

# Redes Neurais Especialistas para a Identificação de Partículas em um Calorímetro Finamente Segmentado

Amanda Camacho Novaes de Oliveira, Werner Spolidoro Freund, e José Manoel de Seixas

Universidade Federal do Rio de Janeiro (UFRJ)  
Rio de Janeiro, RJ, Brasil

E-mails: [amandacnoliveira@poli.ufrj.br](mailto:amandacnoliveira@poli.ufrj.br), [wsfreund@gmail.com](mailto:wsfreund@gmail.com),  
[seixas@lps.ufrj.br](mailto:seixas@lps.ufrj.br)

**Resumo** Apresenta-se uma proposta de aplicação de Ciência de Dados no contexto de Física de Partículas, aplicando um ensemble de redes neurais para efetuar a discriminação de partículas a partir da fusão de dados de diferentes conjuntos de sensores finamente segmentados. A técnica procura identificar raros eventos de interesse durante fase de operação *online*, dentre um grande conjunto total de eventos.

Ao comparar a proposta estendendo a versão em vigor, observa-se um ganho de 1% em detecção e redução de 1,5% (valores centrais) em falso alarme para atuação na filtragem online de elétrons.

**Palavras-chave:** Rede Neural, Classificador, Redes Neurais Especialistas, Calorimetria, Variáveis de Traço, Filtragem *Online*, Data Fusion, Machine Learning, Data Science

## 1 Introdução

A Ciência de Dados é um campo multidisciplinar em amplo crescimento na atualidade, sua pesquisa sofrendo grande aumento com o surgimento das tecnologias modernas de computação. Dentro deste ramo abrangente, o aprendizado de máquina vem se tornando cada vez mais popular, com aplicações em reconhecimento de imagens e fala, detecção de vírus e spam, diagnósticos médicos, controle robótico, entre outros.

Este trabalho exemplifica uma aplicação de aprendizado de máquina, e mais especificamente redes neurais multivariáveis, ao escopo da física de altas energias. Ele está inserido no ambiente de filtragem *online*, e propõe uma abordagem para trazer uma técnica de classificação já implementada (de maneira bem sucedida) para um nível acima de filtragem, fazendo para isso fusão de dados de outros detectores de maneira a suprir a demanda do experimento por métodos mais eficazes, a qual cresceu com o aumento da geração de eventos. O problema se torna ainda mais complexo quando se levam em conta as exigências de latência (a decisão deve ser tomada em questão de milissegundos) e a fina granularidade

dos detectores (com mais de 1 k sensores), exigindo uma combinação de técnicas de aprendizado de máquina com conhecimento especialista de física de partículas para atender as demandas do projeto.

O experimento ATLAS, no qual este trabalho está inserido, é descrito na Seção 2, onde também se descrevem os detectores (Seção 2.2) e o sistema de filtragem (Seção 2.3). A abordagem proposta é explicada na Seção 3, com a descrição das grandezas empregadas na discriminação (Seção 3.1), da topologia de redes neurais (Seção 3.2) e do método empregado para a concepção, treinamento e escolha das redes (Seção 3.3). Os resultados obtidos são expostos e analisados na Seção 4, e o trabalho é concluído na Seção 5.

## 2 Experimento ATLAS

O CERN [1] é uma organização que visa o estudo da estrutura fundamental da matéria, e usa aceleradores de partículas e detectores construídos especialmente para esse propósito. Dentre os aceleradores, o maior é o LHC [2] (Grande Colisor de Hádrons), sendo o experimento ATLAS [3], o maior dos seus quatro experimentos principais.

Dedicado a uma larga gama de estudos físicos, o ATLAS conta com três sub-detectores, na ordem do mais interno para o mais externo: o Detector Interno ou Detector de Traços (ID), o Sistema de Calorimetria, subdividido em Calorímetro Eletromagnético (ECAL) e Calorímetro Hadrônico (HCAL), e o Espectrômetro de Múons. A ilustração do experimento pode ser observada na Figura 1 abaixo.

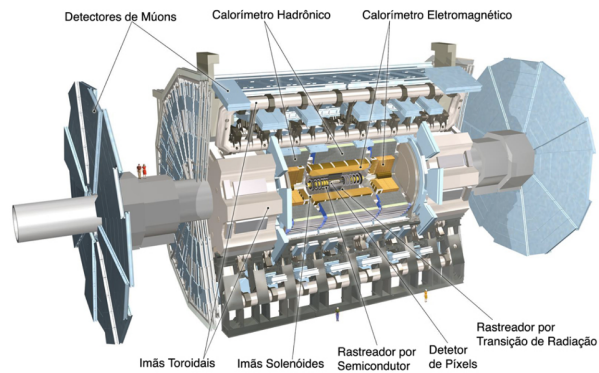


Figura 1: O experimento ATLAS. Adaptado de <https://cds.cern.ch/images>

### 2.1 Detecção de elétrons

A proposta deste trabalho está relacionada à identificação especificamente de partículas de elétrons dentro do experimento ATLAS. Elas são de extrema

importância para a física de partículas por serem um dos resultados finais do decaimento de diversos eventos de interesse.

Muitos fenômenos raros só ocorrem em colisões de alta energia e permanecem neste estado por alguns instantes antes de decaírem em outras partículas mais comuns. É o caso do Bóson de Higgs, por exemplo. Os mecanismos físicos estudados pelo ATLAS normalmente não são detectados diretamente, entretanto fótons e elétrons, que são partículas elementares, aparecem como resultado final de diversos decaimentos, quando a partícula original perde parte de sua energia. Por isso, a partir da detecção destas partículas e da relação que eles têm entre si pode-se vir a identificar uma alta gama de eventos dos quais eles são derivados, remontando assim quais foram os produtos originais provenientes das colisões ocorridas no LHC.

## 2.2 Detectores do ATLAS: Traço e Calorimetria

O Detector de Traços [4] é o mais interno do ATLAS e tem a função de identificar as trajetórias (rastreamento) de partículas carregadas desde a sua colisão até atingirem os calorímetros, de forma que consegue-se o momento das partículas. Ele é subdividido em dois medidores de alta precisão, o Detector de Pixel e o Detector de Rastreamento por Semicondutores (SCT), que exploram a diminuição da resolução de posição (melhorando a resolução do detector), e o Detector de Rastreamento por Transição de Radiação (TRT), que explora medição contínua do traço da partícula de forma a maximizar o comprimento, resultando em melhor resolução de energia total.

O Sistema de Calorimetria [5] é um conjunto de detectores destinados a determinar a energia das partículas, o que é feito a partir da absorção das mesmas pelos calorímetros. Conforme passam pelos detectores, elas vão interagindo com o material e criando uma cadeia contínua de eventos gerando um chuva de partículas. A forma e profundidade dos chuviscos gerados são usadas para determinar o tipo de partícula incidente. O calorímetro deve absorver completamente a energia da partícula e produzir um sinal proporcional à energia depositadas nas suas células.

Os calorímetros são divididos em ECAL e HCAL. O Calorímetro Eletromagnético, ECAL, corresponde ao conjunto de detectores mais internos do sistema, tendo estes a finalidade de absorver toda a energia de partículas eletromagnéticas. Ele foi projetado para resistir a grandes quantidades de radiação e procura absorver as partículas eletromagnéticas em sua totalidade. Já o Calorímetro Hadrônico, HCAL, mais externo, é mais denso para absorver toda a energia restante (que em teoria corresponde à energia de partículas hadrônicas, já que as eletromagnéticas já teriam sido absorvidas), de forma que apenas elementos como múons ou neutrinos seriam capazes de escapar à detecção de energia do ATLAS. O sistema de calorimetria é composto por sete camadas de células de detecção, compondo um total de aproximadamente 100 milhões de canais de leitura. Embora utilizem tecnologias diferentes, ambos possuem o mesmo princípio, eles medem a energia da partícula original através de um processo destrutivo de multiplicação de partículas chamado chuva.

Para a tarefa de identificação de elétrons usam-se apenas os dados de traços e de calorimetria, sem avaliar os sinais provenientes do detector de Múons, um detector projetado para a detecção destas partículas em específico.

### 2.3 Sistema de Filtragem

A detecção de eventos de interesse exige alta eficiência, já que eles são muito raros e a taxa de geração de eventos durante a colisão é muito alta (até 40MHz). Como o LHC tem taxa de colisão muito maior que a sua capacidade armazenamento, faz-se necessário um sistema de filtragem que aplica uma cadeia sequencial de seleção. Tem-se o Primeiro Nível de Filtragem, o L1, que precisa processar os dados da forma mais rápida o possível, de forma a acompanhar a taxa de produção dos mesmos. Para conseguir baixa latência (e alta frequência), essa etapa é implementada em hardware e não é tão discriminante quanto as outras, que vêm em seguida. Ou seja, no intuito de fazer um processamento mais rápido, ela aceita um maior número de hádrons em conjunto com os elétrons, eliminando apenas o suficiente para atender às capacidades de armazenamento e latência requeridos no próximo nível.

Já o próximo nível, Alto Nível de Filtragem, o HLT, é uma seleção em software com latência bem maior. Ele é dividido em duas etapas: a etapa rápida e a precisa. A rápida realiza a pré-seleção a partir da reconstrução simples das informações de trajetória e calorimetria. Já na etapa precisa se tem processos de decisão e reconstrução mais rigorosos, empregando-se algoritmos mais complexos e mais próximos o possível dos usados na análise *offline* (após a passagem pelo sistema de filtragem). Pode-se observar um esquema do funcionamento do sistema de filtragem na Figura 2 abaixo. Este projeto foi desenvolvido como uma alternativa para a identificação de elétrons durante a etapa rápida do HLT, fazendo uma reconstrução rápida tanto de traços quanto de calorimetria e retirando o máximo de hádrons com o mínimo de perda de elétrons no processo.

Percebe-se que os primeiros níveis da filtragem são mais rápidos, mas necessariamente menos discriminantes. É melhor que passem mais sinais que não são elétrons do que as informações dos elétrons em si sejam perdidas. Então se não se pode fazer uma discriminação tão boa, é melhor carregar mais hádrons, mas se certificar que são registrados (quase) todos os elétrons. Afinal, os sinais rejeitados pela filtragem online não serão gravados para futura análise, não poderão ser verificados novamente.

## 3 Estratégia de discriminação

A estratégia de discriminação de elétrons atualmente utilizada é a reconstrução da informação anelada de calorimetria alimentando um discriminador

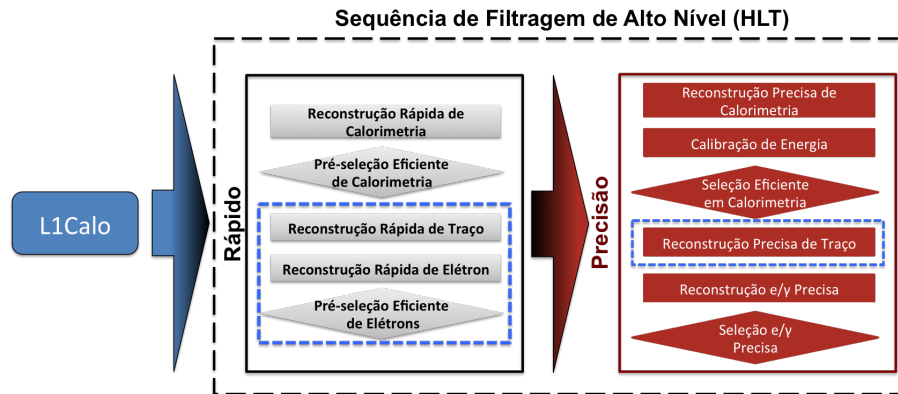


Figura 2: Estrutura do Sistema de Filtragem do ATLAS para eventos com elétrons. A seleção se inicia com o algoritmo do L1Calo, responsável por encontrar regiões de interesse com dados de Calorimetria. Em seguida, aplicam-se cortes altamente discriminantes para a redução da taxa de eventos. Posteriormente opera o HLT, dividido em dois estágios, rápido e preciso, aquele atuando como pré seleção e este aplicando análise o mais similar possível daquela realizada em *offline*

neural, segundo o algoritmo *Ringer*. A proposta utilizada envolve usar os resultados já obtidos por essas redes e juntá-los com uma rede alimentada pelas variáveis do Detector Interno na estratégia de *Expert Neural Networks* ou Redes Neurais Especialistas.

### 3.1 Grandezas físicas empregadas

Para esta fase do projeto, que utiliza eventos simulados para as condições de operação de 2014, possuem-se três variáveis de traço pertinentes à identificação de elétrons:  $\frac{E}{P}$ ,  $\Delta\eta$  clusterizado e  $\Delta\Phi$  clusterizado. A primeira,  $E/P$  nos dá a razão entre a energia e o momento da partícula, a segunda é a diferença em  $\eta$  entre o baricentro do cluster e do traço extrapolado para a primeira camada eletromagnética e a terceira a diferença em entre o baricentro do cluster e do traço extrapolado para a segunda camada eletromagnética.

Além das de traço, usamos um conjunto de 100 grandezas de calorimetria, os anéis de energia provenientes da técnica *Ringer* [6]. Sabendo que o chuvaire de partículas possui geometria cônica, o algoritmo *Ringer* vai explorar esse formato e organizar a informação de calorimetria em anéis de energia. Esses anéis são centrados na célula mais energética da camada e são obtidas as informações de anéis pela soma das energias das células dos vizinhos que compõem uma dada circunferência. Ao final é composto um conjunto de 100 anéis, os quais irão alimentar a rede neural. Os anéis são normalizados antes de servirem como

variáveis de entrada para as redes, sendo todos divididos pelo somatório da energia de todos os anéis, na chamada normalização por *Norma 1*.

Trabalhando com a organização dos sinais de calorimetria em informação anelada nos possibilita armazenar as informações de posição e formato dos chuveiros de partículas, tendo em vista que eles têm formato cônico. Essa descrição mostra alto potencial discriminativo por representar justamente a diferença lateral e longitudinal do chuveiro para o discriminador, que é a principal diferença entre os chuveiros de partículas eletromagnéticas e hadrônicas.

O *Ringer* atua com um *ensemble* de redes neurais que faz 20 partições em posição e energia, e esse total de partições toma a decisão da filtragem *online* baseada em calorimetria. Esses resultados foram utilizados para obter a informação de calorimetria do nosso discriminador, que usará além disso as informações de traços para determinar a sua própria classificação final. No entanto neste trabalho o método foi aplicado a apenas uma das partições, a de posição  $\eta$  menor que 0,8 e de energia  $E_T$  entre 30 GeV e 40 GeV.

### 3.2 Topologia

Foram usadas redes neurais MLP com uma camada escondida e um neurônio na camada de saída. Inicialmente foram treinadas redes separadas para sinais de traço e de calorimetria, de forma que cada uma se torna especialista em extrair as características mais discriminantes das informações de cada detector. As redes de calorimetria usadas foram as do próprio *Ringer*, possuindo 100 neurônios na camada de entrada (um para cada anel) e 13 neurônios na camada escondida. Como já foi explicado os anéis de entrada foram normalizados com *Norma 1*. As redes de traço foram montadas com três neurônios na camada de entrada e dois na camada escondida.

Treinadas as redes especialistas, são retiradas as camadas de saída de cada uma e suas camadas escondidas alimentam uma terceira rede neural, a rede global. Esta rede é montada com cinco neurônios na camada escondida e é responsável por combinar as informações das redes anteriores e tomar a decisão final de discriminação [7]. Dividindo a análise do conjunto de dados em duas sub-tarefas pode-se examinar a informação disponível de maneira otimizada, em teoria conseguindo um sistema mais discriminante e aumentando a eficiência do mesmo.

### 3.3 Método

As redes foram treinadas para obter o máximo de separação entre elétrons e hádrons (mínimo erro quadrático), com o objetivo de gerar saída de valor +1 para sinais de elétrons e -1 para sinais hadrônicos. O método de treinamento utilizado foi o *RMSprop* (*Root Mean Squared Propagation*), disponível na infraestrutura de redes neurais Keras [8]. O treinamento não foi limitado por número de épocas, ao invés disso é finalizado quando o erro médio quadrático do conjunto de validação fica crescente durante 50 épocas consecutivas. Utiliza-se esse

método de *callbacks* para evitar que a rede caia em mínimos locais, uma vez que a otimização dos pesos da rede neural não se trata de um problema de otimização convexa. Ainda para evitar a escolha de mínimos locais pelo treinamento o treino de cada rede é realizado com 100 diferentes inicializações. Cada rede foi treinada independentemente para obter os valores objetivos de saída, e ao final da fase de treino são escolhidas dentre as inicializações aquelas redes de maior índice SP para cada detector, este dado pela fórmula:

$$SP(\%) = 100\% \times \sqrt{\sqrt{P_D \cdot (1 - FA)} \times \left( \frac{P_D + (1 - FA)}{2} \right)} \quad (1)$$

onde  $P_D$  é a probabilidade de detecção de elétrons e  $FA$  é o falso alarme, ou seja, a probabilidade de se identificar hádrons como elétrons. Quanto maior o índice SP, maior a eficiência da discriminação, uma vez que este é um parâmetro que liga as ambas a detecção e o falso alarme, fornecendo um compromisso entre os mesmos.

O conjunto de dados foi subdividido em 10 conjuntos, denominados pacotes, para a validação cruzada segundo o método *Jack-Knife* [9]. Foram utilizados nove pacotes para treino e um para teste, com rotatividade do pacote de validação, gerando um total de 10 configurações de treinamento, denominados sorteios. Dessa forma foram realizados um total de 1.000 treinamentos para cada rede, oriundos das 100 inicializações para cada pacote. Não foi utilizado um pacote de validação porque os dados utilizados já são de simulação e a validação será feita com um novo conjunto de dados, esses oriundos de colisão de partículas de fato.

O treinamento foi realizado primeiro para as redes especialistas. Escolhidas as redes a serem utilizadas, para cada pacote foram retiradas as camadas de saída de cada uma e seus neurônios da camada escondida foram conectados à entrada da rede global (daquele mesmo pacote). As redes especialistas foram congeladas durante o treinamento final, pois considera-se que elas já haviam sido otimizadas para a melhor extração de suas características discriminantes durante a primeira fase de treinamento. No fim a rede final é novamente escolhida pela que apresentou o maior índice SP.

As variáveis alimentadas às redes são normalizadas em partes. As de calorimetria com a *Norma 1*, já mencionada e descrita pela fórmula:

$$x_i = \frac{x_i}{\sum x} \quad (2)$$

Enquanto as de traços foram normalizadas dividindo cada uma pelo dobro do desvio padrão da mesma obtido no conjunto de teste.

$$x_i = \frac{x_i}{2\sigma_{x_i}} \quad (3)$$

## 4 Resultados

A estratégia foi testada em dados com  $E_T$  (energia da partícula) entre 30 GeV e 40 GeV e de  $\eta$  (ângulo com relação ao eixo de colisão) entre 0 e 0,8. O conjunto total de dados conta com aproximadamente 20.000 sinais de elétrons e 2.500 de hádrons, obtidos da simulação de Monte Carlo de *ATLAS Collaboration, 2014*. É um conjunto que representa as condições esperadas para a operação no nível de filtragem rápida do HLT, e os eventos foram distribuídos entre os pacotes da validação cruzada procurando manter a variação estatística dos dados.

Na Figura 3 abaixo pode-se ver gráficos de comparação entre as saídas do discriminador neural especialista desenvolvido e as da rede de calorimetria do *Ringer*. Fazendo a figura de uma saída contra a outra se consegue verificar onde há perdas e ganhos com o acréscimo das informações de calorimetria, e assim verificar qualitativamente qual a agregação de valor desses dados.

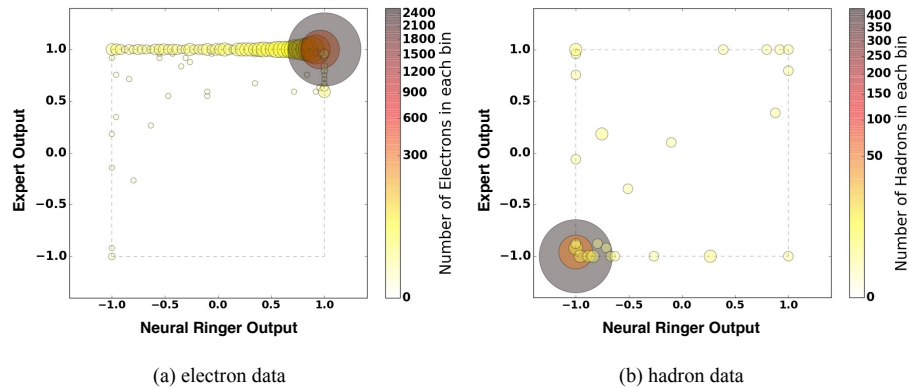


Figura 3: Saídas das redes de calorimetria *Ringer* – no eixo das abcissas – contra as saídas das redes neurais especialistas – eixo das ordenadas – para sinais de elétrons (a) e de hádrons (b),. Usadas as redes treinadas para o 9º sorteio da validação cruzada. As escalas dão o número de sinais no intervalo de valores de saída no qual estão centralizados.

No gráfico 3.a, de sinais eletrônicos, verifica-se que as redes especialistas conseguem fixar no alvo um conjunto de dados que antes estavam distribuídos. Embora haja perda uma piora relativa de alguns casos, há uma clara concentração de eventos no alvo em relação à rede só de calorimetria. O mesmo fenômeno



pode ser observado na Figura 3.b, embora menos evidente e de maneira menos pronunciada. Isso indica que as redes de referência falham na classificação de um conjunto de elétrons que as redes especialistas conseguem classificar bem. Como as redes especialistas analisam mais informação do que as redes de referência, é possível que essa melhora seja resultado da análise dos dados de traço, que possibilitariam a diferenciação de alguns casos que ficariam duvidosos quando só se avalia os dados de calorimetria.

Uma grande vantagem que esse tipo de espectro apresenta é a possibilidade de escolher valores de corte mais frouxos (não tão próximos de  $-1$ ) sem haver grandes perdas na eficiência da detecção de elétrons, que é o grande objetivo de uma rede treinada para operar em *online* no ATLAS. O objetivo é que não se deixe de salvar para a discriminação *offline* eventos que possam vir a ser de interesse. Por sua vez um corte de valor numérico maior nos gera a possibilidade de diminuir os índices de falso alarme da discriminação, pois não são considerados como elétrons o conjunto de hádrons que ficaria entre o corte mais rígido e o mais frouxo.

Na Tabela 1 podem ser observadas as comparações dos valores numéricos entre as eficiências das duas propostas envolvidas. Percebe-se que, como esperado pelos gráficos da Figura 3, de fato houve melhora nas eficiências com o emprego das redes especialistas propostas. É notável um aumento na detecção de elétrons e a diminuição dos níveis de falso alarme, uma boa perspectiva para a avaliação da proposta. Há um aumento de mais de um ponto percentual no índice SP da validação cruzada, que era a grandeza que se procurava maximizar, demonstrando uma melhora sistemática na eficiência.

Outro fator importante de se observar é o desvio padrão das eficiências obtidas. Embora não se possuíssem os desvios da rede de referência, apenas suas eficiências, pode ser observado que, para o índice SP, o desvio padrão foi de  $\approx 0.17\%$  em relação à medição de SP obtido, que é uma porcentagem pequena do valor total.

Tabela 1: Comparação das eficiências das redes especialistas com a referência

	Detecção - $P_D$ (%)	Falso Alarme - $FA$ (%)	Índice SP (%)	Corte
Alternativa	$99,55 \pm 0,18$	$1,7 \pm 0,3$	$98,9 \pm 0,2$	0,76367
Referência	98,33	3,2	97,6	0,645646

Embora designado para operação *online* não foi feita aqui uma análise acerca da latência do método utilizando redes especialistas porque a mesma já foi testada para o *Ringer*, e este supre todos os requisitos de tempo de execução. Como o acréscimo dos dados do Detector Interno implica apenas em mais poucas camadas de redes neurais, que possuem baixa latência, essa questão foi desconsiderada quando se avaliou a viabilidade do método proposto.

## 5 Conclusão

Neste trabalho foi explicitada uma proposta desenvolvida para realizar a discriminação de elétrons atuando na etapa rápida do alto nível de filtragem do ATLAS, HLT, baseado na reconstrução de ambos as variáveis do Detector Interno e a informação anelada do Sistema de Calorimetria. O discriminador neural implementado foi avaliado com relação às redes neurais de referência da equipe sobre um conjunto de dados de simulação das condições esperadas de atuação. Durante a validação da proposta é observado um aumento na eficiência de seleção em relação à referência, que utiliza apenas dados dos calorímetros, demonstrando a aptidão do sistema baseado em redes neurais especialistas em desempenhar a tarefa para a qual foi projetado.

Embora tenha sido desenvolvido para a etapa rápida de filtragem do HLT, o método pode ser implementado para outras etapas de identificação de elétrons do ATLAS, como a etapa precisa do HLT e a operação *offline*, desde que os métodos de reconstrução dos sinais de traço e calorimetria sejam ajustados adequadamente.

Para a melhoria do discriminador aqui proposto e analisado pretende-se fazer uma análise das possíveis normalizações das variáveis de traço, para facilitar a extração das informações por esta rede especialista e com isso o aumento da eficiência dela e do sistema como um todo. Além disso pretende-se expandir o treinamento de redes especialistas para as outras partições de posição e energia e se realizar o *ensemble* de todas as redes, a partir do qual serão comparadas as eficiências com uma discriminação usando as redes de calorimetria do *Ringier* com o algoritmo de *Likelihood* [10].

Em desenvolvimento agora estão a modificação do projeto para expandir o número de variáveis de traço utilizadas, uma vez que os novos dados (a partir de 2016) contam com novas variáveis devido a mudanças técnicas no Detector de Traços do ATLAS. Além disso o projeto está em fase de exportação para ser utilizado na discriminação *offline*, onde os desafios envolvidos são outros, já que necessita-se da melhor reconstrução o possível da física envolvida, mesmo que a latência do algoritmo seja alta.

## Agradecimentos

Essa pesquisa foi desenvolvida com o apoio de CNPq, CAPES, RENAFAE (MCTI), FAPERJ e do Núcleo Avançado de Computação de Alto Desempenho (NACAD) da COPPE, Universidade Federal do Rio de Janeiro (UFRJ); bem como da União Européia, que apoiou o trabalho através do projeto E-PLANET. Os autores agradecem também à colaboração do Experimento ATLAS pelas frutíferas discussões a respeito deste trabalho.

## Referências

1. *Portal oficial do CERN*. Acessado em 05 de Março de 2017. URL: <http://cern.ch/>.
2. Lyndon Evans e Philip Bryant. “LHC Machine”. Em: *JINST* 3 (2008). Ed. por Lyndon Evans, S08001. DOI: 10.1088/1748-0221/3/08/S08001.
3. ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. Em: *Journal of Instrumentation* 3.08 (2008), S08003. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08003>.
4. ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997.
5. *ATLAS Calorimeter Performance: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996.
6. ATLAS Collaboration. “Ring-shaped Calorimetry Information for a Neutral Egamma Identification with ATLAS Detector”. Em: *Journal of Physics: Conference Series* 762.1 (2017), p. 012049. URL: <http://stacks.iop.org/1742-6596/762/i=1/a=012049>.
7. F. W. Soares, D. O. Damazio e J. M. Seixas. “Particle discrimination using matched filters and expert neural networks”. Em: *Electronics, Circuits and Systems, 1999. Proceedings of ICECS '99. The 6th IEEE International Conference on*. Vol. 1. 1999, 595–598 vol.1. DOI: 10.1109/ICECS.1999.812355.
8. François Chollet. *Keras*. <https://github.com/fchollet/keras>. 2015.
9. H. Abdi e L.J. Williams. “JackKnife”. Em: *Encyclopedia of Research Design*. Thousand Oaks (CA): Sage., 2010, pp. 655–660.
10. Andreas Hocker et al. “TMVA - Toolkit for Multivariate Data Analysis”. Em: *PoS ACAT* (2007), p. 040. arXiv: physics/0703039 [PHYSICS].