

# COMPARAÇÕES ENTRE REDES MLP E RBF NO RECONHECIMENTO DE CARACTERES EM IMAGENS DE PLACAS DE AUTOMÓVEIS

EUGENIO S. CANER, MARCIO P. DE ALBUQUERQUE, MARCELO P. DE ALBUQUERQUE, ALINE R. GESUALDI

*Coordenação de Atividades Técnicas, Centro Brasileiro de Pesquisas Físicas  
Rua Dr. Xavier Sigaud, 150, 22290-180, Urca, Rio de Janeiro, RJ, Brasil*

*E-mails: caner@cbpf.br, mpa@cbpf.br, marcelo@cbpf.br, aline@cbpf.br*

**Abstract** This paper compares the performances of a multilayer perceptron neural network and a radial basis function neural network, in pattern classification applications. Comparison of both theoretical and applied properties is conducted based on analysis of the published results as well as experiments conducted by authors.

**Keywords** Multilayer perceptron network (MLP), radial basis function network (RBF), neural discriminator, car plate.

**Resumo** Este trabalho compara os desempenhos de uma rede neural Perceptron Multi-Camadas e de uma rede neural de Funções Radiais de Base, em aplicações de classificação de padrões. A comparação das propriedades teóricas e práticas entre as baseou-se na análise de resultados publicados assim como de experiências conduzidas pelos autores.

**Palavras-chave** Perceptron Multi-Camadas, Funções Radiais de Base (RBF), discriminadores neurais, placas de veículos.

## 1 Introdução

Dentro de um sistema de reconhecimento de padrões, a classificação é uma das fases mais importantes. Nos últimos anos o uso de Redes Neurais Artificiais (RNA) tem ganhado popularidade para realizar esta tarefa nas mais diversas aplicações. Muitas publicações recentes descrevem como usar as RNA em sistemas de OCR-(Optical Character Recognition), reconhecimento de caracteres manuscritos e reconhecimento de caracteres em imagens de placas de automóveis (Hsin-Chia, 1998; Gesualdi, 2000; LeCun, 2003, Erkki, 2003).

A RNA é uma das metodologias de pesquisas que se desenvolvem com maior rapidez. Constantemente, surgem novos modelos tanto no ambiente científico como no de engenharia. Comparar teoricamente as melhorias dos novos modelos em relação aos anteriores é uma tarefa difícil. Uma das maneiras que permitem facilitar a análise é a realização de experimentos que comparem os resultados para uma dada aplicação e que no final se tornem em recomendações práticas para a implementação das RNA.

Neste trabalho compararemos as redes MLP (*Multi-Layer Perceptron*) treinadas com regularização com as redes RBF (*Radial Basis Function*) numa aplicação de reconhecimento de caracteres em imagens de placas de automóveis. Também compararemos, usando o mesmo conjunto de dados, com os resultados obtidos numa aplicação de reconhecimento de placas de automóveis, onde foi usada como classificador uma rede MLP clássica (Mello, 2001).

O trabalho está organizado da seguinte forma: A Seç. 2 apresenta as características das redes MLP e RBF. Na Seç. 3 descreveremos o experimento realizado e os resultados alcançados, a Seç. 4 apresenta as conclusões deste trabalho.

## 2 Redes MLP e RBF

### 2.1 Redes MLP

As redes MLP são os modelos de RNA mais utilizados (Haykin, 1999). São formadas por várias camadas de neurônios, cada uma com uma função específica. Os neurônios de uma determinada camada se conectam aos neurônios da camada seguinte. Uma rede de três camadas é apresentada na Figura 1.

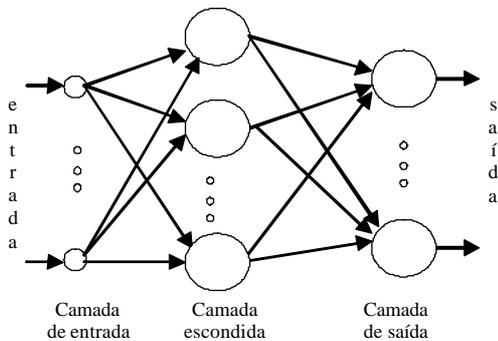


Figura 1. Topologia típica para uma rede MLP de três camadas.

Cada neurônio da rede calcula a função ganho para o produto interno do respectivo sinal de entrada com seu vetor de pesos. A função ganho é tipicamente a tangente hiperbólica ou a função sigmóide.

Geralmente, a rede MLP é treinada com o algoritmo *backpropagation* (Haykin, 1999) tendo como função-objetivo típica o erro médio quadrático. O procedimento de treinamento supervisionado aplicado à rede pode ser considerado como um problema de otimização não linear, no qual os pesos são atualizados iterativamente para que a função-objetivo definida, seja minimizada.

### 2.1.1 Regularização

Um dos problemas que ocorre durante o treinamento das redes MLP é o *sobre-ajuste* (*overfitting*). Este ocorre quando a rede memoriza os exemplos do treinamento, mas não aprende a generalizar às situações novas. Existem vários métodos para melhorar a generalização das RNA (Haykin, 1994; Anders e Korn, 1999). Um dos métodos é chamado *regularização*.

O objetivo da regularização é encontrar um balanço entre o número de parâmetros da RNA e a eficiência de ajuste, penalizando os modelos grandes. A função-objetivo é modificada de tal maneira que o algoritmo de treinamento consiga “podar” eficazmente a rede, levando os parâmetros irrelevantes a zero. Por exemplo, no caso do erro médio quadrático (*mse*) como função-objetivo, é agregado um termo que consiste na média da soma dos quadrados dos pesos e polarização da rede (*msw*). Veja as expressões:

$$msereg = \mathbf{g} \cdot mse + (1 - \mathbf{g}) \cdot msw \quad (1)$$

onde *msereg* é a nova função-objetivo,  $\mathbf{g} > 0$  é a constante de desempenho e

$$msw = \frac{1}{n} \sum_{j=1}^n \mathbf{w}_j^2 \quad (2)$$

Em (MacKay 1992, Foresee e Hagan, 1997), é colocado um método chamado de *regularização bayesiana* que determina de maneira automática os parâmetros ótimos da regularização. A vantagem principal deste método é que mesmo o modelo de RNA esteja sobre-parametrizado, o modelo comporta-se como uma rede pequena.

### 2.3 Redes RBF

Recentemente as redes de funções radiais de base vêm ganhando bastante atenção no domínio das RNA (Orr, 1999; Bors, 2001).

A estrutura de uma rede RBF consiste de apenas três camadas de neurônios. A primeira camada implementa uma conexão direta do espaço vetorial de entrada para a segunda camada ou camada escondida. Cada um dos neurônios que forma a camada escondida representa um centro (*cluster*) no espaço dos dados de entrada. Sendo assim, a camada escondida pode ser representada como uma série de tais *clusters*. Estes neurônios têm como função de ativação as funções radiais de base.

A função radial de base mais utilizada é a função de Gauss. Sua saída é determinada pela distância Euclidiana entre o centro e o vetor de entrada. Cada função requer um parâmetro vetorial - o centro, e um parâmetro escalar que define a *largura* da função radial. A última camada (camada de saída) transforma o espaço vetorial interno, implementando uma soma ponderada (processo linear) dos neurônios escondidos (Bors, 2001).

### 2.4 Comparação das redes MLP com a RBF

Ambos os modelos são aproximadores universais, ou seja, sempre existe uma rede RBF capaz de realizar a mesma função que uma rede MLP e vice-versa. Entretanto as redes MLP constroem aproximações *globais*. Elas são capazes de generalizar em regiões do espaço de entrada onde há poucos dados de treinamento – extrapolação. As redes RBF constroem aproximações *locais*, com uma aprendizagem mais rápida e menor sensibilidade na ordem de apresentação dos exemplos, (Bors, 2001).

A relação entre a entrada e as conexões sinápticas são diferentes: uma rede MLP calcula o produto escalar entre ambas e uma rede RBF calcula a distância Euclidiana ou distância de Mahalanobis.

Na Tabela 1 apresentamos vários trabalhos com resultados comparativos entre as redes MLP e RBF. A maioria indicou a preferência pelas redes RBF. O objetivo de nosso experimento descrito na próxima seção é determinar, no caso de reconhecimento de caracteres de placas de automóveis, qual rede é a mais apropriada.

Tabela 1. Conclusões de trabalhos comparando as redes MLP e RBF no campo de reconhecimento de padrões.

Pesquisadores	Área	Conclusões
Dong, 2003	Classificação imagens de satélites	Tempo de produção menor das redes RBF
Hawickhor, 1995	Reconhecimento de voz	Menor tempo de treinamento e generalização das redes RBF
Lu, 1999	Equalização de canais	Rede MLP com menos neurônios, rede RBF com menor taxa de erro
Fernandes, 1999	Detecção de sinais	Rede RBF com maior tempo de produção e menor tempo de treinamento
Roppel, 1999	Reconhecimento de odor	Rede MLP com taxa de identificação maior
Bettiol, 2001	Classificação de contingências em sistemas elétricos	Boa taxa de classificação das RBF.
Park, 2002	Identificação de dinâmicas em sistemas de potências.	A rede RBF com maior convergência, menor tempo e complexidade de treinamento.
Yampolskiy, 2003	Reconhecimento de algarismos manuscritos	Rede RBF com menor taxa de erro e menor tempo de treinamento

### 3 O experimento

#### 3.1 Descrição do conjunto de dados

Para a realização do experimento utilizamos uma biblioteca imagens de 2616 caracteres de algarismos e 1962 letras de placas de veículos do Rio de Janeiro. Estas imagens foram recortadas manualmente das imagens das placas, posteriormente redimensionadas para um tamanho de  $5 \times 6$  pixels e formatadas em vetores. Veja na Figura 2 exemplos das imagens nesse conjunto de dados.

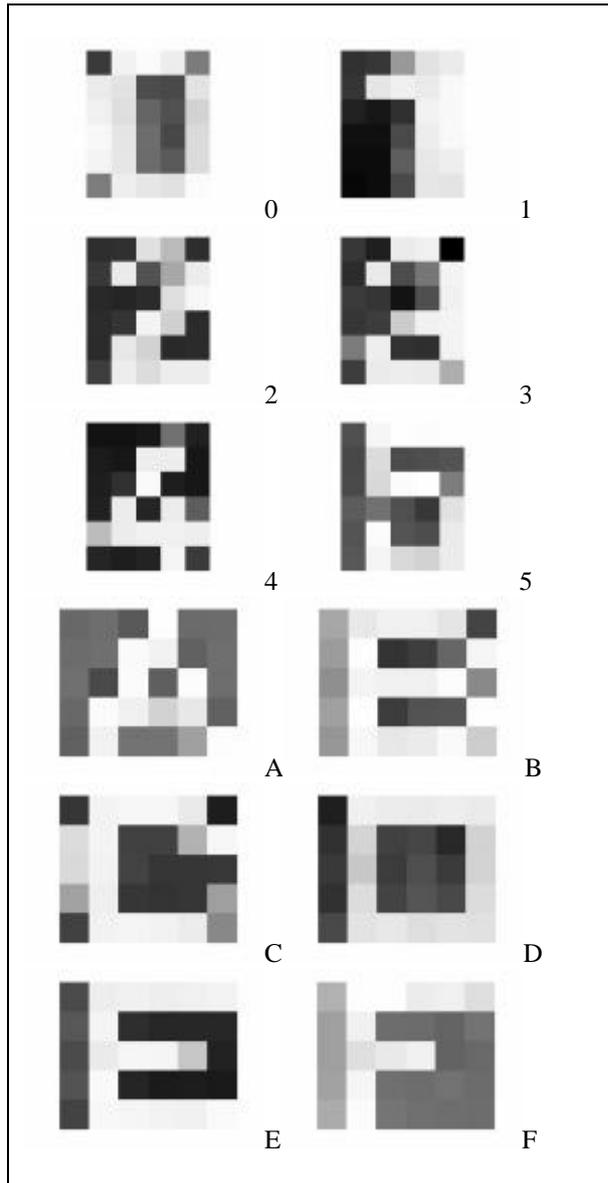


Figura 2. Exemplos do conjunto de dados

Esta biblioteca serviu de base para o desenvolvimento de um sistema de reconhecimento de placas descrito em (Mello, 2001), e com o qual realizaremos nossas comparações finais.

#### 3.2 Descrição do experimento

Simulamos uma rede MLP com regularização bayesiana e uma rede RBF separadamente para ambos conjuntos de dados: de imagens de algarismos e letras. Portanto teremos estruturas de RNA com a camada de entrada de 30 elementos ( $5 \times 6$  pixels) e a saída de 10 elementos para o caso dos algarismos e 26 para as letras.

Os critérios de comparação das RNA escolhidos são: *i*) a porcentagem de acerto (eficiência) no reconhecimento de caracteres, numéricos e letras, nunca antes vistos pela rede, ou seja, para um conjunto de teste; *ii*) o tempo de treinamento; *iii*) tempo de produção, ou seja, o tempo requerido pela RNA para reconhecer um determinado caractere; e *iv*) a robustez ou sensibilidade em presença de entradas significativamente diferentes aos dados de treinamento e teste.

A simulação das RNAs foi implementada usando o Matlab versão 6.0. Usamos a função *trainbr* que implementa uma MLP com otimização Levenberg-Marquard combinada com a regularização bayesiana (Foresee e Hagan, 1997) e *newrb* para simular uma rede RBF.

O treinamento das RNA foi repetido várias vezes para assegurar a obtenção de resultados consistentes e reproduzíveis. Em cada um dos experimentos determinamos os valores ótimos dos parâmetros das RNA.

#### 3.3 Otimização das redes MLP

Foi implementada uma topologia de três camadas. Para determinar o número ótimo de neurônios na camada escondida, testamos num intervalo de 5 a 20 neurônios. Os melhores resultados obtidos foram com 9 neurônios para a rede dos algarismos e 13 neurônios para o caso das letras. A Figura 3 mostra a eficiência com a variação do número de neurônios na camada escondida.

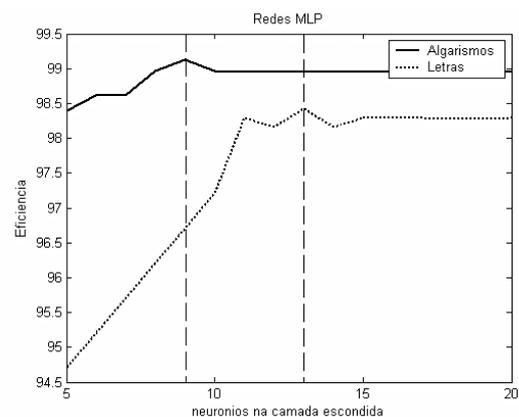


Figura 3. Eficiência alcançada pelas redes MLP (eixo vertical) e sua relação com o número de neurônios na camada escondida para a rede dos algarismos e das letras.

#### 3.4 Otimização das redes RBF

O parâmetro otimizado na simulação da rede RBF foi o *spread*, que caracteriza o a largura da função RBF,

e definido como o valor de entrada para o qual a função é 0,5 (ver Figura 4). Realizamos um teste do valor do *spread* para um intervalo de 1 a 10. Comprovamos que o melhor desempenho resulta para um *spread* com valor de 2 na rede para algarismos e 4 na rede das letras. (ver Figura 5)

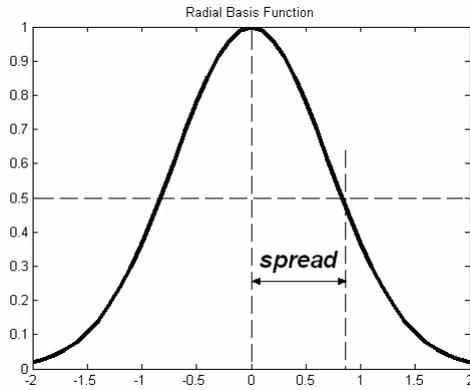


Figura 4. Exemplo de um *spread* de 0.83 para RBF.

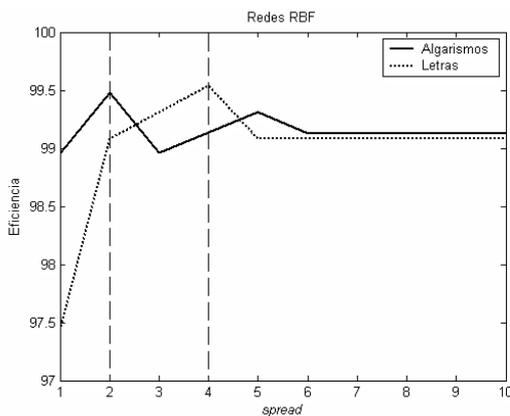


Figura 5. Eficiência no reconhecimento alcançada pelas RBF (eixo vertical) e sua relação com o *spread* para a rede dos algarismos e das letras.

### 3.5 Comparação de resultados

Depois de realizar os ciclos de treinamento e teste obtivemos os seguintes resultados.

**Tempo de treinamento:** Para um dado conjunto de dados o treinamento da RBF foi em média 8 vezes mais rápida.

**Tempo de produção:** Não obtivemos diferenças significativas entre as redes MLP e RBF. Provavelmente devido à programação orientada a objetos do software Matlab. Certamente como a estrutura das redes RBF é muito maior que às MLP, levariam um tempo produção maior numa aplicação real.

**Eficiência:** Calculamos a porcentagem de caracteres corretamente classificados do conjunto de teste. Comparamos também com os resultados obtidos numa rede MLP de treinamento clássica. Apresentamos os resultados na tabela abaixo.

Tabela 2. Comparação das eficiências do conjunto de teste para as diferentes redes analisadas.

	MLP (Mello, 2001)	MLP- com Regularização	RBF
algarismos	97,3%	99,1%	99,5%
letras	97,7%	98,4%	99,5%

**Robustez:** Foram simuladas entradas artificialmente modificadas. Adicionamos aos dados de entrada do conjunto de teste ruídos com distribuição normal, média 0 e desvio padrão entre 0 e 100. Para cada nível de ruído realizamos 100 amostragens aleatórias, adicionando-o às imagens do conjunto de teste e calculando a eficiência média. A Figura 6 mostra a eficiência no reconhecimento com a variação do nível de ruído (amplitude do desvio padrão) para cada conjunto de dados (algarismos e letras) e as redes analisadas.

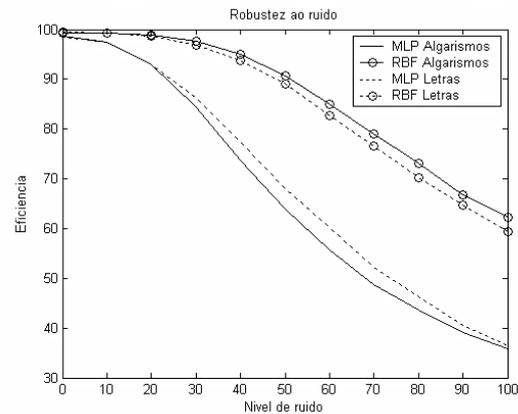


Figura 6. Eficiência no reconhecimento alcançada pelas redes (eixo vertical) e sua relação com o nível de ruído (amplitude do desvio padrão).

Outro teste realizado foi recriar imagens de algarismos e letras artificialmente deslocados, simulando erros no recorte das imagens. Testamos com dois tipos de deslocamento: um para a direita e outro para abaixo. No deslocamento para a direita, cada coluna da imagem recebe os valores da coluna anterior e a primeira coluna recebe o valor 0. No deslocamento para abaixo, cada linha da imagem recebe os valores da linha anterior e a primeira linha recebe o valor 0. Os resultados mostraram que as duas redes foram muito sensíveis, porém a RBF teve uma melhor eficiência.

## 4 Conclusão

Baseado no experimento realizado, onde testamos cada rede sob idênticas condições podemos afirmar que as redes RBF são mais apropriadas para aplicações de reconhecimento de imagens de algarismos e letras de placas de automóveis. Podemos destacar as seguintes vantagens:

- erro final atingido por uma rede RBF foi menor que o de uma rede MLP;

- a convergência de uma rede RBF pode chegar a uma ordem de grandeza mais rápida do que a convergência de uma rede MLP;
- maior robustez da rede RBF a entradas ruidosas.

Os resultados obtidos são consistentes com as conclusões de vários outros pesquisadores em aplicações relacionadas ao reconhecimento de padrões por meio de RNA. Extensões desse trabalho podem ajudar a melhor caracterizar a eficiência para imagens deslocadas.

### Agradecimentos

Gostaríamos de agradecer à Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) e ao Centro Brasileiro de Pesquisas Físicas/MCT pelo apoio a este trabalho.

### Referências Bibliográficas

- Anders, U. e Korn, O. (1999). Model Selection in Neural Networks, *Neural Networks*, 12, pp. 309—323.
- Bettiol, A. L., Tesh, J. R. e Todesco, J. L. Classificação de contingências para a avaliação de estabilidade transitória em sistemas elétricos usando redes neurais artificiais. (2001). *XVI Seminário Nacional de Produção e Transmissão de Energia Elétrica*. Campinas, São Paulo.
- Bors, G. (2001). Introduction of the Radial Basis Function (RBF) Networks. *Online Symposium for Electronics Engineers*, issue 1, vol. 1, DSP Algorithms: Multimedia, <http://www.osee.net/>, pp. 1-7.
- Dong, Q. (2003). MLP and RBF comparison in a classification problem of a satellite image. Disponível em: <http://www.stanford.edu/qfdong/Work/NN.pdf>
- Erkki, O., Laaksonen, J., e Tuura, L. (2003). Application of Statistical and Neural Classifiers to Recognition of Handwritten Digits. *Quinquennial report*. Disponível em: <http://www.cis.hut.fi/research/reports/>.
- Fernandes, M. A., Neto, A. D. e Bezerra, J. B. (1999). Aplicação das Redes RBF na Detecção Inteligente de Sinais Digitais. *Proceedings of the IV Brazilian Conference on Neural Networks* - pp. 226-230, São José dos Campos.
- Foresee, F. D., e Hagan, M. T. (1997). Gauss-Newton Approximation to Bayesian Regularization, in *IEEE International Conference on Neural Networks* (Vol. 3), pp. 1930—1935, New York. IEEE.
- Gesualdi, A. R., Albuquerque, M. P., Seixas, J. M., Calôba, L. P. e Albuquerque, M .P. (2000). Recognition of characters in plates of Brazilian particular vehicles using neural networks. *Proceedings of XIII Brazilian Symposium on computer Graphics and image processing*, 333.
- Hawickhorst, B. A , Zahorian, S. A, e Rajagopal, R. A. (1995). Comparison of Three Neural Network Architectures for Automatic Speech Recognition. *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol. 5, Fuzzy Logic and Evolutionary Programming, pp. 221-226.
- Haykin, S. (1999). Neural Networks: A comprehensive foundation, *Prentice Hall*, ISBN 0-13-908385-5. p.396
- Hsin-Chia, F. e Yeong, Y. X. (1998). Multilingual handwritten character recognition by Bayesian decisionbased neural networks. *IEEE transactions on signal processing*, Vol. 46, No. 10.
- LeCun, Y., Boser, B., Denker, J., Henderson, D., e Jackel, L. D. (2003). Handwritten digit recognition with a back-propagation network. *Advances in Neural Information Processing Systems*, Vol. 2.
- Lu, B. e Evans, B. L. (1999) Channel Equalization by Feedforward Neural Networks. *Proc. IEEE Int. Sym. on Circuits and Systems*, Orlando, FL, vol. 5, pp. 587-590.
- MacKay, D. J. C. (1992). Bayesian Interpolation, *Neural Computation*, 4, 415—447.
- Mello, A. G. (2001). Sistema neuronal de reconhecimento de placas em veículos automotores particulares, COPPE/UFRJ/MSc.
- Park, J., Harley, R G., e Venayagamoorthy. G. K.. (2002). Comparison of MLP and RBF Neural Networks Using Deviation Signals for On-Line Identification of a Synchronous Generator. in *Proc. Of IEEE PES Winter Meeting*, New York, Vol.1, pp. 274-279.
- Roppel, T., Wilson, D., Dunman, K., Becanovic, V., e Padgett, M. L. (1999). Design of a Low-Power, Portable Sensor System Using Embedded Neural Networks and Hardware Preprocessing. *Intl Joint Conf Neural Networks*, Washington D.C., June 1999.
- Orr, M. J. L. (1999). Introduction to Radial Basis Function Networks. Disponível em [www.anc.ed.ac.uk/papers/recad.ps](http://www.anc.ed.ac.uk/papers/recad.ps).
- Yampolskiy, R. and Novikov, D. (2003) Experimental Study of the Choice Between MLP and RBF Neural Networks for Character Recognition, *Western New York Image Processing Workshop*.