

DETECÇÃO DE SINAL EM CALORIMETRIA DE ALTAS ENERGIAS UTILIZANDO REDES NEURAIS ARTIFICIAIS

MATEUS HUFNAGEL MARANHA DE FARIA*, LUCIANO MANHÃES DE ANDRADE*, JOÃO PAULO BITTENCOURT DA SILVEIRA DUARTE*, EDUARDO SIMAS FILHO†, JOSÉ MANOEL SEIXAS‡

**Universidade Federal de Juiz de Fora
Juiz de Fora, Minas Gerais, Brasil*

†*Universidade Federal da Bahia
Salvador, Bahia, Brasil*

‡*Universidade Federal do Rio de Janeiro
Rio de Janeiro, Rio de Janeiro, Brasil*

Emails: mateus.hufnagel@engenharia.ufjf.br, luciano.andrade@engenharia.ufjf.br, joao.duarte@engenharia.ufjf.br, eduardo.simas@ufba.br, seixas@lps.ufrj.br

Abstract— Signal superposition is a recurrent problem when the information is sent through a communication channel in a rate higher than its response time. Modern particle colliders, running with collision rate per second of tens of millions of subatomic particles exhibit signal superposition in their front-end electronics, hindering the selection of events of interest made by the trigger systems of the particle detector. Currently proposed deconvolution methods consider the front-end electronics as a linear transmission channel whose transference function can be cancel using linear channel equalization techniques, allowing the cancellation of the inter-symbolic interference and the correct interpretation of data, previously superimposed. Optimal approximations of this equalizer are proposed using FIR filters. However, this linear channel compensation procedure does not take into account the intrinsic non-linearities present in the sensors outputs. In this work, instead of using a linear channel equalization process, a neural network (NN) classifier is used to perform the detection procedure of the signal of interest. Simulations were used to compare results of the detection efficiency of two the methods in different event rate scenarios.

Keywords— calorimeter, neural network, detection, signal superimposition.

Resumo— A sobreposição de sinais é um problema recorrente quando se envia informação através de um canal de comunicação à uma taxa superior ao tempo de resposta desse canal. Colisionadores de partículas modernos operam em uma taxa de dezenas de milhões de colisões por segundo, o que geralmente leva a uma sobreposição de sinais na eletrônica de *front-end*, dificultando a seleção de eventos de interesse realizada pelo sistema de *trigger* nos detectores de partículas. Métodos de desconvolução atualmente propostos consideram a eletrônica de *front-end* como um canal de transmissão linear, no qual sua função de transferência pode ser cancelada utilizando técnicas lineares de equalização de canal, permitindo o cancelamento da interferência inter-simbólica e a correta interpretação dos dados anteriormente sobrepostos. Aproximações ótimas desse equalizador são propostas usando filtros FIR. No entanto, esse procedimento de compensação linear não leva em consideração as não-linearidades nas saídas dos sensores. Nesse trabalho, ao invés de utilizar um processo de equalização linear, um classificador neural (RNA) é usado para realizar o procedimento de detecção do sinal de interesse. Resultados advindos de simulação comparam a eficiência dos dois métodos em diferentes cenários de taxas de eventos.

Palavras-chave— calorímetro, rede neural, detecção, sobreposição de sinal

1 Introdução

A sobreposição de sinais é um problema recorrente quando se envia informação através de um canal de comunicação à uma taxa superior ao tempo de resposta do mesmo, ocasionando a deformação do sinal e dificultando a equalização do canal. A Física experimental de Altas Energias (*High Energy Physics* - HEP) busca encontrar os componentes fundamentais da matéria e entender como os mesmos interagem entre si. Os experimentos desta área podem ser grandes instalações projetadas, construídas e operadas por pesquisadores de diferentes países. Na área, podemos mencionar o Colisionador Relativístico de Íons Pesados (*Relativistic Heavy Ion Collider*) (Harrison et al., 2003) no Laboratório Nacional de Brookhaven, Estados Unidos, o Complexo de Pesquisa e Acelerador de Prótons do Japão (J-PARC) (Nagamiya, 2012)

em KEK e o Grande Colisor de Hádrons (*Large Hadron Collider* - LHC), localizado no CERN (*Centre Européenne pour La Recherche Nucléaire*), Suíça (Evans and Bryant, 2008).

Um grande número de importantes descobertas em HEP vêm dos colisores de partículas, que são projetados para acelerar e colidir pacotes de partículas com o objetivo de produzir os componentes fundamentais da matéria. Recentemente, uma partícula conhecida como Bóson de Higgs, foi identificada pelo LHC. O Bóson de Higgs foi previsto teoricamente por estar relacionado com a origem da massa, porém nunca havia sido observado experimentalmente (Aad et al., 2012; Atlas, 2016a).

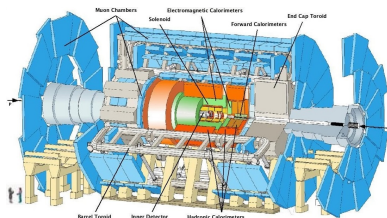


Figura 1: O detector ATLAS e seus sub-sistemas.

1.1 LHC e o CERN

O LHC, é atualmente, o maior acelerador de partículas do mundo. Construído à 100 metros de profundidade na divisa franco-suíça em um túnel circular de cerca de 27 km de circunferência, é o principal acelerador do CERN.

Associados ao LHC, existem vários detectores, localizados em pontos de colisão específicos e projetados para investigar aspectos relevantes para a física de partículas como: a origem da massa, matéria escura, novas partículas e fenômenos ainda não previstos teoricamente. (Atlas, 2016b).

Dentre estes experimentos, destacam-se o ALICE (*A Large Ion Collider Experiment*), CMS (*Compact Muon Solenoid*), LHCb e o ATLAS (*A Toroidal LHC Apparatus*) (Sotto-Maior, 2015), sendo este último o ambiente para o qual este trabalho foi desenvolvido.

1.2 O ATLAS e seu Calorímetro Hadrônico

O ATLAS é o maior experimento do LHC. Funciona através de uma colaboração internacional de 3800 cientistas de cerca de 180 instituições de todo o mundo, representando 38 países de todos os continentes. Assim como o CMS, é um experimento de propósito geral dentro da física de partículas, projetado para explorar todo o potencial de descobertas do LHC (Atlas, 2016a).

O ATLAS pesa cerca de 7000 toneladas e mede cerca de 25 metros de altura e cerca de 45 metros de largura. Tem formato cilíndrico e é constituído por sub-sistemas, que são organizados em camadas sobrepostas ao longo do túnel de feixe, ilustrados na Figura 1. Dentre eles destacam-se o Detector Central (*Inner Detector*), os Ímãs Solenoidais (*Solenoidal Magnets*), os calorímetros Hadrônico e Eletromagnético, o Toróide (*Toroid Magnets*) e o Espectrômetro de Múons (*Muon Spectrometer*) (Sotto-Maior, 2015; Atlas, 2016b).

Os calorímetros têm como propósito explorar os chuveiros de partículas que ocorrem no material detector, com o intuito de medir a energia das partículas produzidas nas colisões. Dentre os calorímetros, o Hadrônico (Tilecal) é responsável por medir a energia das partículas hadrônicas (ex.: prótons e nêutrons). É um sistema finamente seg-

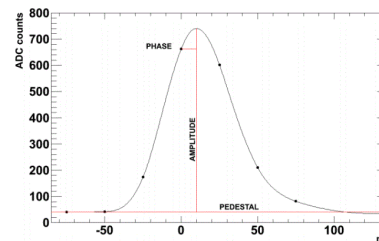


Figura 2: Forma do pulso do Tilecal. A taxa de amostragem do ADC (40 MHz) é sincronizada com os eventos de colisão (BC).

mentado, totalizando cerca de 10.000 canais de leitura, permitindo uma boa localização espacial do chuveiro de partículas (Grahn, 2008; Puzo, 2002; Aad et al., 2008; Sotto-Maior, 2015).

Considerando a alta resolução presente no Tilecal e a alta taxa de cruzamento de feixes (*bunch-crossing* - BC), de 40 MHz, é então gerada uma grande quantidade de dados pelo experimento. Com o intuito de lidar corretamente com todos esses dados, um sistema de *trigger* online, composto de três níveis, é necessário para realizar a seleção de informação relevante à ser armazenada (Watts, 2003; CMS, 2006). Uma estratégia comumente utilizada é usar o primeiro nível de *trigger* (L1) implementado em hardware dedicado, o qual tem a tarefa de decidir, de forma muito rápida, enquanto lida com uma taxa de eventos extremamente alta. Os eventos que foram aceitos pelo L1 são enviados para uma análise mais precisa em um *trigger* de nível superior, o qual é geralmente implementado em software. Apenas os eventos aprovados nesse nível são gravados em mídia permanente para análise posterior (offline).

O calorímetro tem um papel central no sistema de *trigger* (Wigmans, 2000). Esse sub-detector é responsável por medir os perfis de deposição de energia das partículas que chegam à ele. O nível crescente de taxa de evento no LHC gera uma superposição dos sinais nos canais de leitura dos calorímetros. Um sinal de leitura típico do Tilecal dura cerca de 150ns (veja Figura 2) abrange vários BC's adjacentes (1 a cada 25ns no LHC), podendo sofrer de superposição de sinal em condições de alta luminosidade. Essa justaposição de sinais degrada a forma do pulso original, comprometendo a performance dos métodos de detecção do sinal e estimação de energia atual, como o filtro FIR Linear (M de A Filho et al., 2015).

2 Definição do Problema

A primeira tarefa do sistema de *trigger* do L1 é identificar o índice do BC com depósito de energia relevante. Para esse propósito, a informação de saída do calorímetro é crucial. No entanto, em colisores com alta taxa de evento esse procedimento

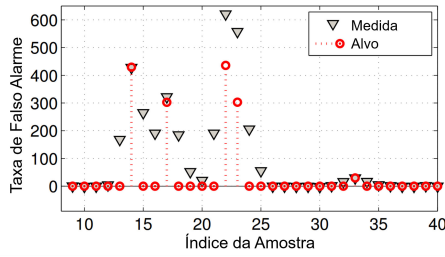


Figura 3: Resultado de sobreposição de sinal em condições de alta luminosidade.

acaba se tornando um desafio, uma vez que os sinais do calorímetro acabam por sobrepor-se.

A Figura 3 exemplifica este problema. Observa-se, nesta figura, a deposição real de energia em cada BC dada pelas amostras circulares. A informação digitalizada da saída do calorímetro é a convolução linear de resposta ao impulso de sua eletrônica de *front-end* (como o pulso da Figura 2) e o padrão de deposição de energia real. De modo a recuperar este valor real, técnicas de desconvolução linear podem ser aplicadas (M de A Filho et al., 2015). Para tal, a implementação de um filtro de Resposta ao Impulso Finita (do inglês, (FIR) é adequado aos níveis de *trigger* inferiores devido à sua rápida e fácil implementação nas modernas FPGA's (Meyer-Baese, 2007).

Apesar de apresentar uma performance satisfatória sob certas condições operacionais, a desconvolução linear é uma técnica não-ótima para o procedimento de compensação do calorímetro. Várias fontes de ruído não-Gaussiano - como desvio de fase do pulso, diafonia (*cross-talk*), deformação do pulso em altas amplitudes e ruído eletrônico - deformam o pulso de referência do calorímetro. Esses ruídos introduzem não-linearidades as quais não são levadas em consideração pelo filtro linear implementado.

Neste trabalho, é proposto um detector de deposição de energia nos BC's baseado em um classificador neural Perceptron Multi-camadas. O método é capaz de lidar com não-linearidades intrínsecas, ao mesmo tempo que é implementado um processo conhecido como *feed-forward*, onde as amostras entram na rede e passam através das camadas de neurônios até a saída, sem realimentação (Haykin, 1994). Este processo é adequado para os sistemas de *trigger* inferiores que funcionam à uma alta taxa de eventos e necessitam de rápida ação.

3 Desconvolução através de filtros FIR

O processo de desconvolução, baseado em filtros FIR, utiliza uma janela composta das W amostras mais recentes da saída do sistema digital em questão de forma a recuperar a informação correspondente da amostra de índice central dessa janela

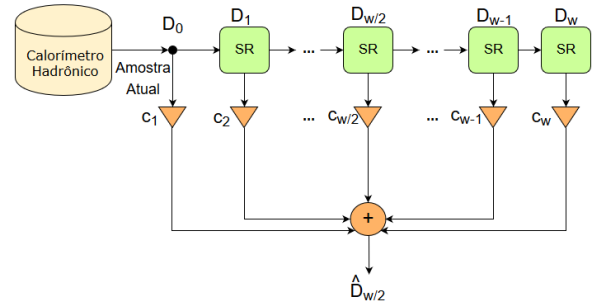


Figura 4: Processo de desconvolução utilizando filtro FIR.

através da convolução linear das W amostras da janela. A Figura 4 mostra esse processo. Uma cascata de W registradores de deslocamento (SR) armazenam os últimos W valores do ADC a partir da saída do sistema digital, os quais são representados por D_1 a D_W nesta figura. Essas amostras são combinadas linearmente através de c_i coeficientes, produzindo a saída $\hat{D}_{W/2}$, que representa a estimativa da energia alvo na posição central dentro da janela. A detecção do BC com sinal é feita aplicando-se um patamar de decisão diretamente à este valor. Deste modo, esse processo de desconvolução implica em um atraso temporal de $W/2$ amostras.

Existem na literatura várias abordagens para encontrar os coeficientes do filtro de desconvolução. Quando pares de amostras de alvos e valores medidos são conhecidos com antecedência, por exemplo, através de simulação de Monte Carlo (Mooney, 1997), o método mais adequado para o projeto do filtro é a formulação dos Mínimos Quadrados Linear (Linear LS). Supondo-se uma matriz $\mathbf{H}$ de dimensões $W \times N$, correspondente à N realizações de uma sequência com W valores de ADC da saída do sistema, e um vetor $\mathbf{d}$ contendo os N valores alvo associados à amostra central de cada janela. O sistema de equações que aproxima linearmente o processo de desconvolução é

$$\mathbf{H}\mathbf{c} = \mathbf{d} \quad (1)$$

onde

$$\mathbf{c} = [c_1, \dots, c_w] \quad (2)$$

é composto pelos coeficientes do filtro que devem ser determinados para minimizar o erro médio quadrático (MSE) entre o valor estimado e os valores-alvo reais. A solução para o vetor $\mathbf{w}$ para a função custo do MSE é (Kay, 1993)

$$\mathbf{c} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{d} \quad (3)$$

4 Método Proposto

Uma rede neural perceptron multi-camadas (MLP) foi usada neste trabalho para substituir os filtros FIR para a detecção do sinal no calorímetro.

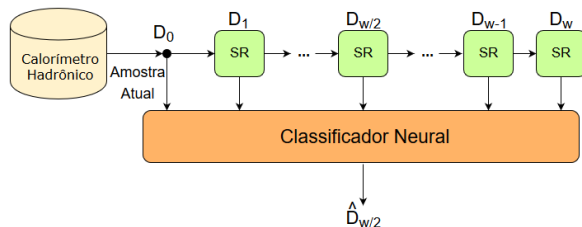


Figura 5: Detecção do sinal utilizando um classificador neural MLP.

Foi usada uma camada escondida de neurônios e a função de ativação dos neurônios utilizada foi a tangente-hiperbólica. A escolha do número de neurônios escondidos foi feita experimentalmente através de testes exaustivos entre diferentes configurações, avaliando a performance de detecção do sinal. Para o treinamento das redes, foi utilizado o algoritmo de Levenberg-Marquardt (Liu, 2010).

Como ilustrado na Figura 5, a MLP foi alimentada a partir de uma janela de amostras de energia. Usando o conjunto de dados para treinamento, a rede neural foi projetada para detectar quando o sensor do calorímetro foi ativado por uma real interação de partículas e descartar os sinais falsos produzidos por interações anteriores (devido ao tempo de resposta do sensor).

O conjunto de dados utilizado para treinamento e teste do método proposto foi obtido através de simulações de Monte Carlo do Tilecal. O simulador foi ajustado para criar cenários próximos aos operacionais, adicionando características do sistema como: ruído de fundo eletrônico, desvio de fase do pulso característico e diferentes condições operacionais do detector ao variar a ocupância do canal (que é diretamente proporcional à taxa de eventos), influenciando diretamente na intensidade do ruído de empilhamento de sinais (M de A Filho et al., 2015).

Como modo de comparação da performance dos detectores foi utilizada a curva de Característica de Operação do Receptor (ROC) (Van Trees, 2001). A curva ROC ilustra como a performance de discriminação (em termos das taxas de eficiência de detecção e falso alarme) varia enquanto o limiar de decisão do classificador é alterado.

5 Resultados

Um total de 80.000 amostras de energia simuladas compõem o conjunto de dados utilizados, o qual foi dividido em subconjuntos de treinamento, validação e teste, correspondendo à 60%, 20% e 20% do conjunto de dados completo, respectivamente. Os subconjuntos de treinamento e validação representam uma tomada de dados com 45% de ocupância (i.e., a probabilidade de ocorrência de uma colisão em cada BC). Foram utilizados 9 subconjuntos de teste diferentes com ocupância do sinal variando

Tabela 1: Taxa de falso alarme mínima por número de atrasos para o classificador neural.

Atrasos	5	7	9	11	15	23
FA (%)	11,0	10,1	9,6	9,7	9,5	8,5

de 10% à 90%, cada um deles composto de 16.000 amostras. Os dados das curvas ROC contendo as taxas de falso alarme (FA) pelas de probabilidade de detecção (PD) foram gravadas para cada topologia de rede como sujeito de comparação e projeto do classificador neural. O processo de treinamento foi inicializado 3 vezes.

O primeiro passo para o projeto dos classificadores foi definir uma ocupância de teste fixa em 30% para comparar o classificador proposto com o linear. O projeto do classificador neural teve início ao definir o número de atrasos de entrada. Para isso, o número de neurônios na camada oculta foi fixado em 10, e a PD do classificador foi definida em 90%. A Figura 6 ilustra a taxa de FA enquanto o número de atrasos de entrada do classificador cresce para cada inicialização. A Tabela 1 mostra o melhor desempenho obtido em percentual pelos três treinamentos da Figura 6. Foi observado que resultados similares foram alcançados a partir de 9 atrasos de entrada, apesar da menor taxa de FA encontrada foi com 23 atrasos. Esse tipo de aplicação requer uma resposta rápida e é preferível utilizar um número reduzido de atrasos, como 9 neste caso.

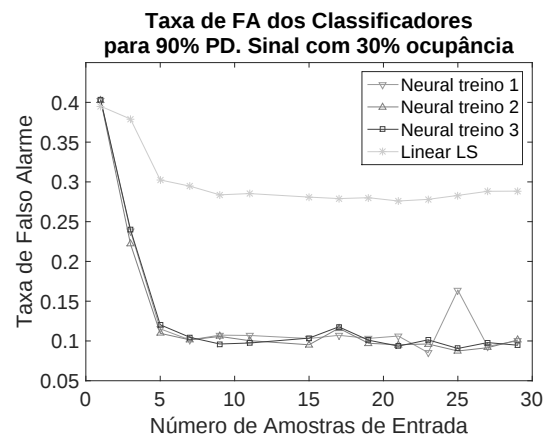


Figura 6: Performance dos classificadores neural e linear LS para diferentes números de atrasos.

O próximo passo do projeto do classificador é escolher o número de neurônios escondidos. Nesse caso, foram utilizadas 9 atrasos de entrada para alimentar os classificadores e o procedimento de treinamento foi também inicializado três vezes, variando o número de neurônios escondidos de 1 à 12. Pode-se observar na Figura 7 utilizando o mesmo critério aplicado na escolha do número de atrasos, que resultados similares foram encontrados a partir de 7 neurônios ocultos, como mos-

Tabela 2: Taxa de falso alarme mínima obtida por número de neurônios do classificador neural.

Neurônios	3	4	6	7	12
FA (%)	17,7	11,3	11,5	10,4	10,0

Tabela 3: Taxa de falso alarme mínimo por número de atrasos obtidos pelo método linear.

Atrasos	17	19	21
FA (%)	27,9	28,0	27,6

trado na Tabela 2 e na Figura 7. O classificador mencionado por último (7 neurônios) deve ser considerado em uma implementação da técnica proposta em um sistema embarcado online, por apresentar um custo computacional reduzido.

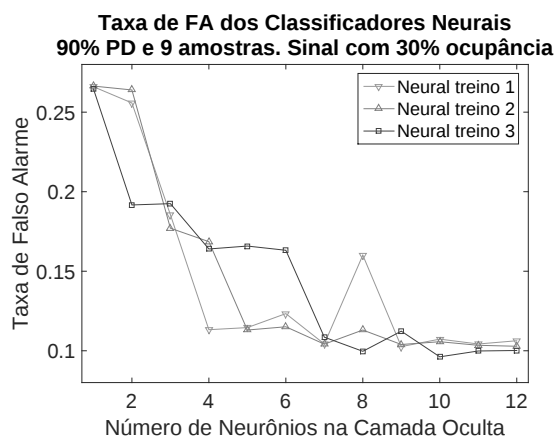


Figura 7: Performance do classificador neural para diferentes números de neurônios.

Os resultados obtidos a partir do projeto do classificador proposto foram comparados com o método linear mais utilizado. Para isso, um detector linear foi projetado para o conjunto de dados utilizado escolhendo o número apropriado de atrasos de entrada. Pode-se observar na Figura 6 e na Tabela 3 que a taxa de FA pouco muda a partir de 9 atrasos de entrada, porém, a menor taxa de FA foi atingida para 21 atrasos.

Para uma comparação adequada entre os métodos, as curvas ROC dos classificadores que alcançaram a menor taxa de FA, para uma ocupância de 30%, são ilustrados na Figura 8. Observa-se que um número de neurônios na camada escondida maior que 7 pouco reduz a taxa de FA do classificador neural. Para comparar a capacidade de generalização dos classificadores, foram utilizados conjuntos de dados simulados para diferentes valores de ocupância do canal, a probabilidade de detecção foi fixada novamente em 90% e os resultados estão ilustrados na Figura 9.

Para valores de ocupância do canal inferiores à 70%, os benefícios de utilizar o método não-linear são mais relevantes e os classificadores com um maior número de neurônios na camada oculta

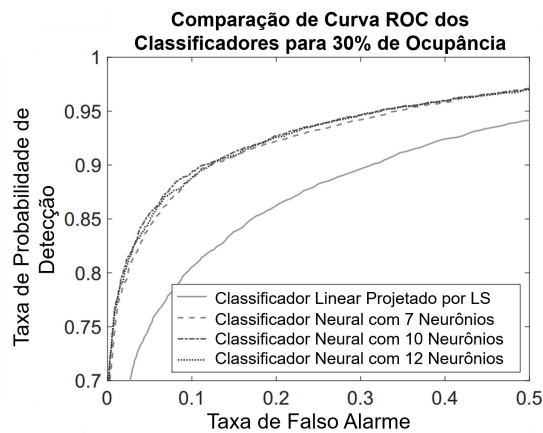


Figura 8: Curvas ROC dos melhores classificadores.

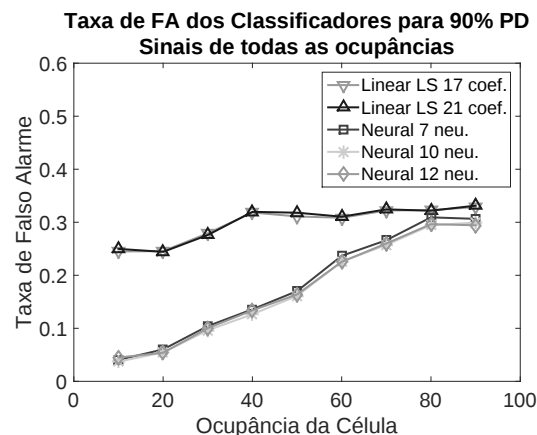


Figura 9: Comparação da capacidade de generalização dos classificadores para um cenário com diferentes ocupâncias.

não obtiveram melhor performance quando comparado ao de 7 neurônios. Enquanto a ocupância cresce além deste valor, o desempenho de ambos os métodos tendem à se tornar equivalente. É importante ressaltar que na maioria dos casos, a ocupância típica dos canais do Tilecal (atualmente em cerca de 30%) está no alcance em que o método proposto apresenta melhores resultados.

Como o método proposto requer implementação online em hardware dedicado para atender os requisitos de latência de tempo do primeiro nível de trigger, uma simples estimativa dos requisitos computacionais foi feita. Para isso, foram consideradas três operações digitais: multiplicações, somas e acessos a tabela de pesquisa (*look-up tables* - LUT).

Como ilustrado na Tabela 4, o número de operações digitais aumentam consideravelmente para o método não linear, no entanto, seus efeitos na latência de tempo do sistema depende do tipo de hardware utilizado. Uma estimativa mais precisa dos requisitos do método proposto está sendo realizada utilizando uma implementação em FPGA

Tabela 4: Comparação das operações digitais necessárias para implementação dos métodos.

Método	Mult.	Somas	LUT
Linear LS	21	20	0
Neural (7 neu.)	78	70	7
Neural (10 neu)	111	100	10
Neural (12 neu)	133	120	12

e será assunto de trabalhos futuros.

6 Conclusões

Calorímetros são sistemas de medição os quais representam um papel importante para a identificação de partículas em física de altas energias. Os colisores de partículas mais modernos estão sendo projetados para operar com altas taxas de colisão com o intuito de aumentar a probabilidade de encontrar um fenômeno físico relevante. Enquanto a taxa de colisão cresce, sobreposição de sinais torna-se um problema sério para detecção e estimação nos calorímetros. Este trabalho propôs um sistema de detecção de sinal em calorímetros de alta energia, baseado em uma rede neural *feed-forward*. Resultados utilizando um conjunto de dados simulados mostraram que a performance de detecção aumentou consideravelmente, quando comparado à utilização do detector linear. Uma implementação em FPGA do sistema proposto está sendo desenvolvida e será apresentada em futuros trabalhos.

Agradecimentos

Os autores gostariam de agradecer ao CNPq, CAPES, FAPERJ, FAPESB, RENAFAP, MCTI, CERN (Suíça) e União Européia pelo apoio financeiro.

Referências

Aad, G., Abajyan, T., Abbott, B. and Abdallah (2012). Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc, *Physics Letters B* **716**(1): 1–29.

Aad, G., Abat, E., Abdallah, J., Abdelalim et al. (2008). The atlas experiment at the cern large hadron collider, *Journal of Instrumentation* **3**(8).

Atlas, C. (2016a). Atlas discover, Acesso em 30 de Março de 2016. <http://atlas.cern/discover/>.

Atlas, C. (2016b). Atlas masterclass. hands on particle physics, Acesso em 7 de Abril de 2016. <http://atlas.physicsmasterclasses.org/pt/>.

CMS, C. (2006). The cms high level trigger, *The European Physical Journal C-Particles and Fields* **46**(3): 605–667.

Evans, L. and Bryant, P. (2008). Lhc machine, *Journal of Instrumentation* **3**(08).

Grahn, K.-J. (2008). *ATLAS Calorimetry: Hadronic Calibration Studies*, PhD thesis.

Harrison, M., Ludlam, T. and Ozaki, S. (2003). Rhic project overview, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **499**(2): 235–244.

Haykin, S. (1994). *Neural networks. A Comprehensive Foundation*, Macmillan College.

Kay, S. M. (1993). *Fundamentals of statistical signal processing, volume I: estimation theory*, Prentice Hall - Alan V. Oppenheim, Series Editor.

Liu, H. (2010). On the levenberg-marquardt training method for feed-forward neural networks, *Natural Computation (ICNC), 2010 Sixth International Conference on*, Vol. 1, IEEE, pp. 456–460.

M de A Filho, L., Peralva, B. S., M de Seixas, J. and Cerqueira, A. S. (2015). Calorimeter response deconvolution for energy estimation in high-luminosity conditions, *IEEE Transactions on Nuclear Science* **62**: 3265 – 3273.

Meyer-Baese (2007). Digital signal processing with field programmable gate arrays, **65**.

Mooney, C. Z. (1997). *Monte carlo simulation*, Vol. 116, Sage Publications.

Nagamiya, S. (2012). J-parc project and its science, *Progress in Particle and Nuclear Physics* **67**(2): 580–593.

Puzo, P. (2002). Atlas calorimetry, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **494**: 340 – 345.

Sotto-Maior, B. (2015). *Reconstrução de Energia para Calorímetros Finamente Segmentados*, PhD thesis.

Van Trees, H. L. (2001). *Detection, Estimation, and Modulation Theory, Part I*, Wiley. New York.

Watts, G. (2003). Review of triggering, *Nuclear Science Symposium Conference Record, 2003 IEEE*, Vol. 1, IEEE, pp. 282–287.

Wigmans, R. (2000). *Calorimetry: Energy measurement in particle physics*, Vol. 107, Oxford University Press.