
UM CLASSIFICADOR NEURONAL COMPACTO E EFICIENTE COM CAPACIDADE DE IDENTIFICAR CONTAMINAÇÃO EM DADOS EXPERIMENTAIS

Denis Oliveira Damazio*
damazio@lps.ufrj.br

José Manoel de Seixas*
seixas@lps.ufrj.br

A. C. Soares*
cecilia@lps.ufrj.br

*Laboratório de Processamento de Sinais, COPPE/EE/CAP, Universidade Federal do Rio de Janeiro
CP 68504, 21945-970, Rio de Janeiro, Brasil

ABSTRACT

A neural classifying system is developed to identify three particle classes in experimental high-energy physics. The system makes use of the extraction of principal discriminating components to obtain compactness and high classification efficiency, even identifying outsiders in experimental data sets. More than 97% of analysed events are correctly classified.

KEYWORDS: Neural Networks, Pattern Recognition, Preprocessing.

RESUMO

Um sistema classificador neuronal é desenvolvido para identificar três classes de partículas em física experimental de altas energias. O sistema usa a extração de componentes principais de discriminação para combinar compactidade e alta eficiência de classificação, identificando, inclusive, a contaminação presente nos dados experimentais. Mais de 97% dos eventos analisados são corretamente classificados.

PALAVRAS-CHAVE: Redes Neurais, Reconhecimento de Padrões, Pré-processamento.

1 INTRODUÇÃO

Na física de altas energias, os problemas estudados fornecem base para desafios interessantes na área de engenharia e automação de processos. No presente caso, um problema de reconhecimento *online* de padrões, associado ao desenvolvimento de um detector de partículas, nos leva à um estudo prático de um sistema baseado em redes neuronais. Extraíndo-se as características mais importantes da informação provida do detector, espera-se que se possa reconhecer os diferentes tipos de partículas envolvidas. No caso do reconhecimento *online* de padrões, não só uma eficiência elevada é desejada, mas também existem requisitos severos de velocidade, de forma a se manter o processamento de modo sustentável. Ainda mais, em física de altas energias os detectores são altamente segmentados, gerando uma informação de dimensionalidade elevada. Desta maneira, métodos de pré-processamento das informações brutas recebidas do detector devem compactar a informação de forma eficiente, favorecendo a diminuição do tempo de processamento e mantendo elevada a eficiência de discriminação. Neste artigo, fazemos uso da técnica de componentes principais de discriminação (Caloba et al., 1995) para atingir tal objetivo.

Artigo submetido em 20/12/2000

1a. Revisão em 15/10/2001; 2a. Revisão 18/2/2002;

3a. Revisão 8/7/2003

Aceito sob recomendação dos Eds. Associados Profs. Fernando Gomide e Takashi Yoneyama

O problema de discriminação de partículas a ser discutido envolve a informação fornecida por um calorímetro de altas energias, um detector importante que mede a energia das partículas que com ele interagem no experimento. Na prática, protótipos de uma dada técnica de calorimetria são testados em feixes experimentais de partículas. Apesar da alta qualidade que se obtém hoje em dia com tais feixes, é inevitável a presença de um nível, em nada desprezível, de contaminação do feixe. Assim, o problema de classificação de partículas se torna mais complexo.

Neste trabalho, utilizam-se redes neuronais artificiais e a análise de componentes de discriminação para se identificar as partículas e reconhecer os eventos de contaminação do feixe experimental. Na próxima seção, apresentaremos uma introdução ao ambiente de física experimental de altas energias, com ênfase na descrição do experimento LHC (Large Hadron Collider) que, atualmente, se encontra em construção no CERN (Laboratório Europeu para a Física de Partículas), localizado na Suíça. Na sequência, apresentaremos o calorímetro Tilecal, para o qual o sistema discriminador está sendo desenvolvido. Na quarta seção discutiremos a análise de componentes de discriminação e na quinta mostraremos os resultados obtidos. Por fim, a sexta seção traz as conclusões do trabalho.

2 FÍSICA DE ALTAS ENERGIAS

Os físicos experimentais de altas energias vêm desenvolvendo experimentos que tentam reproduzir condições semelhantes àquelas que teriam ocorrido nos primeiros instantes após o Big Bang, a grande explosão que teria dado origem ao Universo. Tais condições podem ser resumidas em se concentrar quantidades cada vez maiores de energia em volumes cada vez menores. A melhor forma para se reproduzir estas condições é colidir partículas aceleradas, logo, com grande quantidade de energia cinética. Tais colisões dispersam a energia acumulada no processo de aceleração na pequena área a volta das partículas gerando eventos que podem fazer parte de uma nova física. Ao redor destes pontos de colisão, um conjunto complexo de detectores é então colocado, para permitir uma leitura eletrônica das partículas resultantes, na tentativa de reconhecer esta nova física.

O LHC consiste de um complexo acelerador de partículas, com capacidade para produzir o choque entre duas nuvens de 10^{11} prótons, que têm uma energia de 14 TeV (Tera-elétron-Volts) no centro de massa (CERN, 1999). Tal acelerador se situa na região da fronteira Franco-Suíça e ocupa um túnel de 27 Km de comprimento, a

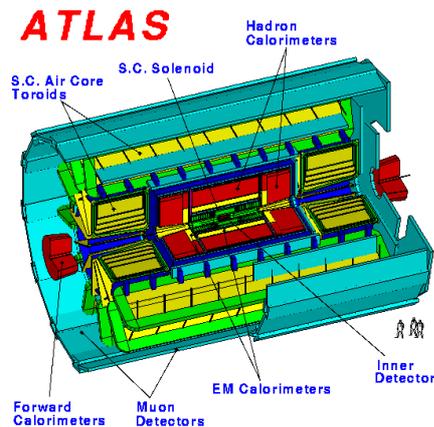


Figura 1: Estrutura do laboratório ATLAS.

cerca de 100 m abaixo do solo, utilizando técnicas de supercondutividade para gerar os poderosos campos magnéticos necessários para manter os seus dois feixes de partículas circulando em sentidos contrários e à velocidades próximas a da luz até os pontos de colisão.

A Figura 1 mostra um dos conjuntos de detectores que será colocado ao redor de um dos pontos de impacto do LHC, o laboratório ATLAS (The ATLAS Collaboration, 1999a). Este conjunto de detectores formará, como se pode ver na figura, um barril a volta do ponto de impacto, sendo que a colisão ocorrerá no meio do mesmo e os seus produtos serão irradiados de dentro para fora do barril, passando por cada um dos detectores. Tais detectores são especializados em detectar diferentes tipos de partícula, e realizar medidas como energia e ângulo de incidência.

Dentre os detectores, se destacam os calorímetros (Wigmans, 1991). Calorímetros são detectores de partículas segmentados (tipicamente, milhares de células de leitura) que medem, com grande precisão, a energia das partículas que com ele interagem, através da absorção total desta energia. Para otimizar esta detecção, encontramos calorímetros especializados em partículas eletromagnéticas (como elétrons) e outros especializados em partículas hadrônicas (por exemplo, prótons, píons). Como cada classe de partículas apresenta uma deposição característica de sua energia sobre as células do calorímetro, a informação obtida pelos calorímetros nos permite discriminar a qual classe pertence determinada partícula incidente.

A Figura 2 mostra o arranjo em módulos do calorímetro Tilecal, que vem a ser o calorímetro hadrônico do ATLAS (The ATLAS Collaboration, 1999b) e alvo do

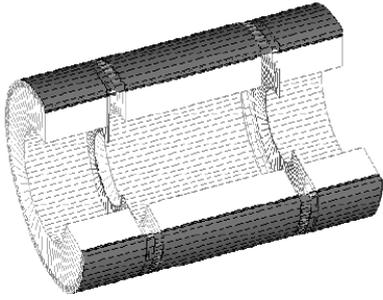


Figura 2: O calorímetro Tilecal com seus blocos detectores do barril e do barril estendido.

presente trabalho. Este calorímetro é formado por dois blocos detectores: área central (barril, com 64 módulos) e área estendida (barril estendido, com duas seções, 64 módulos cada). No total, o Tilecal requer a leitura de cerca de 10.000 canais de medição. Para desenvolvimento do detector, se utilizam feixes de partículas que são lançados contra o protótipo, simulando a situação real de detecção dos produtos das colisões do LHC.

Neste trabalho, utilizamos a informação produzida por duas seções do barril estendido do Tilecal, a partir de testes com feixes de diferentes energias.

Para que possamos melhor compreender o trabalho realizado pela rede neuronal, vamos, na próxima seção, estudar um pouco o que acontece durante a realização destes testes com feixes de partículas, analisando a resposta do detector a este evento.

3 O CALORÍMETRO

O TileCal é formado por telhas de material cintilante, colocadas entre camadas de ferro. As partículas, ao interagirem com o ferro, considerado material passivo, sofrem decaimentos em outras partículas, o que resulta numa dispersão de energia na região onde este processo de decaimento ocorre. Estas novas partículas voltam a colidir com átomos de ferro, gerando mais dispersão de energia e mais partículas. Este processo se repete até que toda a energia da partícula inicial seja dispersa na estrutura do calorímetro, formando uma cascata de partículas. O material dito ativo (telhas cintilantes) emite luz na presença da irradiação de energia, gerando uma quantidade de luz proporcional à energia dispersada. Desta maneira, pode-se ter acesso à informação sobre quanto de energia foi dispersada numa dada área do calorímetro. Esta luz é levada para o exterior do detector através de fibras óticas, que são agrupadas pela região do calorímetro a que atendem, formando células na superfície do detector. Duas leituras, uma para cada

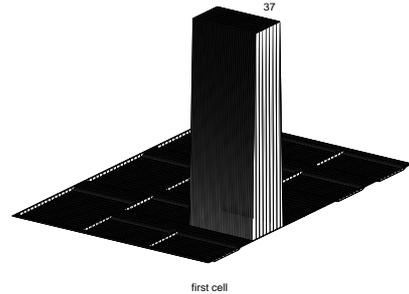


Figura 3: Típico evento de elétrons no módulo do barril estendido.

superfície livre de cada célula (superior e inferior), são obtidas por evento. Estas células representam segmentações radiais e longitudinais do calorímetro, que visam permitir um detalhamento fino do perfil de deposição da energia das partículas no detector. O sinal luminoso produzido em cada célula é transmitido para tubos fotomultiplicadores de alta eficiência, os quais transformam em sinal elétrico a informação luminosa, fornecendo o valor de energia depositada. Cada um dos módulos das extensões do barril produz 28 canais de leitura associados a leitura dos sinais luminosos provindos de cada uma de suas 14 células.

O perfil de deposição de energia de partículas em calorímetros é claramente dependente da partícula que o originou. Para podermos verificar isso, vejamos a Figura 3, que mostra um padrão típico de deposição de energia para elétrons num módulo do barril estendido. Nesta figura, podemos ver uma imagem em que cada torre tem sua altura proporcional ao valor de energia depositado na célula correspondente a uma área no calorímetro (base da torre). Em geral, um elétron ao incidir no calorímetro perde quase toda sua energia na primeira célula por ele atingida, formando um perfil de pequena penetração na estrutura do detector. Na figura, este fato pode ser verificado pela grande presença de energia (uma torre muito alta) em apenas uma célula e praticamente nenhuma energia no restante das células do detector.

Ao contrário dos elétrons, os píons tendem a depositar sua energia após terem penetrado uma certa distância no detector (ver Figura 4). Também o perfil de deposição pode ser considerado mais disperso, posto que tipicamente mais células a volta do ponto de impacto são sensibilizadas.

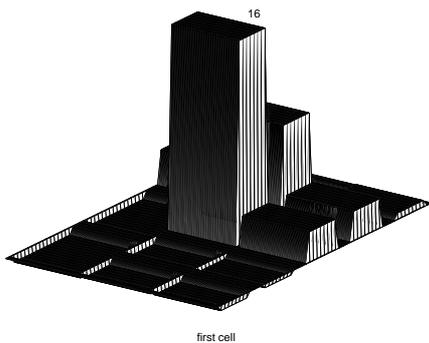


Figura 4: Típico evento de píons no módulo do barril estendido.

Através destes exemplos, fica clara a possibilidade de se usar o calorímetro para distinguir partículas, explorando-se a sua informação de fina granularidade. Assim, o problema de discriminação de partículas pode ser enfocado como um problema de reconhecimento de padrões, para o qual redes neurais vêm se mostrando altamente eficientes.

A Figura 5 mostra o arranjo experimental que foi utilizado para testarem-se dois módulos protótipos da última geração para o barril estendido. Cinco módulos de uma geração anterior de protótipos do Tilecal foram também utilizados neste arranjo experimental (dois módulos acima e três módulos abaixo), com vistas a compensar o esperado vazamento de energia, uma vez que os dois módulos em teste são incapazes de, por si só, absorver completamente a energia das partículas incidentes. De fato, como já comentado, cada seção do barril estendido se completa com 64 módulos idênticos aos dois aqui estudados, sendo um só módulo muito estreito para conter toda a cascata.

Um total de 1242 eventos de elétrons, 1631 píons e 1168 múons foram adquiridos, utilizando-se um feixe de partículas de 100 GeV incidente no centro da junção dos dois protótipos principais, de modo a se obter máxima absorção de energia. Um total de 256 canais de leitura foram digitalizados para cada evento (56 para os módulos do barril estendido e 200 dos módulos auxiliares), definindo com precisão o perfil de deposição de energia para cada classe de partícula.

Apesar da alta qualidade do feixe de partículas utilizado experimentalmente, é inevitável algum nível de contaminação do mesmo. Assim, durante a produção de píons, o acelerador para testes acaba também gerando múons, o que mascara o desempenho do calorímetro para partícu-

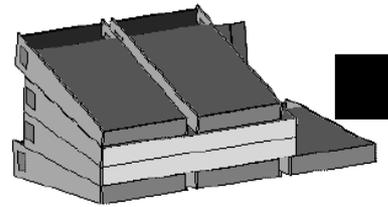


Figura 5: Arranjo experimental para a coleta de dados do Tilecal.

las hadrônicas como os píons. Ainda pior, em testes com elétrons, o acelerador, além destes, gera píons e múons de contaminação, distorcendo qualquer estudo da resposta do calorímetro para partículas eletromagnéticas. Apenas o feixe de múons costuma ser, na prática, livre de contaminação significativa. Para as altas energias envolvidas no problema aqui estudado (100 GeVs), a contaminação por múons não chega a ser um problema, já que os mesmos costumam depositar pouca energia no detector (de 1 a 10 GeVs), possibilitando uma retirada desta contaminação através de um simples corte no valor da energia total depositada no calorímetro. Assim, neste trabalho, lidaremos apenas com a contaminação de píons no feixe de elétrons.

Devido à contaminação, a tarefa de discriminação de partículas se torna bastante complexa na prática. Como poderemos observar, no entanto, uma rede neuronal, durante seu treinamento, é capaz de identificar a contaminação para cada tipo de feixe, de tal modo que, retendo-se apenas os eventos válidos para cada classe de partículas, se pode medir mais acuradamente a correspondente resposta do calorímetro a cada um dos tipos de partículas.

4 COMPONENTES DE DISCRIMINAÇÃO

Para identificar a contaminação do feixe de partículas e discriminar os eventos pertencentes a cada classe, utilizou-se uma rede neuronal multicamadas, sem realimentação, com neurônios do tipo tangente hiperbólica e treinada com o algoritmo de *backpropagation* (Hertz et al., 1990).

Esta rede apresenta 3 neurônios de saída, de tal modo que cada neurônio fica associado a uma única classe. Para o treinamento supervisionado da rede, considera-se um dado feixe de partículas, independente de haver ou não contaminação, e associa-se um vetor alvo dos eventos deste feixe, que contém 1 para o neurônio associado à classe do respectivo feixe e -1 para os demais.

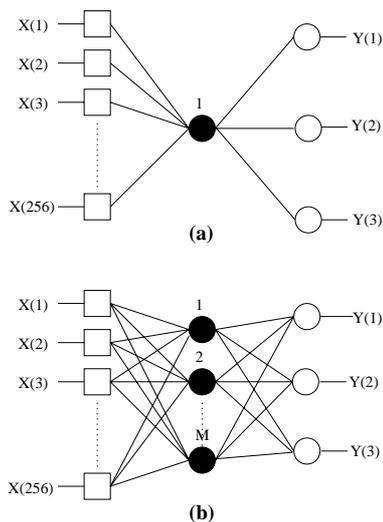


Figura 6: Extração da primeira (a) e da m-ésima componente principal de discriminação(b).

Assim, por exemplo, no caso do feixe ser de píons e, respectivamente, os neurônios de saída estarem associados às classes de elétrons, píons e múons, os vetores alvo para os eventos produzidos por este feixe serão iguais a $(-1,1,-1)$.

Para formar o conjunto de treinamento, tomou-se a metade dos eventos de cada um dos conjuntos de dados adquiridos para cada classe de feixe de partículas, usando-se a outra metade para formar o conjunto de teste que avalia a generalização da rede e, assim, a real performance do classificador em operação prática. Deve-se também ressaltar que a informação de energia depositada em cada célula foi normalizada pela energia total depositada no calorímetro, antes de alimentar os nós de entrada da rede neuronal.

O projeto da rede neuronal classificadora envolveu a extração das componentes principais de discriminação. Estas componentes, ao contrário das componentes principais (Oja, 1995) não visam descrever o conjunto de dados da melhor forma possível, mas sim realizar, de forma otimizada, a discriminação dos conjuntos de dados. Assim, esta extração visa identificar as direções ortogonais privilegiadas do espaço original de dados de entrada que sejam capazes de revelar mais intensamente as sutis diferenças entre as classes de partículas que se deseja identificar, mesmo em condições adversas de contaminação. Tipicamente, apenas umas poucas componentes precisam ser extraídas, para que se possa atingir uma elevada eficiência de classificação, de tal modo que se obtém uma rede neuronal compacta e eficiente (Caloba et al., 1995).

A Figura 6 mostra como é feita a extração destas componentes. Esta topologia permite que a extração das componentes principais seja feita de forma integrada com o projeto do discriminador, de tal modo que uma única rede realiza a extração das componentes e classifica os padrões baseada na informação projetada nestas mesmas componentes. A primeira componente é obtida treinando-se uma rede neuronal através de backpropagation (outros métodos como quickprop podem ser usados), com um único neurônio na camada escondida, para que se atinja a melhor eficiência possível de classificação para as três classes de partículas. Aí se encontra outra diferença com relação às componentes principais, já, que no caso destas últimas a figura de mérito para o treinamento é o erro entre a saída e a entrada (que se quer que sejam iguais), enquanto o estudo das componentes principais de discriminação usa a média geométrica das eficiências de cada classe como figura de mérito do treinamento. Ao final do treinamento, a primeira componente se encontra revelada no vetor de pesos que conecta a camada de entrada à camada escondida da rede.

A extração das demais componentes segue uma mesma seqüência. Primeiramente, adiciona-se um neurônio à camada escondida, treinando-se a rede neuronal novamente para que se obtenha a melhor eficiência de classificação. Durante a fase de treinamento, apenas o vetor de pesos que liga a camada de entrada ao neurônio que foi adicionado à rede é atualizado, haja vista que os demais vetores de peso que ligam a camada de entrada à camada escondida da rede representam componentes principais extraídas nos passos anteriores e, portanto, são mantidos fixos ao longo do processo de extração desta nova componente. O vetor de pesos que conecta a camada de saída à camada escondida também é treinado cada vez que se adiciona uma nova componente, pois deseja-se que a rede possa combinar otimamente a informação das componentes que se tornam disponíveis a cada novo passo da extração. Continuam-se adicionando neurônios (componentes de discriminação) à camada escondida da rede classificadora até que a adição de uma nova componente não represente uma melhora significativa na eficiência de classificação.

Como a rede encontrada deve ser utilizada de forma *online*, a rede definida pela extração de componentes principais de discriminação é também extremamente vantajosa, já que sua implementação *online* pode ser realizada com os mesmos códigos computacionais existentes para redes multi-camadas comuns. Isso leva a uma simplicidade muito grande, quando comparamos com estruturas mais complexas geradas, por exemplo, por métodos correlacionais contrutivos (Fahlman e Lebiere, 1991).

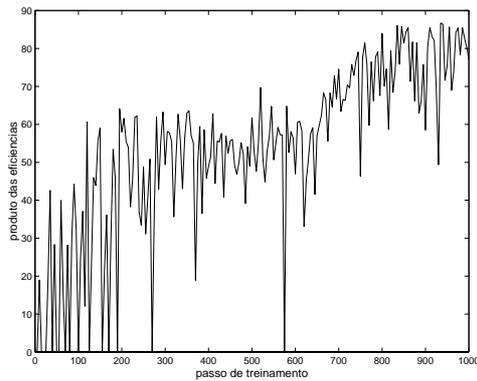


Figura 7: Evolução da média geométrica das eficiências das três partículas no conjunto de treino

5 RESULTADOS

Para extração das componentes principais de discriminação, é importante que se atinja um alto nível de eficiência em cada nova componente que esteja sendo extraída. Assim, devemos analisar a curva de evolução das eficiências de discriminação do nosso problema ao longo do processo de extração de componentes. Durante a extração da primeira componente, um problema, entretanto, pode surgir. Dado o fato de que existe uma contaminação de píons no conjunto de elétrons, a eficiência para esta última partícula deve ser baixa. No caso estudado, métodos clássicos de análise apontam para uma contaminação de cerca de 27%, logo a eficiência na classificação de elétrons não deveria passar de 73%. Devido ao elevado nível de contaminação, a rede composta por um único neurônio na camada escondida pode não identificar plenamente a contaminação do feixe. Isso é ruim, pois se extrairmos a primeira componente erradamente, a extração das demais componentes fica comprometida.

Para contornar este problema, basta não incluir a eficiência na classificação de elétrons na figura de mérito do treinamento para a primeira componente principal. A partir da segunda componente, a extração se desenvolve da forma já descrita, pois bastam dois neurônios na camada escondida para que a contaminação do feixe de elétrons seja claramente revelada.

Exploremos como este processo é implementado na prática. Na Figura 7 temos a evolução da média geométrica das eficiências da rede por passo de treino. Podemos notar que as eficiências sobem até um determinado patamar acima dos 85% de acerto, o que, dada a existência da contaminação, é um resultado não muito satisfatório. Se, entretanto, verificarmos qual a melhor eficiência nos baseando apenas no produto das duas últimas, obtemos um ponto em que a eficiência de elétrons fica em 78,8%,

componente	el	pi	mu	tot
1 ^a	78,8%	63,0%	99,5%	78,4
2 ^a	72,6%	97,1%	99,2%	90,2
3 ^a	72,3%	97,2%	99,2%	90,1
4 ^a	72,5%	97,4%	99,2%	90,3

Tabela 1: Tabela da evolução das eficiências quando da extração das componentes.

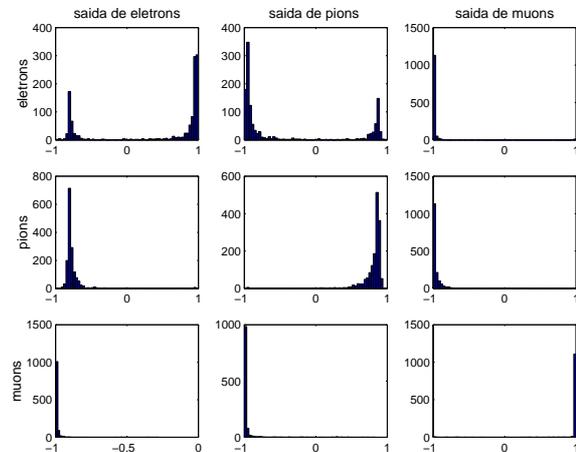


Figura 8: Saídas (colunas) da rede neuronal para os três conjuntos de dados (linhas).

portanto já perto do valor que sabemos ser o ponto correto.

Seguindo o procedimento já descrito para extração das demais componentes, fomos retirando até a quarta componente. Na Tabela 1, podemos ver a evolução da eficiência do classificador neuronal para cada classe de partícula, elétrons (el), píons (pi) e múons (mu) e também a evolução da média das eficiências ponderadas pelo número de eventos. Nota-se que, a partir da segunda componente, já se obtém uma eficiência de classificação semelhante àquela obtida para quatro componentes, com um alto nível de eficiência nos conjuntos de píons e múons. Nota-se também que a eficiência de elétrons acaba por se estabilizar no ponto previsto teoricamente. O fato de extraírem-se mais de duas componentes se justifica na robustez desejada para a identificação de contaminação, permitindo uma análise estatisticamente não tendenciosa da performance do detector para a classe de partículas contaminada. Por outro lado, pode-se observar a capacidade de compactação que a análise de componentes de discriminação obtém, extraindo apenas quatro componentes para realizar a classificação dos padrões em meio à elevada contaminação observada no feixe de elétrons.

A saída do classificador, quando se consideram quatro componentes de discriminação, pode ser observada na Figura 8. Nesta figura, podemos ver uma matriz de histogramas que traz nas suas colunas as três saídas do classificador neuronal. A primeira coluna traz a saída de elétrons, a segunda de píons e a última traz a saída de múons. Cada linha desta figura se relaciona com um dos conjuntos de dados utilizados. De novo, temos elétrons, píons e múons em seqüência, de cima para baixo. Assim, temos, por exemplo, no histograma mais ao alto à esquerda, a saída associada à classe de elétrons para eventos provenientes do feixe (contaminado) de elétrons. À direita desta última, temos a saída de píons para o conjunto de eventos provenientes do feixe de elétrons e assim por diante.

Uma vez que uma saída ativa se encontra associada ao valor +1, e uma saída inativa se encontra associada ao valor -1, para que as distribuições da Figura 8 traduzam um comportamento ótimo do classificador, apenas as distribuições da diagonal principal devem se encontrar próximas de +1. Por seu turno, as distribuições fora da diagonal principal devem ter distribuições próximas de -1. Entretanto, quando se considera que há contaminação significativa no feixe de elétrons, deve-se observar que uma certa quantidade de eventos venha a ser rejeitada pela saída associada aos elétrons, enquanto que a saída associada aos píons, para os eventos do feixe de elétrons, deve identificar tais píons de contaminação. Assim, a Figura 8 mostra que a contaminação no feixe de elétrons faz aparecer um segundo pico, próximo de -1, na distribuição da saída associada à classe de elétrons para os eventos do feixe de elétrons, indicando que estes eventos não pertencem ao conjunto de elétrons. De modo similar, a saída de píons também indica claramente a contaminação do feixe. A Figura 9 mostra que há correlação entre as duas saídas, de tal maneira que os mesmos eventos rejeitados pela saída de elétrons são, em geral, identificados como píons, pela saída associada à classe de píons. Note-se também (ver Figura 8) que a saída associada aos múons não sofre qualquer efeito relativo à contaminação. Da mesma maneira, para eventos do feixe de múons, o classificador neuronal não é confundido pela presença de contaminação no feixe de elétrons, classificando corretamente os múons.

Esta análise mostra que a rede desenvolvida com treinamento *backpropagation* e extraíndo componentes principais, tem ambas as saídas (de elétrons e de píons) concordando sobre uma possível contaminação de píons no conjunto de elétrons. É necessário, entretanto, comprovar que os eventos identificados como contaminação, são realmente os eventos que contaminam o conjunto de dados. Isto pode ser feito se lembrarmos do fato já exposto

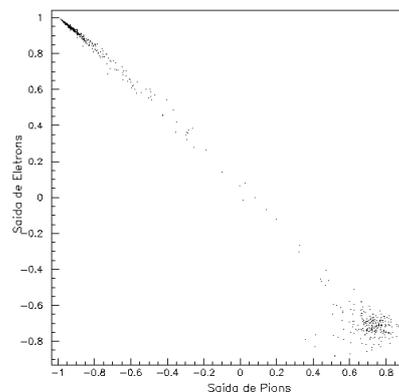


Figura 9: Correlação entre as saídas de elétrons e de píons provando a coerência interna do classificador neuronal.

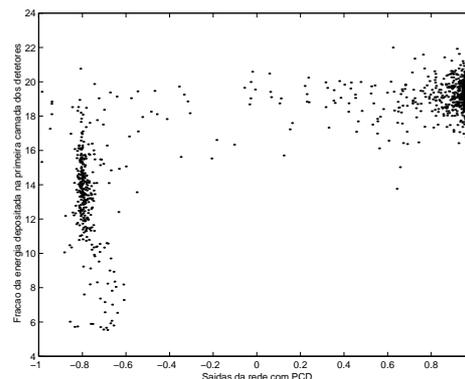


Figura 10: Correlação entre a energia depositada na primeira camada dos dois módulos do barril estendido as saídas da rede.

de que elétrons, ao interagirem com o calorímetro, perdem a maior parte de sua energia logo na primeira camada de células deste detector. Os píons, ao contrário, têm uma tendência a depositar uma maior fração de sua energia no meio do detector, atingindo mais as células da segunda e da terceira camada do calorímetro. Na Figura 10 vemos a correlação entre a saída de elétrons da rede neuronal contra a fração de energia depositada na primeira camada dos dois módulos do barril estendido para o conjunto de elétrons. Vemos que os eventos que depositaram menos energia nesta primeira camada (logo candidatos a píons de contaminação) fizeram com que o classificador neuronal apresentasse a sua saída de elétrons próxima a -1, ou seja, foram etiquetados como píons pela rede neuronal. Da mesma forma, os eventos que depositaram mais energia na primeira camada do detector (logo candidatos a elétrons de verdade) foram os que ativaram a saída de elétrons do classificador neuronal.

Finalmente, para completar a análise neuronal da contaminação em feixe de partículas, uma rede neuronal com três camadas foi livremente treinada, visando identificar a contaminação do feixe, conforme trabalhos anteriores (Damazio e Seixas, 1999). Utilizando-se do método de *backpropagation*, uma rede neuronal com 10 neurônios na camada escondida e 3 nós de saída foi treinada com o mesmo conjunto de dados. A Figura 11 mostra a correlação entre as saídas de elétrons para ambas redes neuronais classificadoras, ou seja, a rede treinada livremente e aquela treinada com componentes principais. Nota-se que há forte correlação na identificação da contaminação do feixe de elétrons.

A redução de dimensionalidade realizada pelo classificador neuronal treinado com componentes principais de discriminação, reduzindo a análise à projeção do espaço original em apenas três ou quatro componentes, é bastante importante quando se considera que este sistema de discriminação deve operar de forma *online*. Para tal, uma implementação em processadores digitais de sinais (Analog Devices Inc., 1995) ou em placas PowerPC (VME) vem sendo considerada, para a introdução do processamento neuronal no próximo período de testes do TileCal. Note-se que cada novo neurônio na camada escondida da rede corresponde a mais 256 multiplicações e acumulações, o que torna a compactação obtida pela análise de componentes principais de discriminação extremamente importante.

6 CONCLUSÕES

Foi desenvolvido um sistema neuronal de discriminação de partículas para auxiliar a avaliação de performance de um protótipo do calorímetro hadrônico de altas energias (Tilecal) que participará do importante experimento LHC. O sistema desenvolvido foi capaz de manter um alto nível de acerto na classificação de elétrons, píons e múons (acima de 97%), no caso de conjuntos de dados livres de contaminação, enquanto identificava corretamente a contaminação que ocorre em situações práticas de teste experimental.

A análise de componentes principais de discriminação foi capaz de reduzir drasticamente a alta dimensionalidade do espaço original de dados, mostrando que apenas quatro componentes são capazes de revelar as principais características dos perfis de deposição de energia das diferentes classes de partículas. Esta redução de dimensionalidade tem importância na implementação *online* do sistema classificador, visando-se rejeição dos eventos de contaminação do feixe durante a aquisição de dados.

A análise neuronal aqui desenvolvida é suficientemente

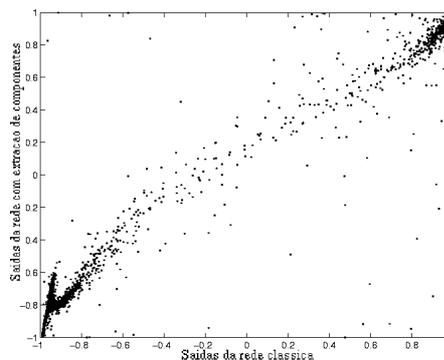


Figura 11: Correlação entre as saídas de uma rede neuronal padrão e de uma rede treinada usando-se PCD.

genérica para ser aplicada em vários problemas da calorimetria. Da mesma maneira, o processamento neuronal pode ser utilizado para a análise *offline* de dados contaminados (Seixas et al., 1999), ampliando a sua aplicabilidade.

AGRADECIMENTOS

Agradecemos ao apoio dado a este trabalho pelo CNPq, CAPES, FAPERJ e CERN (Suíça). Agradecemos, particularmente, aos nossos colegas da colaboração Tilecal, pela cessão do conjunto de dados experimentais que é usado neste trabalho e pelas discussões frutíferas sobre os resultados obtidos.

REFERÊNCIAS

- Analog Devices Inc. (1995). *DSP/MSP Products Reference Manual*.
- Caloba, L., Seixas, J. e Pereira, F. (1995). Neural discriminating analysis for a second-level trigger system, *Computing in High Energy Physics*.
- CERN (1999). LHC, the Large Hadron Collider Project. *<http://www.cern.ch/LHC/>
- Damazio, D. e Seixas, J. (1999). Outsider identification in a neural particle discriminator based on calorimetry, *Artificial Inteligency in High-Energy and Nuclear Physics* p. 6.
- Fahlman, S. e Lebiere, C. (1991). The cascade-correlation learning architecture, *School of Computer Science*.
- Hertz, J., Krogh, A. e Palmer, R. (1990). *Introduction to the Theory of Neural Computation*, Addison-Wesley.

Oja, E. (1995). *Neural Network for Chemical Engineers*, Elsevier Science.

Seixas, J., Maidantchik, C., Damazio, D. e Lacerda, F. (1999). A remote neural discriminator for the scintillating tile calorimeter of ATLAS., *VIII International Conference on Calorimetry in High Energy Physics Submetido a World Cientific* p. 6.

The ATLAS Collaboration (1999a). ATLAS, A Toroidal Lhc ApparatuS.
*<http://atlasinfo.cern.ch/Atlas/>

The ATLAS Collaboration (1999b). Tilecal, tile calorimeter detector unit.
*http://atlasinfo.cern.ch/Atlas/SUB_DETECTOR/TILE/tilecal.html

Wigmans, R. (1991). Advances in hadron calorimetry, *Rev. Nucl. Sci.* **41**: 133–185.