

Redes Neurais Convolucionais para Mitigação de Ruído Cruzado em um Sistema de Calorimetria de Altas Energias com Elevada Segmentação

André P. Conrado Rodrigues¹, Antônio C. L. Fernandes Jr¹,
Edmar Egídio P. de Souza¹, Eduardo F. de Simas Filho¹, Paulo César M. de A. Farias¹,
Mateus Hufnagel M. de Faria², Luciano Manhães A. Filho², Bertrand Laforge³

¹Laboratório de Sistemas Digitais, PPGEEC/UFBA

²Universidade Federal de Juiz de Fora

³LPNHE - Sorbonne University, Paris

Emails: {andre.paiva, antonio.lopes, edmar.egidio, eduardo.simas, paulo.farias}@ufba.br,
mateus.hufnagel@cern.ch, luciano.andrade@ufjf.br, laforge@lpnhe.in2p3.fr

Resumo—Sistemas de instrumentação eletrônica com fina segmentação de sensores podem ser afetados por interferência cruzada (*crosstalk*), ocasionada principalmente pelos efeitos de acoplamento indutivo e capacitivo entre os sensores, que são muito próximos entre si. Este cenário pode ser observado em experimentos de física de partículas dotados de sistemas de calorimetria, utilizados para medição da energia das partículas. Nestes sistemas, o *crosstalk* degrada a estimação da amplitude e da fase do sinal coletado, e por consequência, perturba a medição precisa de propriedades subatômicas, impactando na qualidade da reconstrução do sinal de interesse. Com base na disposição matricial de células de um calorímetro de argônio líquido, este trabalho adota uma metodologia baseada em aprendizagem de máquina para redução de interferência cruzada, utilizando arquiteturas de Redes Neurais Convolucionais. A metodologia proposta para estimação de energia para a célula central do arranjo trouxe resultados satisfatórios, obtendo coeficiente de correlação de até 99% entre a energia da partícula estimada pelo modelo neural e a energia sem efeito de *crosstalk*.

Palavras-chave—Redes Neurais Convolucionais, Calorimetria, Detecção de Partículas, Física de Altas Energias, *Crosstalk*.

I. INTRODUÇÃO

Experimentos de Física de Partículas impulsionam a fronteira tecnológica, devido à necessidade de elevada precisão na medição de propriedades de interações a níveis nucleares e subatômicos [1]. Infraestruturas tecnológicas complexas são implementadas para realizar medições precisas, desde a instrumentação eletrônica dos sensores, que coletam informações da interação das partículas com a matéria, até *hardware* dedicado para processamento da informação e sistemas computacionais para avaliação e refinamento das medidas [2], [3]. Em experimentos que coletam informações da subestrutura de partículas por meio de colisões, detectores de partículas são acomodados em torno do ponto de colisão dos feixes, com o intuito de medir as propriedades do subproduto das partículas que interagem com o sistema de detecção [4].

No acelerador de partículas LHC (*Large Hadron Collider*), atualmente em operação no CERN (Centro Europeu para Pesquisa Nuclear), diversos experimentos são acomodados

em diferentes pontos de colisão; um deles é o experimento ATLAS (*A Toroidal LHC Apparatus*), que engloba vários subdetectores com princípios de construção particulares e diferentes finalidades. Considerando apenas o calorímetro eletromagnético, há mais de 100.000 canais de leitura finamente segmentados que capturam o espalhamento lateral e longitudinal de partículas incidentes [5]. Experimentos desta natureza viabilizam a medição precisa de propriedades subatômicas como energia e trajetória da partícula. Tais informações são de vital importância para a reconstrução de eventos físicos gerados com as colisões [2]. Por meio da observação destes eventos, pesquisadores buscam novas descobertas que podem elucidar algumas das muitas questões em aberto relacionadas ao Modelo Padrão e à natureza fundamental da matéria, bem como validar modelos teóricos experimentalmente [6].

Dentre os subdetectores empregados, há o calorímetro eletromagnético de argônio líquido, que possui células dispostas em camadas com subdivisões de diferentes granularidades, conciliando a separação de partículas incidentes fisicamente próximas com uma boa absorção de energia. Este tipo de calorímetro é projetado para interagir de modo mais intenso com partículas que interagem com a força eletromagnética, ou seja, elétrons (e^-), pósitrons (e^+) e fótons (γ) [4].

Um dos desafios associados à detecção de partículas por meio dos calorímetros supracitados é a presença de *crosstalk* (XT), que é a interferência cruzada nos sinais gerados pelas células de detecção por conta de acoplamentos capacitivos, indutivos e resistivos entre células vizinhas, os quais são inerentes à construção eletrônica e mecânica dos dispositivos e à condição de interação com níveis elevados de energia [7], [8]. O *crosstalk* deforma o sinal observado, causando impactos indesejados no processo de reconstrução de eventos [9], e a modelagem dos efeitos de *crosstalk* não é trivial, justificando a busca por alternativas orientadas a dados.

Trabalhos anteriores averiguam o uso de técnicas de aprendizado de máquina e processamento digital de sinais para diversas aplicações na área de física de altas energias, como a

filtragem *online* de eventos para gravação e análise [10]–[12], estimação de energia a partir de leituras de calorímetros [13], simulação de colisões e respostas de detectores [14] e atenuação de efeitos de interferência cruzada [7], [8].

Este artigo traz resultados de experimentos preliminares de aplicação de Redes Neurais Convolucionais (CNNs) para mitigação de *crosstalk* em eventos gerados por simulação com o *framework* Lorenzetti Showers [14]. A escolha pelas CNNs é devida à disposição física das células do calorímetro em formato matricial, de modo que a convolução poderia trazer vantagens na captura de relações de interferência cruzada entre células vizinhas. Neste trabalho, a colisão de um único elétron no centroide de uma célula foi considerada como um “evento”, e os dados utilizados foram *clusters* de células de tamanho 3×3 , com 5 amostras ao longo do tempo contendo o sinal contaminado por *crosstalk* indutivo e ruído eletrônico. A estimação de energia para a célula central do *cluster* foi satisfatória, mas para as células vizinhas o desempenho foi mediano. A estimação do tempo de voo da partícula foi insatisfatória, evidenciando a necessidade de avaliar novas estratégias de treinamento e/ou arquiteturas que levem a sequência temporal das amostras em conta.

O texto está organizado como descrito a seguir. A Seção II aborda sistemas de calorimetria para Física de Altas Energias, com enfoque em calorímetros eletromagnéticos. Na Seção III constam maiores detalhamentos sobre o *crosstalk*. A metodologia adotada neste trabalho é apresentada na Seção IV, com os resultados expostos e discutidos na Seção V. Por fim, a Seção VI apresenta as conclusões e perspectivas futuras.

II. SISTEMAS DE CALORIMETRIA ELETROMAGNÉTICA PARA FÍSICA DE ALTAS ENERGIAS

Detectores instalados em aceleradores de partículas tipicamente compreendem uma câmara em formato cilíndrico posicionada em um ponto estratégico de colisão. A câmara engloba múltiplos subdetectores em disposição praticamente simétrica [9]. Na literatura, é convencionado um sistema de coordenadas para o interior do detector, esquematizado na Figura 1. O ponto nominal de colisão é tomado como origem e reside no centro da câmara, e o feixe de colisão é representado pelo eixo z . O ângulo ϕ é utilizado para caracterizar a direção de jatos e partículas em relação à circunferência da câmara, e o ângulo θ diz respeito à direção em relação ao comprimento do detector, utiliza-se o conceito de “pseudorapidez” [5], definido matematicamente como $\eta = -\ln(\tan \frac{\theta}{2})$.

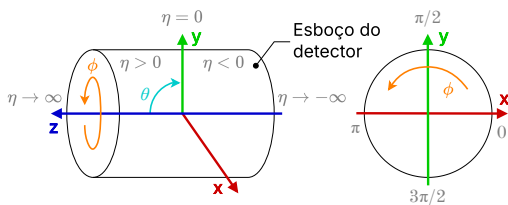


Figura 1. Sistema de coordenadas de referência para pseudorapidez e unidades angulares. Adaptado e revisado de [15].

No ponto de colisão, aglomerados de prótons colidem numa taxa da ordem de milhões de vezes por segundo, fenômeno este chamado de *bunch crossing* [8], [16]. As colisões geram novas partículas que, ao interagirem com os subdetectores, geram chuveis de partículas que se propagam para camadas mais externas da câmara. Uma representação computadorizada deste processo é exibida na Figura 2.

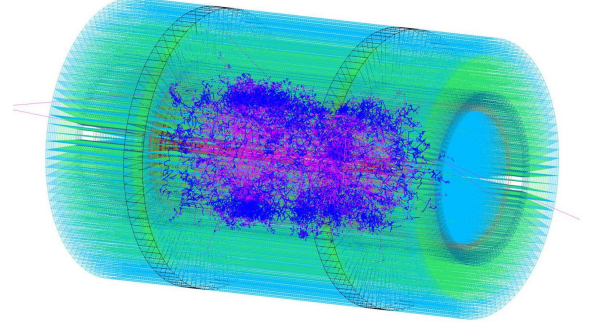


Figura 2. Formação de chuveis de partículas no interior do detector. Adaptado de [14].

Dentre os subdetectores empregados, encontram-se calorímetros. No ramo de física de partículas, a calorimetria compreende técnicas e dispositivos de instrumentação dedicados a identificar partículas originadas por um determinado evento, e caracterizar seus perfis de deposição de energia, ângulo/direção de incidência e tempos de voo [2], [4].

A Figura 3 mostra a disposição de calorímetros de propósito geral cuja modelagem é empregada para fins de simulação no Lorenzetti [14]. Neste diagrama, nota-se a presença de duas classes de calorímetros: os eletromagnéticos (EMCal), localizados em camadas internas do detector, e os hadrônicos (HCal) no trecho externo. Ambos cobrem completamente o detector ao redor de ϕ . O EMCal é disposto em duas zonas do detector para possibilitar uma maior cobertura em termos de η , compondo as regiões do barril (EMB) e *end-cap* (EMEC). O EMB absorve partículas em um intervalo de $|\eta| < 1,475$, enquanto o EMEC fica a cargo da região $1,375 < |\eta| < 3,2$.

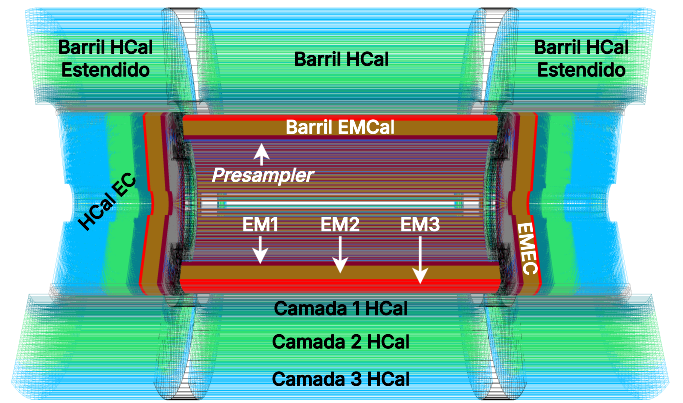


Figura 3. Recorte de um conjunto de calorímetros de propósito geral para detectores de partículas. Adaptado de [14].

O calorímetro eletromagnético é segmentado em camadas sobrepostas (EM1, EM2 e EM3), cada uma composta de células de leitura com sua granularidade particular dada em termos de $\Delta\eta \times \Delta\phi$, como ilustrado na Figura 4.

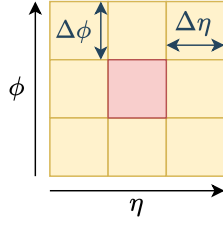


Figura 4. Disposição e granularidade das células de leitura do EMCal.

A camada EM1 consegue capturar uma posição mais precisa em relação a η ; já a deposição de energia captada se mostra mais intensa na EM2 para $e^\pm/\gamma$, sem deposição significativa de energia no HCal [17]. Deste modo, consegue-se uma informação mais acurada quanto à posição de incidência da partícula, ao mesmo tempo que se obtém o perfil de deposição característico da partícula incidente. A Tabela I exibe as dimensões das células do EMCal [14].

Tabela I
GRANULARIDADE DAS CÉLULAS DE LEITURA DO EMCal POR REGIÃO

Camada	Região	Cobertura	$\Delta\eta \times \Delta\phi$
EM1	Barril	$0,00 < \eta < 1,55$	$0,003 \times 0,1$
		$1,37 < \eta < 1,80$	$0,003 \times 0,1$
		$1,80 < \eta < 2,00$	$0,025 \times 0,1$
		$2,00 < \eta < 2,37$	$0,006 \times 0,1$
		$2,37 < \eta < 3,20$	$0,1 \times 0,1$
EM2	Barril	$0,00 < \eta < 1,50$	$0,025 \times 0,025$
		$1,35 < \eta < 2,50$	$0,025 \times 0,025$
		$2,50 < \eta < 3,20$	$0,1 \times 0,1$
EM3	Barril	$0,00 < \eta < 1,58$	$0,05 \times 0,1$
		$1,35 < \eta < 2,50$	$0,05 \times 0,025$
		$2,50 < \eta < 3,20$	$0,1 \times 0,1$

Os sinais provenientes das células de leitura são proporcionais à energia depositada; tais sinais são amplificados e passam por uma etapa de *shaping*, que consiste na adequação do sinal para um pulso característico. Este pulso é então amostrado numa taxa de 40 MHz, mesma taxa de ocorrência de *bunch crossing* [5], [18]. A Figura 5 exibe o sinal triangular com decaimento linear proveniente da leitura, o pulso após a adequação do sinal e as amostras coletadas.

III. O PROBLEMA DO CROSSTALK

Com a sequência de amostras do pulso gerado a partir da leitura, é feita a reconstrução da energia e do tempo de voo, como parte do processo de caracterização da partícula incidente. A reconstrução é executada por meio do uso do Filtro Ótimo (OF) [8], [19], que de modo simplificado, pode ser expressado pelas Equações 1 e 2,

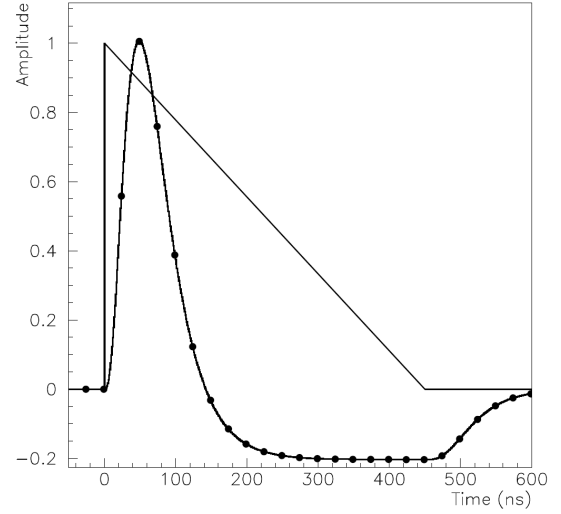


Figura 5. Sinal triangular capturado pelos eletrodos do EMCal, pulso gerado após *shaping* e amostras (pontos sobre o pulso). Fonte: [5].

$$A = \sum_{i=1}^n a_i S_i, \quad (1)$$

$$\tau = \frac{1}{A} \sum_{i=1}^n b_i S_i, \quad (2)$$

onde A representa a amplitude do sinal, τ é o tempo de voo da partícula, a_i e b_i são coeficientes de reconstrução respectivamente relacionados à amplitude e ao tempo de voo, e n é o número de amostras consecutivas utilizadas para reconstrução do sinal.

O problema é que o método de reconstrução não dispõe de estratégias para eliminar ou atenuar as interferências inerentes à própria construção do calorímetro, que causam distorções nas informações recuperadas [9]. As amostras envolvidas nos cálculos do OF já estão contaminadas pelo *crosstalk*, e podem ser representadas pela Equação 3,

$$S_j^{\text{rec}}(t_i) = S_j^{\text{truth}}(t_i) + \sum_k XT_{k \rightarrow j}(t_i) + b_j(t_i), \quad (3)$$

na qual i representa o índice da amostra, e j uma determinada célula de interesse com sua k -ésima célula vizinha. $S_j^{\text{rec}}(t_i)$ representa a amostra contabilizada na reconstrução, $S_j^{\text{truth}}(t_i)$ é a amostra da verdadeira energia depositada na célula, $XT_{k \rightarrow j}(t_i)$ é a influência do *crosstalk* causado pela célula vizinha k sobre a célula j , e $b_j(t_i)$ consiste num modelo de ruído térmico da célula j [8].

O *crosstalk* é causado pelo compartilhamento de cargas por entre os canais de leitura por conta de acoplamentos elétricos indesejados e intrínsecos da construção do dispositivo, de modo que tais acoplamentos ocasionam efeitos capacitivos, indutivos e resistivos [9]. Tipicamente, o efeito resistivo não é considerado nos estudos para mitigação de XT por não causar

impactos na forma de onda do sinal ou atrasos na temporização, que são os principais efeitos a serem atenuados [8]. A Figura 6 representa de modo simplificado os acoplamentos capacitivos e indutivos entre uma célula de interesse (destacada em vermelho) e seus vizinhos imediatos.

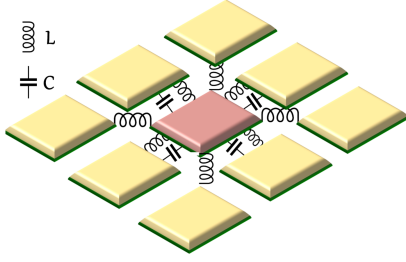


Figura 6. Representação de acoplamentos capacitivos e indutivos relacionados ao *crosstalk* entre células do EMCal. Fonte: [7].

A mensuração do impacto do XT sobre a recuperação do perfil de deposição de energia pode ser definido matematicamente de duas formas [8], [9], as quais tomam como base as indicações presentes na Figura 7:

- **Sob o pico:** Proporção da intensidade da contribuição do XT no momento do pico do pulso característico ($X(T_{Max})$) em relação à amplitude de pico ($WavePeak(T_{Max})$). Medida considerada mais realista em relação aos impactos no tempo de voo e na amplitude, mas que sofre variações com as flutuações do *crosstalk* e com deslocamentos de fase no processo de amostragem.
- **Pico-a-pico:** Consiste na razão entre as amplitudes de pico do XT (X_{max}) e do pulso característico ($WavePeak(T_{Max})$). Considerada uma medida mais pessimista, mas mais robusta aos fatores que causam instabilidade na estimação sob o pico.

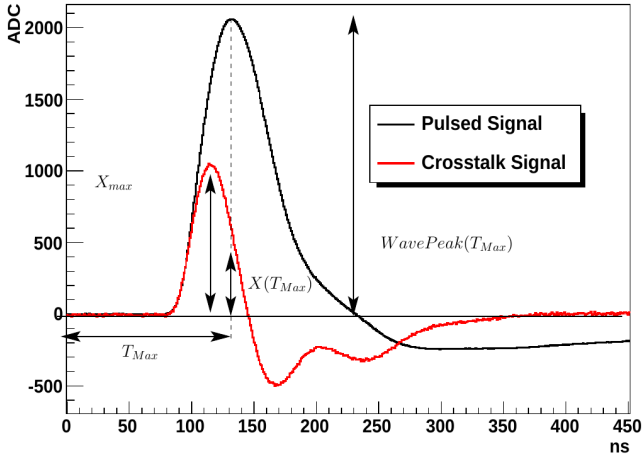


Figura 7. Representação gráfica das definições de *crosstalk*. Fonte: [20].

A Figura 8 exibe uma representação dos efeitos do *crosstalk* sobre o pulso característico. O sinal com pico V_{max} representa o pulso sem interferência, enquanto a curva tracejada com pico em V_X representa o pulso deformado por influência do XT.

A curva de menor amplitude com pico em X_{max} representa a influência do XT em si.

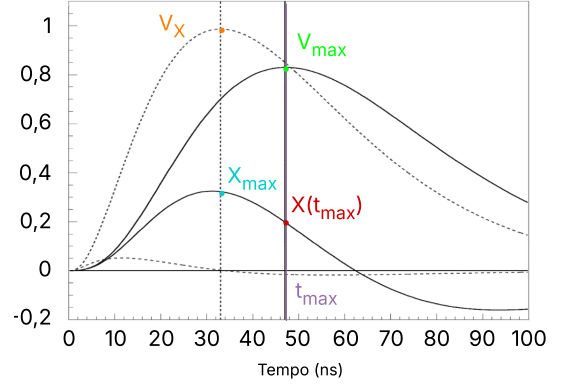


Figura 8. Exemplo gráfico de deformação causada no pulso pelo *crosstalk*. Adaptado de [21].

IV. METODOLOGIA

Os experimentos consistem no uso de eventos simulados para treinar redes neurais convolucionais em estratégia supervisionada, objetivando estudar o comportamento destas redes como potenciais mecanismos de redução de *crosstalk* em sinais oriundos de calorímetros eletromagnéticos.

A. Geração do conjunto de dados

Os conjuntos de amostras utilizados para treinamento e avaliação de desempenho foram obtidos a partir de eventos simulados por meio do *framework* Lorenzetti [14], com a disposição dos calorímetros de acordo com a Figura 3. Foram gerados vários conjuntos de dados com aproximadamente 10000 eventos cada para diferentes configurações de tamanho de *cluster* e de magnitude de contribuições (capacitiva e indutiva) no XT.

Os eventos simulados representam a incidência de um único e^- no EMCal com energia de 50 GeV, atingindo o centroide de uma célula da camada EM2. Foi considerado que a colisão responsável pela incidência do elétron ocorreu exatamente no centro do detector, no momento zero do *bunch crossing*, ou seja, sem flutuações na posição do vértice de colisão. A modelagem do ruído eletrônico também está incluída nas amostras geradas, e é representada por uma distribuição gaussiana de média zero e desvio padrão de 60 MeV.

Os dados do conjunto compreendem *clusters* de formato 3×3 com 5 amostras no tempo igualmente espaçadas por 25 ns, com $\eta \approx 0,7125$ e $\phi \approx 1,828$. A contribuição indutiva adotada para o *crosstalk* simulado foi de 1,1% do pico da energia verdadeira pela definição pico-a-pico; não foi considerada contribuição capacitiva.

B. Modelos, Treinamento e Avaliação de Desempenho

Os modelos CNN treinados e avaliados seguem o padrão arquitetural exibido na Figura 9, recebendo como entrada um conjunto de amostras de um *cluster* ao longo do tempo, e retornando como saída o conjunto de amostras filtrado. Foi

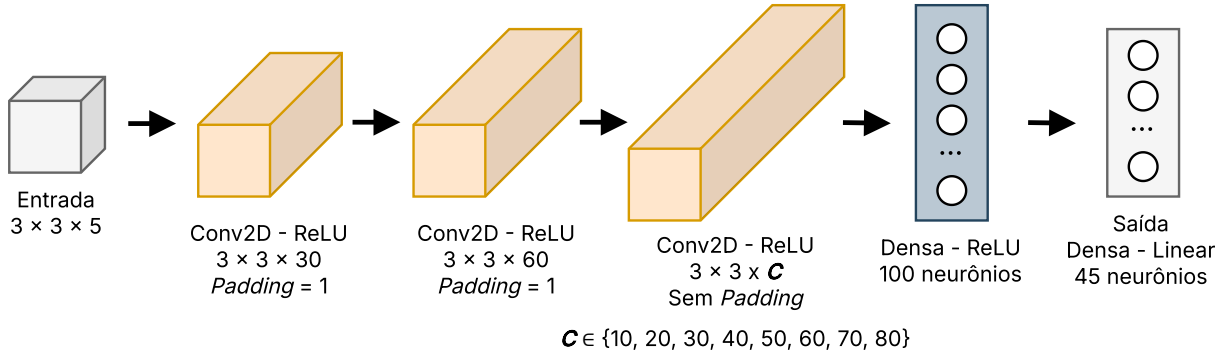


Figura 9. Diagrama da arquitetura dos modelos CNN treinados, com excursão do número de canais da terceira camada convolucional (parâmetro C).

utilizado um tamanho de *kernel* de 3×3 para as três camadas convolucionais utilizadas, e a quantidade de canais na terceira camada foi excursionada (parâmetro C). A arquitetura foi inspirada na disposição matricial das células do EMCaI, motivando o uso de camadas convolucionais. Camadas posteriores completamente conectadas foram usadas para tentar capturar relações implícitas entre as células para além da vizinhança imediata.

A representação de dados na entrada foi formatada de modo que a matriz 3×3 representa o *cluster* cuja célula central foi atingida em seu centroide por um elétron, e os 5 canais representam as amostras no tempo para cada uma das células. Após as camadas convolucionais, há uma camada densa oculta de 100 neurônios, e a camada de saída conta com 45 neurônios, representando as 5 amostras das 9 células em sequência temporal.

Convencionamos a indexação do vetor de saída da CNN com a disposição $9k + x$, com k representando o instante de tempo. Para representação espacial, utilizamos uma convenção inteira relativa de (η, ϕ) , assumindo a célula central como $(0, 0)$. As convenções estão ilustradas na Figura 10.

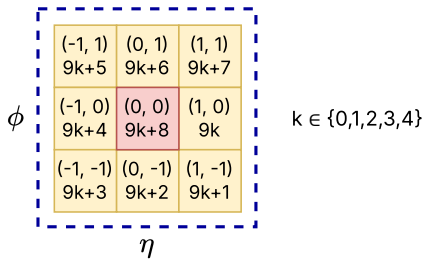


Figura 10. Convenção de indexação para células do *cluster*.

As configurações dos treinamentos podem ser visualizadas na Tabela II. A quantidade de eventos e o tamanho de *batch* foram escolhidos para otimizar o tempo de treinamento. A estratégia *K-Fold* foi utilizada com número de *folds* típico usado comumente na literatura. Valores padrão foram utilizados para os parâmetros β_1 e β_2 do otimizador Adam [22]. Os outros parâmetros foram obtidos de modo heurístico.

A função Erro Quadrático Médio (MSE) utilizada como custo é definida pela Equação 4,

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2, \quad (4)$$

na qual $\hat{y}_i$ é o valor predito, y_i é o valor de referência e n é a quantidade de exemplos levados em conta.

Tabela II
CONFIGURAÇÕES DOS TREINAMENTOS

Número de eventos	65536
Tamanho de <i>batch</i>	8192
Estratégia de validação	<i>K-Fold</i>
Número de <i>folds</i>	5
Iterações por <i>fold</i>	5
Função de custo	MSE
Otimizador	Adam
β_1	0,9
β_2	0,999
Taxa de aprendizagem	$2,5 \times 10^{-3}$
Decaimento da taxa de aprend.	10^{-5}
Número máx. de épocas	500
Critério de parada	25 épocas sem melhorias de custo

A avaliação de desempenho foi feita com base na reconstituição da energia e do tempo de voo por meio da aplicação do OF sobre a saída da rede. O conjunto de testes compreende 50000 eventos com as mesmas características do conjunto de treinamento. A qualidade da reconstrução de cada célula foi aferida por meio de gráficos de dispersão $\hat{y}_i \times y_i$ e do coeficiente de determinação (R^2), definido pela Equação 5,

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (5)$$

no qual $\bar{y}$ representa a média das observações. Valores de R^2 mais próximos de 1 indicam predições mais próximas do esperado. Para cada célula também foi avaliada a evolução da Raiz do Erro Quadrático Médio ($\text{RMSE} = \sqrt{\text{MSE}}$) para observar a influência do excursão de C .

V. RESULTADOS

A Figura 11 exibe *boxplots* com as distribuições de MSE para as melhores iterações de cada *fold*, considerando o excursionamento do número de canais da terceira camada convolucional. Analisando apenas esta Figura, é notável a vantagem de modelos com 30+ canais na terceira camada em relação aos modelos com 10 e 20 canais. No entanto, após este limite de 30 canais, o MSE no conjunto de validação mostra uma tendência de estagnação.

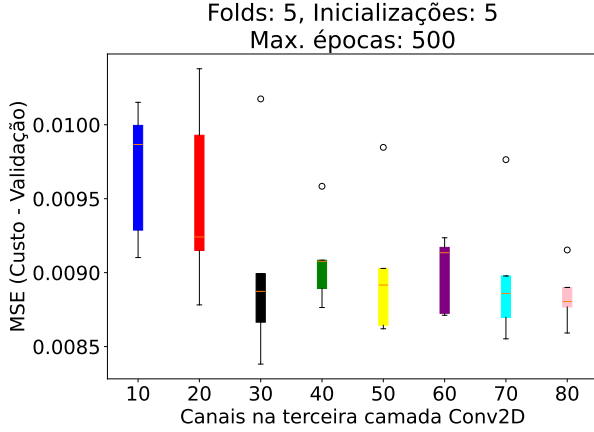


Figura 11. Distribuição de valores da função custo para melhor inicialização de cada *fold*.

O gráfico de dispersão presente na Figura 12 refere-se à variante do modelo CNN com 80 canais na terceira camada, e foi gerado a partir do conjunto de testes. Percebe-se que, apesar de um bom desempenho para reconstrução da energia na região de maiores níveis energéticos, a estimativa de tempo de voo fica restrita a uma faixa de valores que não compreende toda a real distribuição. Este fato se mostra mais evidente para as células com energia abaixo de 5 GeV.

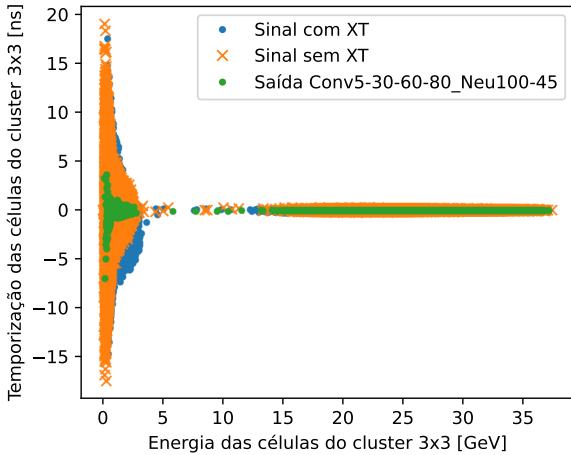


Figura 12. Gráfico de dispersão de energia e tempo dos sinais sem XT, com XT e pós-processados pela rede CNN com $C = 80$.

A recuperação de tempo de voo não apresentou melhorias significativas, com R^2 tendendo a 0, o que evidencia a necessidade de usar estratégias de treinamento que levem em conta a sequência temporal dos dados. Tipicamente em redes CNN bidimensionais, os canais não representam informação temporal, diferentemente da representação de dados adotada neste trabalho para a entrada das redes.

Em relação à recuperação de energia das células, a Tabela III exibe os valores de R^2 para a reconstrução de energia pelo OF a partir das amostras contaminadas e das amostras filtradas, de acordo com a convenção da Figura 10. Foi escolhida a variante do modelo com 60 canais na terceira camada, que apresentou os melhores valores de R^2 dentre os modelos. Os valores negativos de R^2 para o sinal contaminado realçam o impacto do processamento da CNN para a reconstrução. A melhoria foi substancial para a célula central, e significativa para as células vizinhas, mas o desempenho da filtragem destas ainda é passível de melhorias.

Tabela III
 R^2 PARA ESTIMAÇÃO DE ENERGIA PELO OF A PARTIR DE AMOSTRAS CONTAMINADAS E FILTRADAS COM CNN ($C = 60$)

R^2 (contaminado)		
-1,8088	-0,0300	-1,7429
-1,4971	0,5103	-1,0427
-1,8923	-0,2130	-1,8880

R^2 (filtrado)		
0,0612	0,5866	0,0615
0,0715	0,9917	0,4825
0,0136	0,5111	0,0805

A Figura 13 exibe gráficos de dispersão da qualidade de reconstrução da energia para cada uma das células do *cluster*. Desconsiderando a célula central, nota-se que as células (0, 1), (0, -1) e (1, 0) obtiveram coeficientes de determinação maiores que as outras células. A célula (-1, 0), no entanto, não se desempenhou de modo equiparável, representando um desempenho assimétrico para a reconstrução de energia.

Na Figura 14 vemos *boxplots* da RMSE para os níveis de energia de cada uma das células de um *cluster*, considerando o conjunto de dados de teste e as arquiteturas com excursionamento de C . Vale ressaltar que o MSE calculado para otimização da rede neural se dá em função do conjunto de amostras de um *cluster* ao longo do tempo, e não da intensidade de energia reconstruída. Por este gráfico, percebe-se que mesmo com a variação da quantidade de canais na terceira camada, não houve melhora notável nos valores de RMSE e não parece haver correlação direta entre o aumento da quantidade de canais e a melhoria de desempenho para a reconstrução de energia.

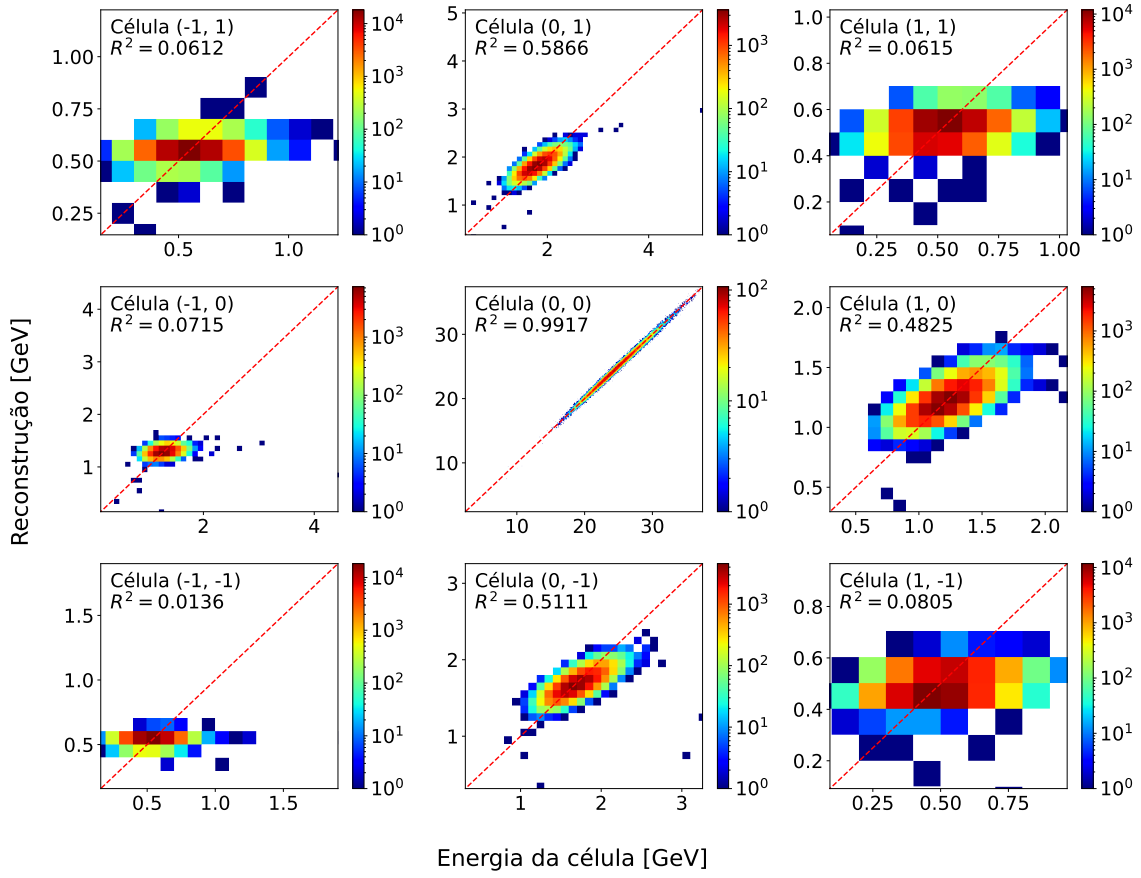


Figura 13. Qualidade de reconstrução de energia de *clusters* 3×3 para o modelo com $C = 60$. A linha tracejada em vermelho representa a tendência caso a predição fosse perfeita. O eixo horizontal de cada gráfico representa a energia da respectiva célula registrada no conjunto de testes, e o eixo vertical representa a estimação por OF a partir da saída da CNN. Resolução de 100 MeV/bin.

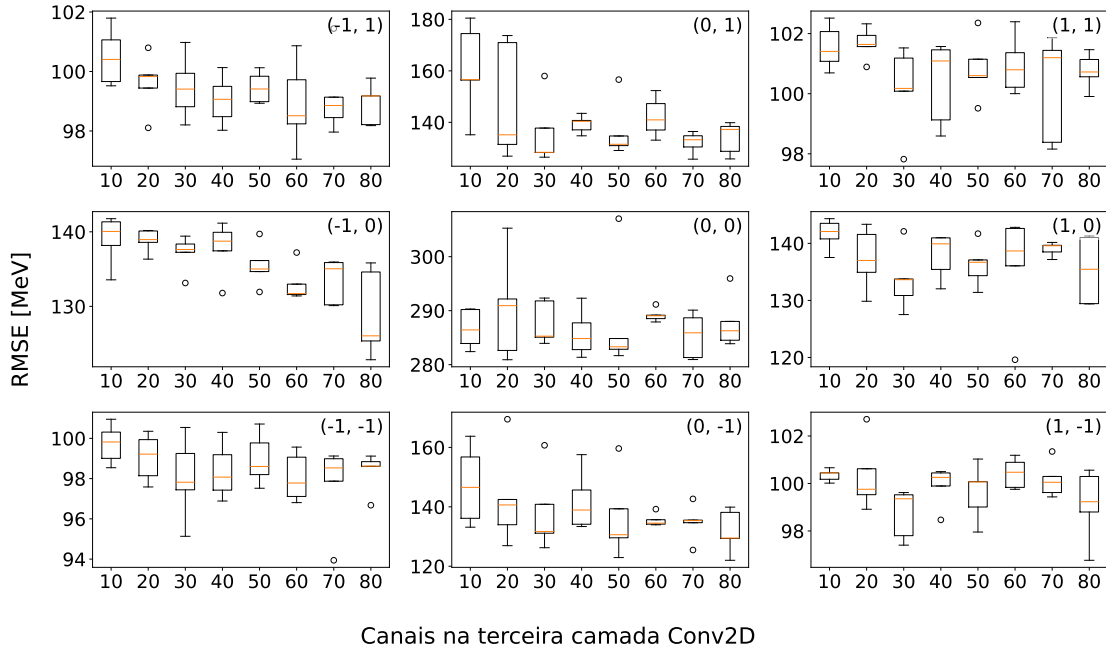


Figura 14. Distribuição do erro de estimação de energia para cada célula do *cluster* 3×3 . Cada *boxplot* representa a distribuição de RMSE dos 5 *fold*s treinados para cada configuração de C , considerando a melhor iteração de cada *fold*.

VI. CONCLUSÃO

Os resultados obtidos mostraram que, em um cenário controlado, as CNNs foram eficazes na recuperação da energia da célula central do *cluster*, com coeficiente de determinação superior a 0,99. Para as células vizinhas, embora tenha havido melhora significativa em relação aos dados brutos afetados pelo *crosstalk*, o desempenho ainda necessita de melhorias, o que sugere a necessidade de avaliar outras estratégias para a arquitetura. Possíveis estratégias para futuras abordagens são o uso de uma função de custo que permita otimizar a rede para prover uma boa reconstrução de energia e de tempo, e compensações para as células de menor nível energético, como a normalização dos dados. No contexto de mitigação de *crosstalk*, a normalização ainda é algo que precisa ser bem discutido e delineado, pois diferentemente do cenário controlado utilizado para os experimentos deste artigo, as partículas incidentes não terão exatamente o mesmo nível de energia.

A estimação do tempo de voo apresentou desempenho insatisfatório, com R^2 tendendo a zero, o que evidencia a limitação das CNNs bidimensionais em capturar a dinâmica temporal das amostras quando estas são tratadas apenas como canais. Este resultado aponta para a necessidade de empregar abordagens que incorporem explicitamente a dependência temporal dos dados, como estruturas recorrentes, ou estruturas híbridas que permitam explorar tanto a característica espacial quanto a característica temporal dos eventos.

Ainda há outros desafios a serem tratados em futuros trabalhos, os quais estão relacionados à variabilidade de diversos fatores, como a diferença de granularidade dos sensores em diferentes regiões do detector, a variação dos níveis de energia das partículas incidentes, a deterioração de resolução dos detectores a níveis de energia menores, a variação do ângulo de incidência e flutuações no ponto de colisão em relação ao ponto nominal.

AGRADECIMENTOS

Os autores agradecem à FAPESB, ao CNPq e à CAPES (programa CAPES/COFECUB) pelo apoio financeiro.

REFERÊNCIAS

- [1] G. Barr, R. Devenish, R. Walczak, and T. Weidberg, *Particle Physics in the LHC Era*. Oxford University Press, 2016. [Online]. Available: <https://doi.org/10.1093/acprof:oso/9780198748557.001.0001>
- [2] M. Livan and R. Wigmans, *Calorimetry for Collider Physics, an Introduction*. Springer Cham, 2019. [Online]. Available: <https://doi.org/10.1007/978-3-030-23653-3>
- [3] M. Delmastro, “Energy reconstruction and calibration algorithms for the ATLAS electromagnetic calorimeter,” Ph.D. dissertation, Università degli Studi di Milano, 2003.
- [4] H. Kolanoski and N. Wermes, *Particle Detectors: Fundamentals and Applications*. Oxford University Press, 2020. [Online]. Available: <https://doi.org/10.1093/oso/9780198858362.001.0001>
- [5] The ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, Aug 2008. [Online]. Available: <https://dx.doi.org/10.1088/1748-0221/3/08/S08003>
- [6] P. Langacker, *The Standard Model and Beyond*, 2nd ed. Taylor & Francis, 2017. [Online]. Available: <https://doi.org/10.1201/b22175>
- [7] M. S. Santos, P. C. M. A. Farias, E. F. Simas Filho, B. Laforge, and J. M. Seixas, “Redes Neurais como um método de redução do Crosstalk entre sensores no calorímetro eletromagnético do experimento ATLAS,” in *XV Congresso Brasileiro de Inteligência Computacional (CBIC)*, 2021.
- [8] —, “Redes Neurais Artificiais como algoritmo de reconstrução da informação depositada no Calorímetro Eletromagnético do Experimento ATLAS reduzindo os efeitos de Crosstalk,” in *XVI Congresso Brasileiro de Inteligência Computacional (CBIC)*, 2023.
- [9] A. L. Solis, “Contribution à la mesure d’énergie et du temps des électrons et des photons dans l’expérience ATLAS et recherche de la production de matière noire en association avec un boson de Higgs,” Ph.D. dissertation, Université Pierre et Marie Curie, 2017. [Online]. Available: <https://theses.hal.science/tel-02430126>
- [10] M. V. de Araújo, J. V. da Fonseca Pinto, and J. Seixas, “Filtragem Online Baseada na Informação do Sistema Calorimetria do Experimento ATLAS Utilizando Modelos de Aprendizado de Máquinas,” in *XV Congresso Brasileiro de Inteligência Computacional (CBIC)*. Joinville, SC: SBIC, 2021.
- [11] E. Simas, P. Farias, E. Souza, J. Marin, W. Freund, J. Seixas, J. Fonseca, and B. Laforge, “Uma Proposta para Calibração do Sistema Online de Seleção de Eventos no Detector ATLAS utilizando Gradient Boosted Decision Trees,” in *XV Congresso Brasileiro de Inteligência Computacional (CBIC)*. Joinville, SC: SBIC, 2021.
- [12] R. Vianna, J. Marin, and J. Seixas, “Transferência de conhecimento para filtragem online de partículas baseada em calorimetria de altas energias,” in *XV Congresso Brasileiro de Inteligência Computacional (CBIC)*. Joinville, SC: SBIC, 2021.
- [13] C. C. Dias, B. S.-M. Peralva, M. Hufnagel, A. M. e Silva, L. M. de Andrade Filho, A. S. Cerqueira, and J. Seixas, “Estimação de energia assistida por redes neurais para o calorímetro de telhas do atlas no lhc,” in *XV Congresso Brasileiro de Inteligência Computacional (CBIC)*. Joinville, SC: SBIC, 2021.
- [14] M. V. Araújo, M. Begalli, W. S. Freund, G. I. Gonçalves, M. Khandoga, B. Laforge, A. Leopold, J. L. Marin, B. S.-M. Peralva, J. V. F. Pinto, M. S. Santos, J. M. Seixas, E. F. Simas Filho, and E. E. P. Souza, “Lorenzetti Showers - A general-purpose framework for supporting signal reconstruction and triggering with calorimeters,” *Computer Physics Communications*, vol. 286, p. 108671, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0010465523000164>
- [15] E. F. Simas Filho, “Análise não-linear de componentes independentes para uma filtragem online baseada em calorimetria de alta energia e com fina segmentação,” Ph.D. dissertation, Universidade Federal do Rio de Janeiro, 2010. [Online]. Available: https://minerva.ufrj.br/F/?func=direct&doc_number=000756421&local_base=UFR01
- [16] L. Evans and P. Bryant, “LHC Machine,” *Journal of Instrumentation*, vol. 3, no. 08, pp. 1–158, Aug 2008. [Online]. Available: <https://dx.doi.org/10.1088/1748-0221/3/08/S08001>
- [17] E. E. P. Souza, “Conjunto de Classificadores Neurais alimentados por informação especialista para filtragem online em um detector de partículas,” Ph.D. dissertation, Universidade Federal da Bahia, 2021.
- [18] C. Gabaldón, “Drift Time measurement in the ATLAS Liquid Argon electromagnetic calorimeter using cosmic muons,” CERN, Geneva, Tech. Rep., 2010. [Online]. Available: <https://cds.cern.ch/record/1301524>
- [19] W. Cleland and E. Stern, “Signal processing considerations for liquid ionization calorimeters in a high rate environment,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 338, no. 2, pp. 467–497, 1994. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0168900294913323>
- [20] J. Labbé and R. Ishmukhametov, “Crosstalk Measurements in the Electromagnetic Calorimeter during ATLAS Final Installation,” CERN, Geneva, Tech. Rep., 2008.
- [21] J. Colas, C. de La Taille, R. Lafaye, N. Massol, P. Pralavorio, D. Sauvage, and L. Serin, “Crosstalk in the ATLAS Electromagnetic Calorimeter; preliminary version,” CERN, Geneva, Tech. Rep., 1999. [Online]. Available: <https://cds.cern.ch/record/683952>
- [22] D. P. Kingma and J. L. Ba, “Adam: A Method for Stochastic Optimization,” *3rd International Conference for Learning Representations*, 2014.