 para teste)
Goal Fitness Train	0.0001
Épocas	5000
Número de Camadas Ocultas	1
Função de ativação	Tangente Hiperbólica
Arquitetura	Completamente conectada

desempenho. A Fig.9 apresenta o *boxplot* dos índices SP encontrados (na faixa de energia E_T : [30, 40] GeV e η : [0, 0,80]), para arquiteturas avaliadas entre 2 e 10 neurônios na camada oculta. Observa-se que a partir de 7 neurônios, não há significativa mudança na distribuição dos valores de SP encontrados.

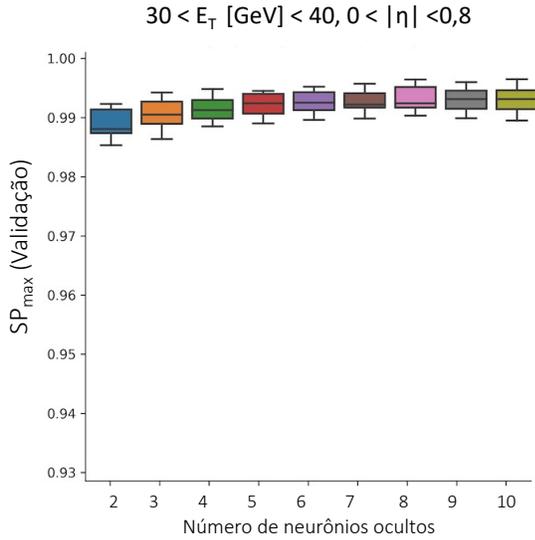


Figura 9: Boxplot dos valores encontrados para o índice SP com o conjunto de teste, para os modelos treinados com 2 a 10 neurônios na camada oculta.

As Tabelas IV e V apresentam os índices de probabilidade de detecção (P_D) e probabilidade de falso alarme (P_F) para os pontos de operação *tight* (o critério *tight* é o ponto de operação do classificador que indica que aquela partícula possui uma alta probabilidade de ser um elétron) e *vloose* dos sistemas de classificação respectivamente. Estas tabelas indicam os resultados de eficiência geral (considerando os resultados de todos os classificadores do *ensemble*). Os resultados obtidos indicam que o método RpRing com 7 neurônios proporciona uma redução na taxa de falso alarme de 0.54 pontos percentuais quando comparado a referência (NeuralRinger), operando no ponto de operação *tight*. Quando o ponto de operação é

ajustado para *vloose*, o RpRing encontra uma redução no falso alarme em 1.37 pontos percentuais, se comparado a referência.

Tabela IV: Eficiência geral usando validação cruzada para cada técnica de classificação para a operação *Tight*.

	P_D [%]	P_F [%]
CutBased	99.05	17.63
NeuralRinger (Referência)	99.05 ± 0.00	1.33±0.05
RpRing (mlp2)	99.04±0.00	1.00±0.01
RpRing (mlp3)	99.05±0.01	0.94±0.01
RpRing (mlp4)	99.05±0.00	0.90±0.01
RpRing (mlp5)	99.04±0.00	0.86±0.01
RpRing (mlp6)	99.05±0.00	0.84±0.01
RpRing (mlp7)	99.05±0.01	0.79±0.01
RpRing (mlp8)	99.05±0.01	0.79±0.01
RpRing (mlp9)	99.05±0.00	0.80±0.01
RpRing (mlp10)	99.05±0.01	0.80±0.01

Tabela V: Eficiência geral usando validação cruzada para cada técnica de classificação *Very Loose*.

	P_D [%]	P_F [%]
CutBased	99.65	32.67
NeuralRinger (Referência)	99.67 ± 0.00	2.86±0.17
RpRing (mlp2)	99.66±0.00	2.26±0.44
RpRing (mlp3)	99.67±0.01	2.73±1.70
RpRing (mlp4)	99.67±0.00	1.68±0.02
RpRing (mlp5)	99.67±0.00	1.59±0.04
RpRing (mlp6)	99.67±0.00	1.55±0.03
RpRing (mlp7)	99.67±0.01	1.49±0.01
RpRing (mlp8)	99.67±0.01	1.49±0.01
RpRing (mlp9)	99.67±0.00	1.51±0.01
RpRing (mlp10)	99.66±0.01	1.52±0.01

Durante o treinamento das redes neurais uma monitoração foi realizada para verificar possível incidência de *overfitting*. A Fig. 10 apresenta como exemplo, duas curvas de treinamento para o modelo neural que obteve melhor desempenho em SP na faixa de energia [30;40] GeV e η [0; 0,80]. Em 10 (a) é apresentado a evolução do erro médio quadrático ao longo das épocas de treinamento indicando que, neste caso, o erro sobre o conjunto de dados de validação não apresenta redução para um número de épocas acima de ≈ 175 . Em 10 (b) o comportamento do índice SP ao longo das épocas de treinamento é apresentado, em que por meio da estratégia de parada antecipada, o treinamento foi interrompido após 25 épocas sem melhorias em relação ao índice SP.

A evolução dos parâmetros α e β para o classificador neural selecionado que obteve melhor desempenho em SP na faixa de energia [30;40] GeV e η [0; 0,80] é apresentada na Fig. 11.

As Fig. 12 e 13 apresentam os histogramas de saída da rede neural selecionada, que obteve melhor desempenho em SP na faixa de energia [30;40] GeV e η [0; 0,80] para sinal e ruído de fundo respectivamente. A função de ativação não linear foi removida do neurônio de saída e uma curva linear foi ajustada em função do nível de empilhamento ($\langle\mu\rangle$). A curva foi ajustada para manter o P_D do classificador RingerRp com o mesmo desempenho do classificador de referência (NeuralRinger). Cada ponto da curva (pontos azuis) indica um

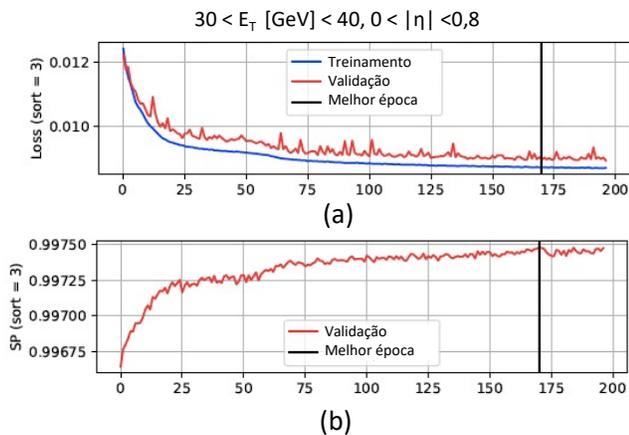


Figura 10: Curvas de monitoramento de treinamento para o classificador neural treinado no intervalo de E_T [30;40] GeV e η [0; 0,80]. Em (a) a curva do erro médio quadrático é apresentada e em (b) a curva do índice SP para cada época de treinamento é mostrada.

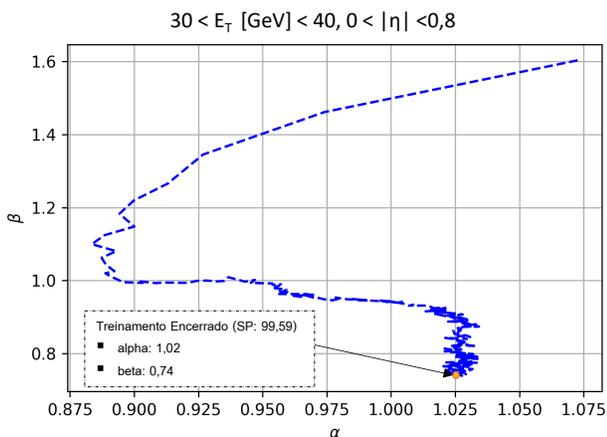


Figura 11: Variação dos parâmetros α e β ao longo das épocas de treinamento.

limiar de decisão que mantém o P_D de 99.18% considerando a variação do empilhamento. Na Fig. 12 observa-se que a maior quantidade de sinais está posicionada a direita da reta que foi ajustada pela intensidade de empilhamento, indicando que a maioria dos sinais foram corretamente discriminados pelo classificador.

Observa-se na Fig. 13 que a maioria das assinaturas de jatos hadrônicos estão a esquerda da reta ajustada em função do empilhamento, indicando que a maior parte destes eventos foram corretamente discriminados pelo RpRing.

VI. CONCLUSÕES

O experimento ATLAS é o maior detector de partículas de propósito geral do LHC, contribuindo significativamente para diversas descobertas científicas. Devido ao intenso ruído de fundo produzido nas colisões e uma elevada quantidade de

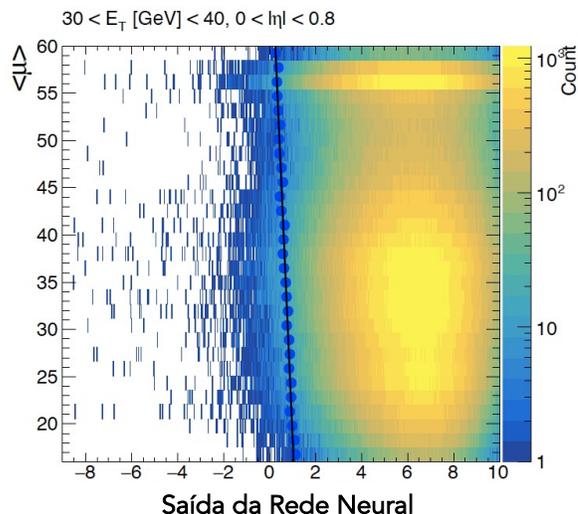


Figura 12: Histograma de saída da rede neural, com exemplos de sinal (elétrons), selecionada com melhor desempenho em SP na faixa de energia [30;40] GeV e η [0; 0,80].

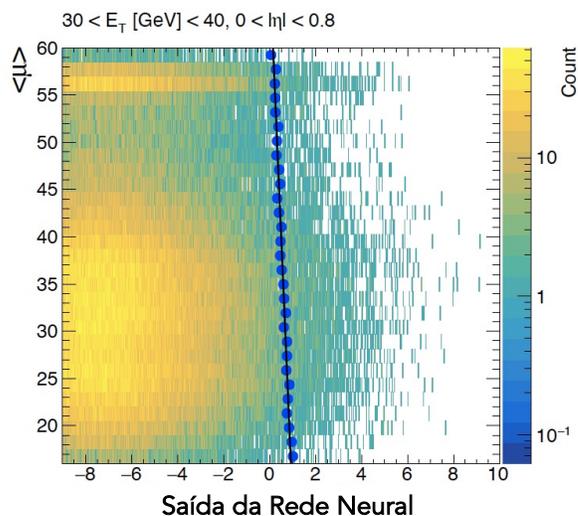


Figura 13: Histograma de saída da rede neural, com exemplos de ruído de fundo (jatos), selecionada com melhor desempenho em SP na faixa de energia [30;40] GeV e η [0; 0,80].

canais de leitura, um sistema rápido e eficiente de filtragem *online* é necessário. O algoritmo NeuralRinger, baseado em um *ensemble* de redes neurais artificiais, opera no sistema de filtragem *online* do ATLAS desde 2017, e a investigação de sua otimização vem sendo realizada, para atendimento das exigências do novo período de tomada de dados. Neste trabalho avaliou-se o uso da informação especialista do perfil lateral do chuva de partículas (R_p multidimensional), junto aos anéis gerados pelo NeuralRinger. Os resultados

da inclusão desta característica nos sinais formatados em anéis indicaram uma redução na taxa de falso alarme de 0,54 e 1,37 pontos percentuais para os pontos de operação *Tight* e *Very Loose* respectivamente (ambas configurações com 7 neurônios na camada oculta), se comparado ao Neural-Ringer usando dados experimentais de 2017. Os resultados encontrados motivam a exploração da metodologia proposta e para trabalhos futuros, dados de colisão de 2018 podem ser utilizados para verificar o impacto da técnica proposta. Redes com arquiteturas convolucionais também podem ser avaliadas.

AGRADECIMENTOS

Os autores agradecem a FAPESB, CNPq, FAPERJ e RENAFAE pelo apoio financeiro. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

- [1] J. Fernández-Tejero, V. Fadeyev, C. Fleta, J. Hacker, M. Ullán, and Y. Unno, "Microelectronic test structures for the development of a strip sensor technology for high energy physics experiments," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 969, p. 163971, 2020.
- [2] T. Boccali, "Computing models in high energy physics," *Reviews in Physics*, vol. 4, p. 100034, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2405428319300449>
- [3] N. Nottbeck, D. C. Schmitt, and P. D. V. Büscher, "Implementation of high-performance, sub-microsecond deep neural networks on FPGAs for trigger applications," *Journal of Instrumentation*, vol. 14, no. 09, pp. P09014–P09014, sep 2019.
- [4] D. Sarkar, M. P., P. S., N. Chouhan, C. Borwankar, A. Bhattacharya, A. Tickoo, and R. Rannot, "Comparison of data storage and analysis throughput in the light of high energy physics experiment mace," *Astronomy and Computing*, vol. 33, p. 100409, 2020.
- [5] S. Benson and K. Gizdov, "Nndrone: A toolkit for the mass application of machine learning in High Energy Physics," *Computer Physics Communications*, vol. 240, pp. 15–20, 2019.
- [6] 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.
- [7] P. Arpaia, G. Azzopardi, F. Blanc, G. Bregliozzi, X. Buffat, L. Coyle, E. Fol, F. Giordano, M. Giovannozzi, T. Pieloni, R. Prevet, S. Redaelli, B. Salvachua, B. Salvant, M. Schenk, M. S. Camillocci, R. Tomás, G. Valentino, F. Van der Veken, and J. Wenninger, "Machine learning for beam dynamics studies at the cern large hadron collider," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 985, p. 164652, 2021.
- [8] A. Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, pp. S08003–S08003, aug 2008. [Online]. Available: <https://doi.org/10.1088/1748-0221/3/08/s08003>
- [9] L. Evans and P. Bryant, "LHC machine," *Journal of Instrumentation*, vol. 3, no. 08, pp. S08001–S08001, aug 2008.
- [10] O. Brüning, H. Burkhardt, and S. Myers, "The Large Hadron Collider," *Progress in Particle and Nuclear Physics*, vol. 67, no. 3, pp. 705–734, 2012.
- [11] R. Wigmans, *Calorimetry: Energy Measurement in Particle Physics*, ser. International series of monographs on physics. Clarendon Press, 2017.
- [12] A. Collaboration, "Evidence for Higgs boson decays to a low-mass dilepton system and a photon in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Physics Letters B*, vol. 819, p. 136412, 2021.
- [13] G. Aad, Abbott *et al.*, "Performance of electron and photon triggers in ATLAS during lhc run 2," *Eur. Phys. J. C*, vol. 80, p. 47. 56 p, Sep 2019.
- [14] "An Ensemble of Neural Networks for Online Filtering Implemented in the ATLAS Trigger System," CERN, Geneva, Tech. Rep., Jul 2018. [Online]. Available: <https://cds.cern.ch/record/2632911>
- [15] J. V. Da Fonseca Pinto, "Ring-shaped Calorimetry Information for a Neural EGamma Identification with ATLAS Detector," CERN, Geneva, Tech. Rep., Mar 2016. [Online]. Available: <https://cds.cern.ch/record/2142761>
- [16] A. Collaboration, "Projected ATLAS electron and photon trigger performance in run 3," *Proceedings of Science*, vol. ICHEP2020, p. 757, 2021.
- [17] D. Acosta, S. Buontempo, L. Calóba, R. DeSalvo, A. Ereditato, R. Ferrari, G. Fumagalli, G. Goggi, W. Hao, A. Henriques, L. Linssen, M. Liyan, A. Maio, M. Mondardini, B. Ong, H. Paar, F. Pastore, E. Pennacchio, L. Poggioli, G. Polesello, F. Riccardi, A. Rimoldi, C. Scheel, J. Seixas, A. Simon, M. Sivertz, P. Sonderegger, M. Souza, Z. Thomé, V. Vercesi, Y. Wang, R. Wigmans, and C. Xu, "Lateral shower profiles in a lead/scintillating fiber calorimeter," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 316, no. 2, pp. 184–201, 1992.
- [18] 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, pp. 143–157, 1996.
- [19] T. Ciodaro, D. Deva, J. de Seixas, and D. Damazio, "Online Particle Detection by Neural Networks Based on Topologic Calorimetry Information," *14th International Workshop On Advanced Computing And Analysis Techniques In Physics Research, Uxbridge, West London, UK, 5 - 9 Sep 2011*, Sep 2011. [Online]. Available: <http://cds.cern.ch/record/1379508>
- [20] W. Spolidoro Freund and A. Collaboration, "An Ensemble of Neural Networks for Online Electron Filtering at the ATLAS Experiment," CERN, Geneva, Tech. Rep., May 2019. [Online]. Available: <https://cds.cern.ch/record/2675025>
- [21] C. V. Scheel, *The spaghetti calorimeter. Research, development, application*. Amsterdam (Netherlands): International Atomic Energy Agency, 1994.
- [22] A. Collaboration, "Electron efficiency measurements with the ATLAS detector using 2012 lhc proton–proton collision data," *Eur. Phys. J. C*, vol. 195, no. 77, pp. 1–45, 2017.
- [23] R. C. Torres, D. E. F. de Lima, E. F. de Simas Filho, and J. M. de Seixas, "Neural online filtering based on preprocessed calorimeter data," in *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2009, pp. 530–536.
- [24] E. F. Simas Filho, J. M. de Seixas, and L. P. Calóba, "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>