

Redes Neurais Especialistas para a Separação Elétron-Jato usando Calorímetros Multi-camadas e Multi-segmentados

A.R. Anjos e J.M. Seixas*

18 de Fevereiro de 2002

Resumo

O experimento ATLAS estará operacional no ano de 2006. O objetivo principal deste experimento é a detecção do bóson de Higgs, usando entre outros tipos de detetores, calorímetros. Um dos canais de detecção mais importantes no experimento é o de elétrons com alta energia transversa, representando de 30 a 40% do total das assinaturas a serem analisadas pelo Sistema de Filtragem. Jatos (de partículas) confundem-se comumente com elétrons pela forma que interagem com os calorímetros. Neste trabalho, apresentamos um sistema de discriminação elétron-jato baseado em redes neurais especialistas, utilizando os dados dos calorímetros do ATLAS. Este sistema, depois de treinado, compacta o espaço de variáveis de entrada (células dos calorímetros) em um subespaço que mantém os aspectos necessários para uma detecção eficiente de elétrons. Os resultados apresentados se mostram melhores que os resultados obtidos usando-se técnicas desenvolvidas no CERN, com o mesmo objetivo.

1 Introdução

O experimento ATLAS, no CERN, Suíça, começará a operar em 2006. Neste experimento o detetor é composto por vários sub-detetores, todos posicionados de forma concêntrica ao redor do ponto de colisão. A física de interesse no experimento é o bóson de Higgs. Prevê-se que seja possível detetá-lo em alguns eventos por hora, colidindo-se pacotes de prótons a 14 TeV no ponto de impacto, cerca de 40 milhões de vezes por segundo.

O bóson de Higgs é extremamente instável, decaindo rapidamente em elementos mais estáveis e menos energéticos. O sistema de detecção adotado no experimento se baseia no estudo dos possíveis decaimentos de um bóson de Higgs para confirmar sua existência. Estes possíveis decaimentos são denominados comumente *assinaturas* do bóson.

A Figura 1 esquematiza o detetor ATLAS. No centro encontram-se os detetores de traços, enquanto que na parte exterior, os detetores de múons. Ao centro destacam-se os calorímetros [1, 2]. Estes últimos são capazes de medir a energia de partículas que atravessam seu volume.

1.1 O Sistema de Filtragem

Espera-se que, em média, um volume de dados de 2 megabytes seja gerado por evento [3]. Para a taxa de eventos por base de tempo no experimento, verifica-se a inviabilidade do armazenamento e posterior análise *off-line* dos dados. No período de 1 segundo, o experimento produzirá $2 \times 40 \times 10^6 = 80$ Terabytes. Em uma hora seriam quase 29 mil Terabytes.

Visando reduzir este imenso fluxo de dados, um sistema de filtragem de *on-line* está sendo projetado [3, 4]. Este sistema permitirá a separação de eventos que representem física ordinária daqueles que representam a física de interesse do experimento.

*Andre.Rabello@ufrj.br, seixas@lps.ufrj.br, UFRJ/EE/COPPE - Laboratório de Processamento de Sinais (LPS)
- CP 68504, 21945-970 Rio de Janeiro, Brasil

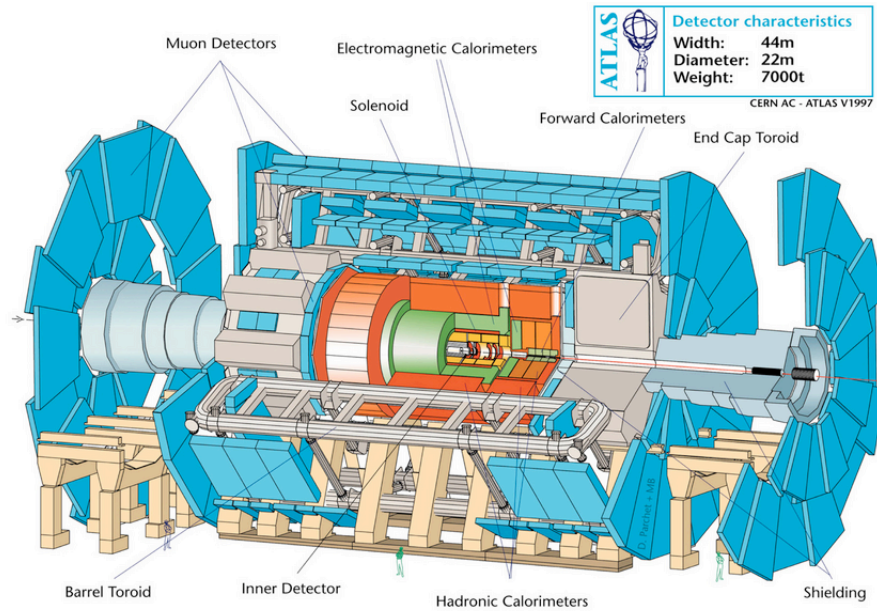


Figura 1: Diagrama esquemático do detector ATLAS.

O Sistema de Filtragem do ATLAS é dividido em 3 níveis conectados em cascata, permitindo uma análise mais refinada (e também mais lenta) de cada evento, a cada nível. Os eventos rejeitados são descartados, não sendo analisados pelos níveis posteriores de filtragem.

O Primeiro Nível (LVL1) é composto de *hardware* de baixa programabilidade, mas extremamente veloz. Seu objetivo é reduzir a taxa de eventos de entrada, de 40 milhões por segundo (40 MHz), para não mais que 100 kHz. O segundo (LVL2) e terceiro níveis (EF) serão construídos utilizando computadores tipo PC, conectados por chaves *gigabit ethernet*, devendo reduzir conjuntamente a taxa de entrada a um valor entre 10 e 100 Hz (100:1 e 10:1 respectivamente).

A seleção de eventos no LVL1 é realizada utilizando-se apenas um subconjunto dos detectores do ATLAS. Esta seleção é feita de forma que as regiões onde houver interação de partículas com o detector (*Regions of Interest*, RoI) sejam destacadas. O LVL2 continua com o processo de seleção, tomando as RoI's e utilizando os dados de todos os detectores para refinar a decisão sobre o evento. Este tipo de análise é chamado de *local*, pois somente leva em consideração a região localizada ao redor das RoI's destacadas pelo LVL1. O EF utiliza todo o detector para analisar o evento, e por isto consegue obter um refinamento ainda maior na decisão sobre um evento.

1.2 Detecção de elétrons no LVL2

No LVL1, muitos dos eventos são aprovados com base na detecção de um elétron com alta energia transversa no calorímetro. A detecção é feita baseando-se na concentração da deposição energética do objeto no calorímetro: quanto mais concentrada a deposição de energia, maior a probabilidade de ser um elétron. Cerca de 30 a 40% dos eventos disparados pelo LVL1 advêm da ocorrência de elétrons, e por isto esta assinatura é considerada uma das mais importantes no ATLAS. Para cada 25000 eventos disparados pelo LVL1 pela detecção de um elétron de alta energia nos calorímetros, apenas 1 será realmente um elétron. O *background* de elétrons é formado por jatos de partículas que interagem com os calorímetros de forma parecida, falseando os primeiros.

Se um elétron foi o precursor do disparo, do sistema de filtragem, o LVL2 começará a análise do evento buscando confirmar este objeto. Para isto, o LVL2 transfere os dados dos *buffers* conectados aos calorímetros correspondentes a RoI de interesse. A Figura 2 mostra a projeção de uma RoI no plano $\eta \times \phi$. Esta figura está organizada conforme a direção de penetração da partícula: de cima

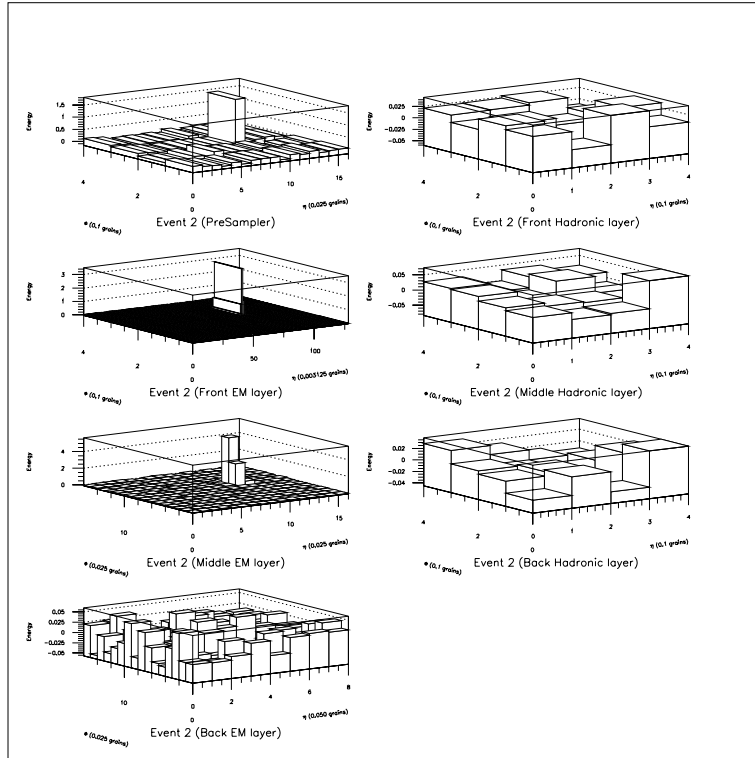


Figura 2: Um elétron típico interagindo com os calorímetros do ATLAS.

para baixo e da direita para esquerda afasta-se do ponto de colisão. Cada camada dos calorímetros possui uma granularidade particular. De uma forma geral, as camadas mais próximas do centro são mais granulares, otimizadas para a detecção de elétrons e as mais afastadas menos, otimizadas para a detecção de jatos de partículas. Cada pequeno bloco no gráfico representa a deposição energética proveniente da interação da partícula com aquela célula do calorímetro: quanto maior, mais energia foi depositada naquela posição. Neste caso podemos ver um elétron típico interagindo. Cada RoI possui cerca de 1000 células de calorimetria.

2 Detecção de elétrons usando Redes Neurais

Técnicas atuais de processamento local e *on-line* baseado em calorímetros utilizam um sistema que é baseado em 2 etapas [5, 6, 7]:

1. Pré-processamento: compactação do espaço de entrada em um número reduzido de características que definam o objeto de estudo;
2. Discriminação: utilizando as características definidas na fase anterior do processamento, constrói-se um discriminador que seja eficiente na detecção do objeto de estudo.

A compactação do espaço de entrada reduz o tempo utilizado na discriminação e também a influência do ruído no processo, já que combina a amostragem energética de diversas fontes em poucas variáveis.

Neste contexto, Redes Neurais[8] podem ser utilizadas tanto para o pré-processamento quanto para a discriminação baseada na compactação dos dados de entrada. Este trabalho aborda a utilização de Redes Neurais em ambos os aspectos. A Figura 3 esquematiza o processo de detecção adotado, que visa partir a tarefa complexa de processamento local do LVL2 em um número de

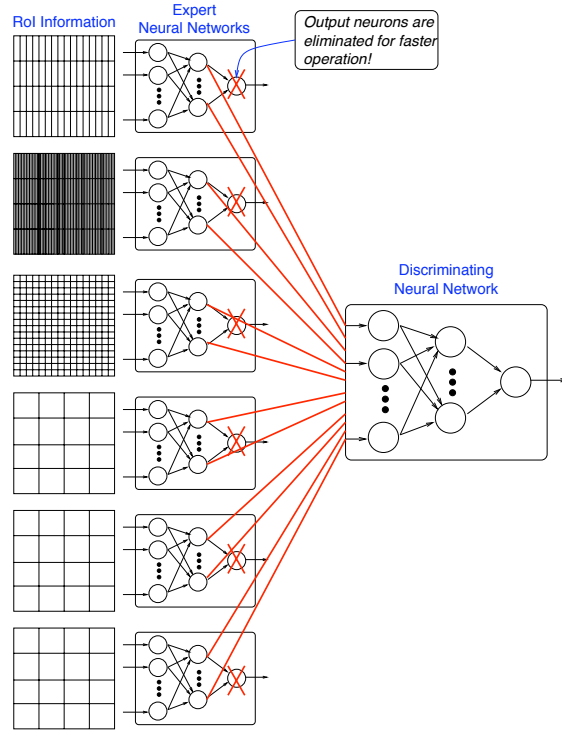


Figura 3: Diagrama esquemático do processamento usando Redes Neurais Especialistas.

redes neurais compactas que se dediquem a analisar cada segmento do sistema de calorimetria em detalhe.

2.1 Pré-processamento

Na Figura 3, é possível distinguir um número de redes neurais que estão conectadas a cada camada do calorímetro. Estas redes são responsáveis pela extração da informação que caracteriza os objetos de interesse (elétrons). O número de neurônios na (única) camada intermediária define a quantidade de características que deseja-se extrair. Estas redes são chamadas *especialistas* [9], pois conseguem caracterizar o objeto de estudo baseando-se somente nos dados da camada a que estão conectadas.

O número de neurônios na camada de entrada da rede dependerá, portanto, do número de células da camada do calorímetro na qual a rede está conectada. A camada de saída é composta por apenas um neurônio, utilizado para definir o ponto ótimo de treinamento do sistema, na identificação do objeto de estudo. O número de neurônios na camada intermediária de cada rede neuronal depende do tipo de camada a qual a rede está conectada: a rede conectada ao pré-irradiador usa 3 neurônios escondidos, aquela conectada à primeira camada EM usa 8, à segunda, usa 4. As redes conectadas às camadas na seção hadrônica possuem todas 2 neurônios na camada intermediária.

Treinamento Para treinarem-se as redes especialistas, utilizou-se um conjunto de dados composto de 600 elétrons e 3600 jatos, que foram igualmente divididos em sub-conjuntos de treinamento e teste. As redes são treinadas até que a eficiência de discriminação entre elétrons e jatos atinja estabilidade em um máximo. Para isto, ao invés de usar-se um critério de parada baseado no erro médio quadrático, utilizou-se uma figura de mérito baseada na eficiência de discriminação para ambas as classes de partículas.

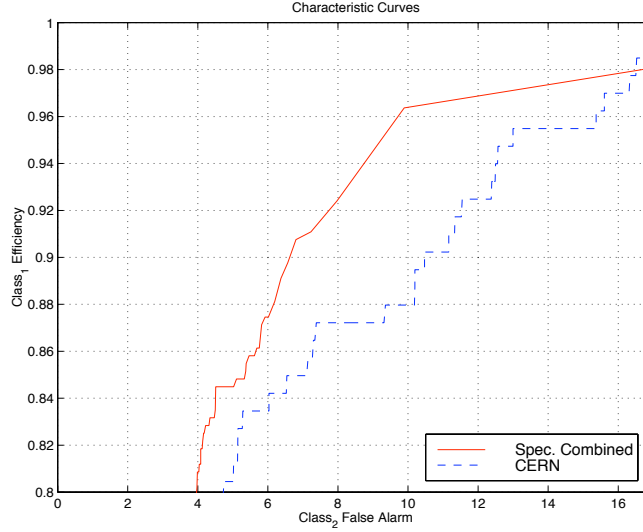


Figura 4: A Região de Convergência (RoC [10]) do discriminador neural.

2.2 Discriminação

Após treinar a rede visando a maximização da detecção de elétrons em cada camada individualmente, os neurônios de saída são eliminados, uma vez que indicam apenas a classificação neural limitada a cada camada, e o treinamento das redes especialistas é congelado. Uma única rede neural é conectada aos neurônios da camada intermediária de todas as redes especialistas, haja vista que a camada intermediária de cada rede especialista retém as características determinantes de cada segmento do sistema de calorimetria para que uma boa identificação das classes de partícula seja obtida. Esta rede de interconexão é treinada utilizando-se o resultado da injeção dos dados originais nas camadas de entrada das redes especialistas.

A rede de saída é treinada até que a eficiência de discriminação entre elétrons e jatos atinja estabilidade em um máximo. A saída desta rede representa a saída do processamento local, ou seja, o resultado da detecção de elétrons no LVL2, usando os calorímetros.

3 Resultados

A Figura 4 mostra graficamente o melhor resultado obtido, aplicando-se a técnica de Redes Neurais Especialistas para a compactação da informação discriminante contida em uma RoI dos calorímetros do ATLAS, e posterior discriminação, também utilizando redes neurais. O gráfico mostra, no eixo horizontal, a taxa de jatos que será admitida dada uma eficiência na discriminação de elétrons (eixo vertical).

Como é possível ver, o processo discriminativo proposto apresenta-se melhor que o método atualmente empregado no CERN [5], representado pela linha pontilhada. Este último, compacta o espaço de entrada em 4 variáveis, aplicando cortes sucessivos em histogramas bi-dimensionais destas variáveis. Estes resultados, porém, são piores que os apresentados em [7], utilizando uma técnica de compactação topológica em anéis, preservando a discriminação neural no fim do processamento.

Ao analisarmos a relevância [11] de cada célula no processo discriminativo, constatou-se que as células no centro da RoI são as que melhor definem a natureza do objeto sendo discriminado. A Figura 5 mostra a relevância na discriminação para as células na terceira camada do calorímetro (indicada como *EM Middle Layer* na Figura 2). Durante a análise da relevância, verificou-se que as últimas camadas do calorímetro, aquelas cuja a granularidade é menor pois são otimizadas para

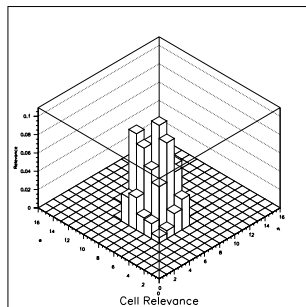


Figura 5: A relevância das células no processo discriminativo (*EM Middle Layer*).

a detecção de jatos, não influenciam no resultado da rede neural de discriminação.

4 Conclusões

O método para a detecção de elétrons descrito possui um desempenho melhor que o método equivalente proposto atualmente no CERN. A análise da relevância de cada camada no processo discriminativo revelou que as camadas do calorímetro com menor granularidade (indicadas na Figura 2 pela legenda *Hadronic Layer*) não são utilizadas pelo discriminador neural de saída, na detecção de elétrons.

Isto indica que a compactação dos dados de entrada utilizando tais camadas não é eficiente usando redes neurais especialistas, possivelmente devido ao excesso de ruído na camada. Elétrons são partículas que normalmente penetram pouco no detetor e, por esta razão, objetos identificados como elétrons pelo LVL1 apresentam pouco ou nenhum sinal nas últimas camadas do calorímetro. Um método híbrido que combine a utilização de redes neurais especialistas com uma compactação eficiente dos dados das últimas camadas do detetor seria, talvez, mais eficiente que o método proposto.

Como é possível notar na Figura 5, não há uma uniformidade na relevância das células centrais, indicando uma possível indecisão da rede neural sob o ponto de vista da distribuição da informação na RoI. Isto acontece pois o pico de deposição energética em cada camada se dá, a cada evento, em uma posição relativa diferente na RoI. Por esta razão, métodos independentes da posição relativa do pico de deposição na RoI apresentam melhores resultados. Um exemplo é o mapeamento topológico em anéis descrito em [7].

5 Agradecimentos

Os autores são gratos à CAPES e ao CNPq pelo apoio à pesquisa conduzida neste trabalho.

Referências

- [1] ATLAS LArg Unit. Liquid argon calorimeter - technical design report. Technical report, CERN, Dezembro 1996.
- [2] ATLAS/Tile Calorimeter Collaboration. Tile calorimeter - technical design report. Technical report, CERN, Dezembro 1996.
- [3] ATLAS Level-1 Trigger Group. First-level trigger technical design report. Technical report, CERN, 1998.

- [4] ATLAS Collaboration. ATLAS daq, ef, lvl2 and dcs technical progress report. Technical report, CERN, 1998.
- [5] S. Gonzalez, T. Hansl-Kozanecka, and M. Wielers. Selection of high-pt electromagnetic clusters by the level-2 trigger of ATLAS. ATLAS-Trigger DAq Note 002, CERN, Fevereiro 2000.
- [6] M.R. Vassali and J.M. Seixas. Principal component analysis for neural electron/jet discrimination on highly segmented calorimeters. *VII Internation Workshop on Advanced Computing & Analysis Techniques in Physics Research*, October 2000.
- [7] A. Anjos and J. Seixas. Neural electron-jet separation using multi-layered and multi-segmented calorimeters. *XXI Encontro Nacional de Física de Partículas e Campos*, Outubro 2000. **Disponível em inglês:** <http://www.lps.ufrj.br/~rabello/publications/index.html>.
- [8] Simon Haykin. *Neural Networks: A comprehensive foundation*. Prentice Hall, 1999.
- [9] J.M. Seixas, L.P. Calôba, A.L. Braga, and E.M.S. Moura. Global decisions with a neural second-level trigger system. *Computing in High Energy Physics*, (San Francisco, Estados Unidos):456–458, 1994.
- [10] H.L. Van Trees. *Detection, Estimation and Modulation*. John Wiley & Sons, 1966.
- [11] A. Gruber, J. Möck, J. Fröchtenicht, D. Goldner, C. Kiesling, T. Klobner, H. Kolanoski, T. Krämerkampfer, and P. Ribarics. Performance of backpropagation mentworks in the second level trigger of the h1-experiment. *New Computing Techniques in Physics Research*, 1994.