

# Detecção de Bordas com Redes Neurais BAM e ART

Ana Paula Abrantes de Castro  
*Instituto Nacional de Pesquisas Espaciais*  
[apaula@lac.inpe.br](mailto:apaula@lac.inpe.br)

José Demisio Simões da Silva  
*Instituto Nacional de Pesquisas Espaciais*  
[demisio@lac.inpe.br](mailto:demisio@lac.inpe.br)

## Resumo

*Esse artigo apresenta um estudo e a implementação de redes neurais na tarefa de detecção de bordas em imagens digitais. O desempenho das redes neurais utilizadas é comparado com o desempenho de métodos de detecção de bordas considerados clássicos na literatura de visão computacional.*

## 1. Introdução

Um esforço considerável tem sido dedicado à solução do problema de reconhecimento e caracterização de objetos presentes em uma imagem. O uso de técnicas óticas de processamento predominou até a década de 70 (Casasent e Psaltis, 1976). No início dos anos 80, as técnicas digitais passaram a ser mais empregadas, com os avanços verificados na microeletrônica e no desenvolvimento de arquitetura paralela de processamento. Atualmente, o amadurecimento das técnicas computacionais inteligentes, como Sistemas Especialistas, Lógica Nebulosa, Redes Neurais e Algoritmos Genéticos, têm permitido novas abordagens para esse problema (Kulkarni, 2001).

As técnicas computacionais inteligentes procuram capturar, tanto na teoria como na prática, as características de fenômenos da natureza. A complexidade do problema de reconhecimento e de classificação de imagens, que é de difícil abordagem algorítmica, tem tornado o uso dessas técnicas cada vez mais freqüente. Em particular, as Redes Neurais Artificiais (RNA's) têm sido mais exploradas (Kulkarni, 2001).

As Redes Neurais Artificiais são dispositivos não-lineares, inspirados na funcionalidade dos neurônios biológicos, aplicados no reconhecimento de padrões, na otimização e na previsão de sistemas complexos (Treleaven, et. al., 1989). São compostas por diversas unidades computacionais organizadas em paralelo e parcial ou totalmente interconectadas. Cada uma dessas unidades, "neurônios artificiais", efetua operações simples e transmite seus resultados às unidades com as quais possui conexão.

Através de um processo de treinamento, as redes neurais passam a ser capazes de reconhecer padrões, mesmo que os dados utilizados nesse treinamento sejam não-lineares, incompletos ou até mesmo contraditórios. A habilidade de manipular dados imprecisos faz com que as redes neurais sejam eficazes em tarefas complexas ou quando um conjunto de regras não pode ser facilmente formulado.

Em visão computacional, tanto a dificuldade de estabelecer algoritmos de reconhecimento quanto à natureza dos dados com os quais se trabalha, indicam as redes neurais como uma solução promissora.

A implementação de um sistema de visão artificial para detecção de bordas em imagem, baseado em redes neurais usando suas características intrínsecas, isto é, sua capacidade de extrair padrões de conjuntos de dados complexos.

A motivação para o desenvolvimento deste trabalho foi um projeto de iniciação científica sobre um sistema de navegação autônoma adaptável, baseado em imagens que utiliza o controle por lógica nebulosa (Castro e Silva, 2001).

O sistema de lógica nebulosa determina a direção a ser seguida e a velocidade a ser empregada para movimento do robô. Foi usado um operador gradiente para extrair informações de bordas determinando as direções das faixas da pista. Este operador apresentou grande sensibilidade a ruídos nas imagens, devido à transmissão por rádio frequência (VHF).

Neste trabalho, as Redes Neurais Artificiais são utilizadas e testadas como uma alternativa ao operador gradiente, buscando menor tempo de detecção de característica e maior robustez a ruído. O desempenho das redes neurais é comparado com o desempenho de operadores tradicionais de visão computacional para detecção de bordas, tais como: Sobel (Prewitt, 1970); Roberts (1965); Canny (1986) (Ding et. al., 2001) e o operador baseado em lógica nebulosa desenvolvida por Cho e Cho (1994).

Nas seções 2 são apresentados dois modelos de redes neurais. Na seção 3 é apresentada a metodologia do trabalho. Na seção 4 alguns resultados obtidos são mostrados. Na seção 5 uma conclusão sobre os resultados obtidos é apresentada.

## 2. Modelos de Redes Utilizados

Nesse artigo são utilizados dois modelos de redes neurais. O primeiro modelo é a rede neural *Bidirectional Associate Memory* (BAM). No início de 1980, Hopfield propôs um modelo de memória auto-associativa para armazenar e recordar informações do mesmo modo que o cérebro humano. A rede proposta era capaz de associar um conceito B a um conceito C e de, dado B recordar-se de C e vice-versa.

No modelo de Kosko (1987) uma BAM é uma combinação heteroassociativa de  $m$  padrões representado em duas camadas, codificando pares de padrões diferentes codificados como pares de padrões espaciais binários (0, 1) ou bipolares (-1, 1), como  $(\bar{X}_k, \bar{Y}_k)$ ,  $k = 1, 2, \dots, m$  usando a regra de aprendizagem de *Hebb* (Haykin 2001). Os vetores das equações 2.1 e 2.2 representam o  $k$ -ésimo par de padrões.

$$\bar{X}_k = (x_{k1}, x_{k2}, \dots, x_{kn}) \quad 2.1$$

$$\bar{Y}_k = (y_{k1}, y_{k2}, \dots, y_{kp}) \quad 2.2$$

A arquitetura de uma memória associativa bidirecional é mostrada na Figura 1.

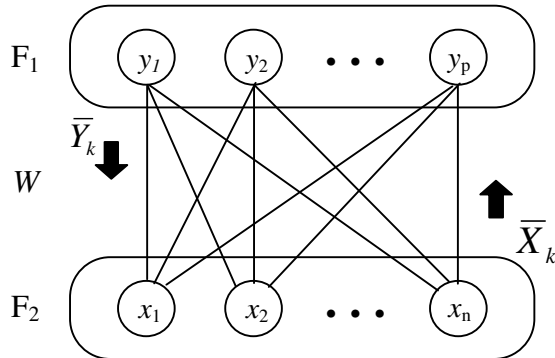


Figura 1: Arquitetura de uma Memória Associativa Bidirecional

A arquitetura da BAM é composta por duas camadas de neurônios ( $F_1$  e  $F_2$ ) com conexões sinápticas bidirecionais entre elas. A rede envia sinais para frente e para trás entre as duas camadas até que todos os neurônios alcancem o equilíbrio, ou seja, até que cada ativação do neurônio permaneça constante por vários passos. A BAM tem habilidade de aprender “on line”, operando de forma discretizada no tempo e incorporando conexões de realimentação na sua arquitetura.

O número total de nós da memória na BAM é a soma dos números de nós de cada camada. O número de conexões é calculado pelo produto do número de unidades em cada camada ( $n_1 = n^\circ$ . de neurônios na camada1  $\times$   $n^\circ$ . de neurônios na camada2).

A aprendizagem na BAM é Hebbiana, sendo dada pela equação:

$$W = \sum_{k=1}^m (\bar{X}_k)(\bar{Y}_k) \quad 2.3$$

onde  $W$  é a matriz de pesos resultante,  $\bar{X}_k$  e  $\bar{Y}_k$  formam  $k$ -ésimo par de padrões e  $m$  é o número total de pares de padrões aprendidos.

Após a fase de aprendizagem, quando a BAM é exposta a um padrão, ela tenta recuperar o respectivo padrão associado. As equações básicas que regem a ativação são:

$$\bar{Y}_k = \bar{X}_k W \quad 2.4$$

$$\bar{X}_k = \bar{Y}_k W^T \quad 2.5$$

onde  $W$  e  $W^T$  - são as matrizes pesos e sua transposta respectivamente. Uma vez que o padrão  $\bar{Y}_k$  é fornecido, a

BAM tenta “relembrar” o padrão  $\bar{X}_k$  e vice-versa. Estes procedimentos de “relembrar” padrões são executados pelas equações (2.4) e (2.5).

A segunda rede neural utilizada é a *Adaptive Resonance Theory* (ART) desenvolvida por Carpenter e Grossberg em 1976 como uma alternativa para superar o problema do dilema estabilidade-plasticidade. A ART é uma rede do tipo auto-organizável (*self-organizing*), cuja idéia central está na autodefinição dos seus parâmetros, constituindo assim, redes que utilizam o paradigma de aprendizado não-supervisionado (Braga et. al., 2000).

A primeira versão da rede ART proposta, foi a rede ART1 que trabalha apenas com valores binários. Posteriormente foi desenvolvida a rede ART2 como uma versão para trabalhar com valores de entrada contínuos. Outros modelos da família ART são as redes ARTMAP, que é uma rede mais sofisticada que permite o aprendizado incremental, ART3, Fuzzy-ART e Fuzzy-ARTMAP (Fausett, 1994).

O modelo utilizado nesse trabalho é a ART1, a arquitetura dessa rede é composta por duas camadas de neurônios (Figura 2). Uma camada de entrada, que processa os dados de entrada e uma camada de saída, que agrupa os padrões de treinamento em *agrupamentos*. Estas duas camadas estão conectadas por meio de dois conjuntos de conexões que unem cada neurônio de uma camada a todos os neurônios da outra.

O primeiro conjunto representado por conexões *para frente*,  $W$ , assume valores reais e segue da camada de entrada para a camada de saída. O segundo conjunto, que contém as conexões de *realimentação*,  $\bar{t}$ , assume valores binários e conecta os neurônios da camada de saída aos neurônios de entrada. Assim, o neurônio  $i$  da camada de entrada está conectado ao neurônio  $j$  da camada de saída através do peso  $w_{ji}$ . Da mesma forma, o neurônio  $j$  da camada de saída está conectado ao neurônio  $i$  da camada

de entrada através do peso  $t_{ji}$ . Entre os neurônios da camada de saída existem pesos que atuam como inibidores laterais. Os pesos *para frente* e de *realimentação* são também conhecidos como *filtros adaptáveis*, por adaptarem, dinamicamente, seus valores para possibilitar o aprendizado de novos padrões. O vetor de pesos de *realimentação* associado a cada neurônio da camada de saída pode ser visto como um protótipo ou exemplo dos padrões pertencentes ao *agrupamento* representado por este neurônio.

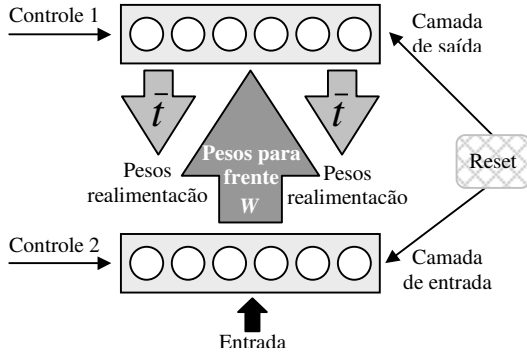


Figura 2 – Arquitetura básica da rede ART, que possui duas camadas, uma de entrada e outra de saída, duas unidades de controle e um mecanismo de reset.

FONTE: Modificada de Beale e Jackson (1990)

A rede também possui um conjunto de pesos bipolares não-adaptáveis entre os neurônios da camada de saída. O peso entre dois neurônios da camada de saída é  $-1$ , se este liga dois neurônios distintos; e  $+1$  se liga a sua saída a si mesmo. Para cada camada da rede ART1 há uma unidade externa de controle:  $C_1$  para a camada de entrada e  $C_2$  para a camada de saída. Estas unidades de controle  $C_1$  e  $C_2$  determinam o fluxo de dados para as camadas de entrada e saída, respectivamente, sendo que a segunda ainda tem a função de habilitar ou desabilitar neurônios da camada de saída.

Entre as camadas de entrada e saída existe um mecanismo de *reset*, responsável por verificar a semelhança entre um vetor de entrada e um dado vetor protótipo, utilizando um limiar de vigilância ( $\rho$ ). Este limiar de vigilância determina se um padrão de entrada pode ser incluído em um dos *agrupamentos* existentes.

O algoritmo de aprendizado da rede ART1 é não-supervisionado e pode ser ativado a qualquer momento, permitindo que a rede aprenda novos padrões continuamente. O treinamento da rede ART1 pode ser rápido ou lento. No primeiro caso, os pesos *para frente* são ajustados para seus valores ótimos em poucos ciclos, geralmente em apenas um ciclo de treinamento. No aprendizado lento, os pesos são ajustados lentamente em vários ciclos de treinamento, possibilitando um ajuste melhor dos pesos da rede aos padrões de treinamento.

Com relação aos parâmetros, durante o treinamento a rede ART se apresenta muito sensível. O parâmetro mais crítico é o limiar de vigilância que controla a resolução do processo de classificação. Se  $\rho$  assume um valor baixo, menor que 0.4, a rede agrupa padrões não muito semelhantes em um mesmo *agrupamento*, criando poucas classes. Por outro lado, atribuindo-se um valor alto para  $\rho$ , pequenas variações nos padrões de entrada levarão à criação de novas classes.

Antes da aprendizagem, os pesos e os parâmetros da rede devem ser inicializados. Todos os pesos de *realimentação* ( $\bar{t}$ ) têm valor 1. Os valores iniciais do vetor de pesos *para frente* ( $W$ ) são determinados por:

$$w_{ji}(0) = \frac{1}{1+n}, \forall i, j \quad 2.6$$

onde  $n$  é o número de neurônios de entrada da rede. O limiar de vigilância também é inicializado, assumindo um valor dentro do intervalo  $0 < \rho < 1$ .

O processo de aprendizagem da rede ART1 envolve três fases: reconhecimento, comparação e busca. Na fase de reconhecimento, cada neurônio da camada de entrada recebe três sinais: um sinal do vetor de entrada, um sinal de *realimentação* da camada de saída e um sinal da unidade de controle  $C_1$ . Um neurônio da camada de entrada somente será ativado se receber pelos menos dois sinais excitatórios. Esta regra é denominada *regra dos dois terços* (RDT). Esta regra é aplicada a cada unidade de entrada, gerando um vetor resultante que é apresentado aos neurônios da camada de saída, através dos pesos sinápticos *para frente* ( $W$ ). Esta apresentação resulta no cálculo das atividades internas de cada neurônio  $j$  na saída, segundo a equação:

$$v_j = \sum_i w_{ji} x_i \quad 2.7$$

onde  $v_j$  é a ativação do neurônio de saída  $j$ , definido pelo produto interno entre seu vetor de pesos  $W_j$  e o vetor de entrada  $\bar{X}$ .

O neurônio da camada de saída que tiver o maior valor de atividade interna é selecionado como o representante do provável *agrupamento* para armazenar o novo padrão. A inibição lateral faz com que somente o neurônio vencedor da camada de saída fique ativo. O neurônio selecionado  $j^*$  envia seu vetor de pesos  $\bar{t}_{j^*}$  para a comparação na camada de entrada.

Na fase de comparação, o valor da unidade de controle  $C_1$  é igual a 0, pois existe um neurônio na camada de saída que está ativo. Novamente, aplica-se a RDT na camada de entrada, gerando-se o *vetor de comparação*  $\bar{Z}$ .

O mecanismo de reset é responsável por testar a similaridade entre o *vetor de entrada* e o *vetor de comparação*. O teste é feito pelo cálculo do índice de

similaridade  $P$  pela equação (2.8), que é comparando com o limiar de vigilância.

$$P = \frac{\left\| \sum x_i z_{ij} \right\|}{\left\| \sum x_i \right\|} \quad 2.8$$

Se  $P$  for maior que  $\rho$ , então o padrão de entrada é incluído no *agrupamento* do neurônio da camada de saída ativado, sendo ajustados os vetores de pesos do neurônio vencedor para incorporar as características do vetor de entrada atual.

Entretanto, se  $P$  não superar o limiar de vigilância, outro neurônio da camada de saída deverá ser verificado para armazenar o vetor de entrada. Assim, a rede ART entra na fase de busca, procurando selecionar um novo neurônio da camada de saída que represente o vetor de entrada corrente. O neurônio da camada de saída rejeitado pelo  $\rho$  é impedido de participar da nova seleção.

O vetor de entrada é reapresentado e a rede entra novamente na fase de comparação. O processo se repete até que se encontre um neurônio de saída que melhor se assemelhe ao vetor de entrada atual, dentro dos limites do limiar de vigilância.

Se nenhum neurônio de saída for encontrado, o vetor de entrada é então considerado como sendo de uma classe desconhecida até então, e a rede ART aloca um neurônio de saída “livre” (não associado ainda a qualquer *agrupamento* existente) para representá-lo.

### 3. Desenvolvimento do trabalho

No trabalho de pesquisa em desenvolvimento, estuda-se o uso de modelos de redes neurais, como alternativa para realizar tarefas na área de processamento digital de imagens. O objetivo é desenvolver algoritmos de processamento de imagens que possuam características similares ao cérebro humano, sendo as redes neurais artificiais apropriadas para tal, devido a sua inspiração neurobiológica. Especificamente, neste artigo apresentamos resultados da aplicação de dois modelos de redes neurais na tarefa de detecção de bordas em imagens.

As redes utilizadas foram a Memória Associativa Bidirecional e a rede da Teoria de Ressonância Adaptável, cujas topologias foram apresentadas anteriormente na seção 2.

Três treinamentos foram realizados para cada rede. No primeiro foram utilizados oito padrões binários de bordas com dimensão  $3 \times 3$  (Figura 3).

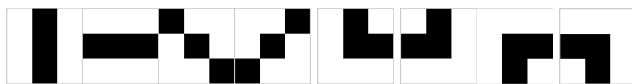


Figura 3 – Padrões Binários utilizados no treinamento das redes neurais.

Os testes de ativação das redes foram realizados sobre imagens em níveis de cinza e coloridas (RGB) com dimensões  $512 \times 512$ . Foi utilizada uma janela deslizante  $3 \times 3$  sobre a imagem principal, para extração de estruturas locais de bordas. Antes da apresentação à rede, essa janela de pixels é normalizada. Quando um padrão de borda é detectado, uma matriz binária equivalente é associada a uma área equivalente numa imagem de saída.

A informação de existência de bordas em uma área é dada pelo cálculo da variância, utilizada para uma pré-seleção das janelas da imagem que devem ser submetidas ao processamento pelas redes. Se não for detectado um padrão em alguma janela, associa-se uma janela  $3 \times 3$  de zeros à área equivalente da imagem de saída.

No segundo treinamento das redes foram adicionados quatro padrões binários (Figura 4).

No terceiro treinamento foram utilizados padrões de bordas bipolares (-1 e 1).

Na seção 4, são apresentados os resultados obtidos com as redes neurais utilizadas, bem como os resultados obtidos utilizando-se algoritmos de detecção de bordas considerados clássicos na literatura.

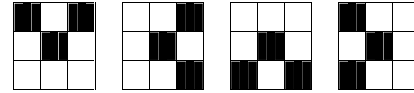


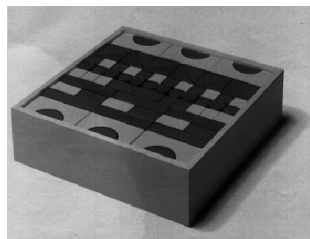
Figura 4 – Padrões Binários adicionais utilizados no treinamento das redes neurais.

### 4. Resultados

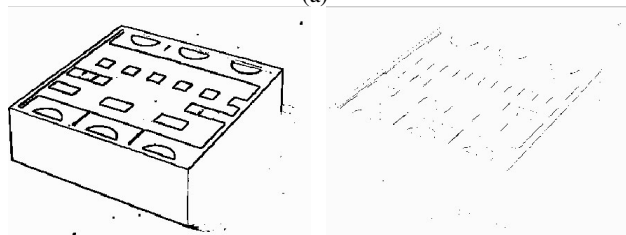
Na Figura 5 apresenta-se o resultado da aplicação das redes numa imagem em níveis de cinza. Observa-se que o resultado para a rede ART (Figura 5.c) foi inferior ao resultado da rede BAM (Figura 5.b), foram utilizados 8 padrões binários. No segundo treinamento foram adicionados mais 4 padrões binários (Figura 4), e visualmente o resultado da rede BAM manteve-se inalterado (Figura 5.d), enquanto que para a rede ART houve uma melhora significativa (Figura 5.d).

Na Figura 6 são apresentados os resultados obtidos com o treinamento da rede BAM, com os padrões bipolares. A análise visual comparativa com os resultados da Figura 5 mostra que não houve melhoras significativas. Não foram feitos testes para o treinamento da rede ART com padrões bipolares.

Na Figura 7 são apresentados os resultados da aplicação de detectores de bordas clássicos e do detector de borda baseado em lógica nebulosa de Cho e Cho (1994), utilizando a imagem da Figura 5.a. A comparação entre os métodos neurais e não neurais, mostra que o operador de Canny (Figura 7.c), visualmente, apresenta um resultado mais otimizado. Entretanto, os resultados das redes neurais mostram bordas mais definidas em termos de intensidade.



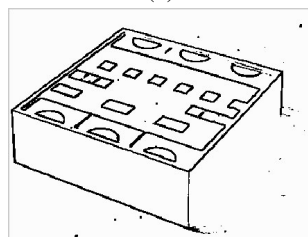
(a)



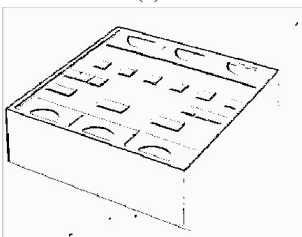
(b)



(c)

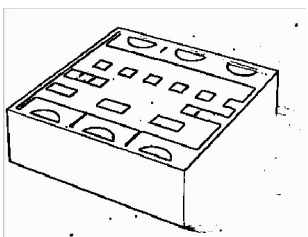


(d)

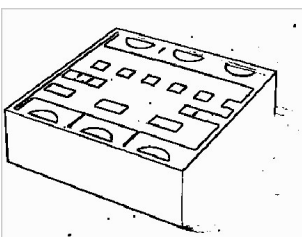


(e)

Figura 5 – (a) Imagem original, (b) BAM 8 padrões, (c) ART 8 padrões, (d) BAM 12 padrões, (e) ART 12 padrões.



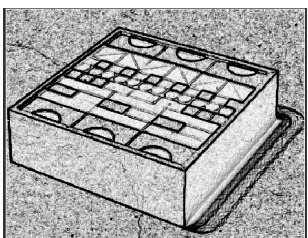
(a)



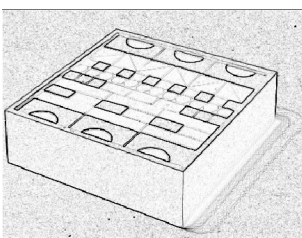
(b)

Figura 6 – (a) BAM 8 padrões bipolares, (b) BAM 12 padrões bipolares.

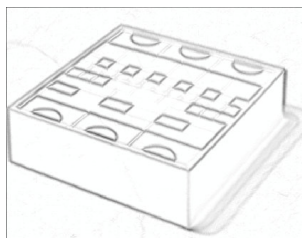
Na Figura 8 apresenta-se os resultados para as redes treinadas com os 8 padrões binários, em uma imagem colorida. Observa-se que ambas as redes tiveram desempenho semelhantes àqueles da Figura 5.



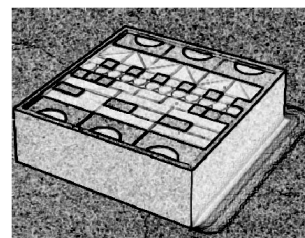
(a)



(b)

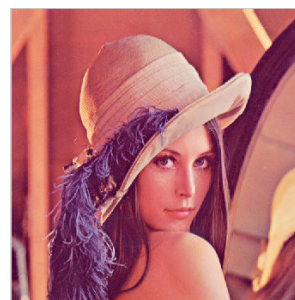


(c)



(d)

Figura 7 – (a) Sobel, (b) Roberts, (c) Canny e (d) Fuzzy



(a)



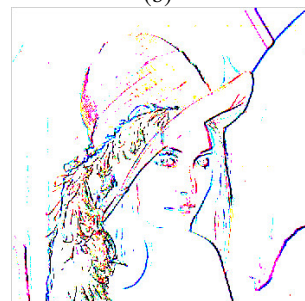
(c)



(b)



(d)



(e)

Figura 8 – (a) Imagem original, (b) BAM 8 padrões, (c) ART 8 padrões, (d) BAM 12 padrões, (e) ART 12 padrões.



(a)



(b)

Figura 9 – BAM:(a) 8 padrões bipolares,(b) 12 padrões bipolares.



A Figura 9 mostra os resultados do treinamento da BAM com padrões bipolares. Também é possível notar que não houve uma melhora significativa.

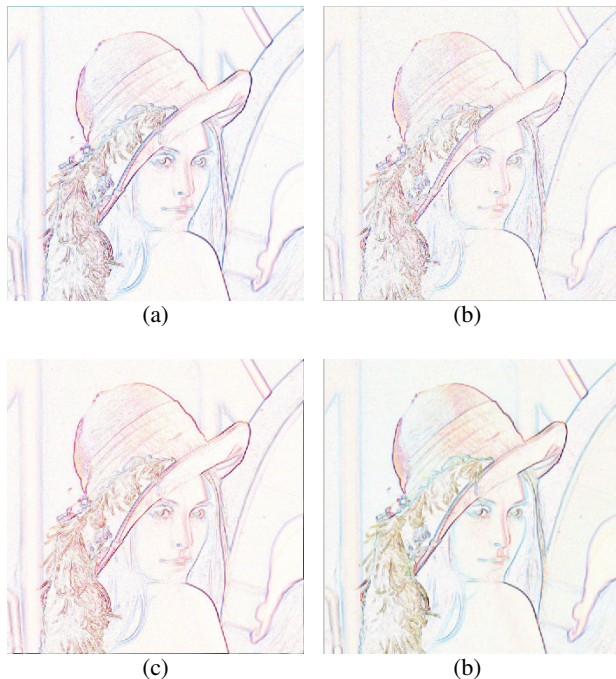


Figura 10 – (a) Sobel, (b) Roberts, (c) Canny e (d) Fuzzy

A Figura 10 apresenta resultados do uso dos mesmos operadores utilizados na Figura 7. Observa-se que os operadores de Sobel e de Cho e Cho (1994) apresentam melhoras significativas, quando comparados à Figura 7. Entretanto, o resultado do operador de Canny ainda é mais homogêneo. Os resultados das redes neurais se mostram semelhantes aos operadores de visão computacional apresentados, com uma definição melhor da intensidade da borda.

## 5. Conclusão

Este trabalho apresenta resultados parciais de uma pesquisa em andamento, sobre o estudo de redes neurais artificiais na detecção de bordas em imagens.

Nos treinamentos das redes, foram usados padrões binários e bipolares, não sendo observados grandes alterações nos resultados.

Observou-se que a rede ART foi influenciada pelo número de padrões utilizados. Provavelmente, o problema com a rede ART está associada à escolha do limiar de vigilância. Foram realizados testes, não reportados neste artigo, usando limiares diferentes e observaram-se sensíveis alterações na resposta da rede ART.

A análise comparativa com os métodos clássicos de detecção de bordas mostra resultados semelhantes, com a

as redes apresentando bordas mais definidas na intensidade.

Apesar do uso de redes neurais na detecção de bordas mostrar-se promissor, novos testes, novos modelos de redes serão testados, procurando evidenciar todas as vantagens na utilização de detectores baseados em redes neurais.

## 6. Referências

- Beale R. & Jackson, T. *Neural Computing: An Introduction*, Bristol, and Publishing Ltda, 1990.
- Braga A. P.; Carvalho A. P. L. F. & Ludermitr T. B. *Redes Neurais Artificiais Teorias e aplicações*, Rio de Janeiro: LTC, 2000.
- Canny, J.F. "A computational approach to edge detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1986, **8**, 679-698.
- Casasent D. & Psaltis D. "Position, rotation, and scale invariant optical correlation". *Applied Optics*, July 1976, **15**(7), 1795-1799.
- Castro, A.P.A. & Silva, J.D.S.; Simoni, P.O. "Image Based Autonomous Navigation Fuzzy Logic Control", *IJCNN - International Joint Conference on Neural Networks*, Washington, 2001, **1**, 2200—2205.
- Cho, S.M. & Cho, J.H. "Thresholding for Edge Detection Using Fuzzy Reasoning Techniques", *IEEE International Conference on Communication Systems (ICCS'94)*, 1994, **3**, 1121-1124.
- Ding, L. & Goshtasby, A. "On the Canny edge detector", *Pattern Recognition*, 2001, **34**, 721—725.
- Fausett, L. *Fundamentals of Neural Networks*, Prentice Hall, Upper Saddle River, New Jersey, 1994.
- Haykin, S. *Redes Neurais Princípios e Prática*, Bookman, Porto Alegre, 2001.
- Kosko B., "Bidirectional Associative Memories", *IEEE Transactions on Systems, Man, and Cybernetics*, 1988, **18**, 49-60.
- Kulkarni, A. D. *Computer Vision and Fuzzy-Neural Systems*, Prentice Hall, Upper Saddle River, New Jersey, 2001.
- Prewitt, J.M.S., *Object enhancement and extraction*, In B. S. Lipkin and A. Rosenfeld, editors Picture Processing and Psychopictorics, Academic Press, New York, 1970.
- Roberts, L.G. *Machine Perception of three-dimensional solids*, In J. P. Tippet et al., editors, Optical and Electro optical Information Processing, MIT Press, Cambridge, MA, 1965.
- Treleaven P.; Vellasco, M., Pacheco, M. "VLSI Architectures for Neural Networks", *IEEE Micro*, December 1989, **9**(6), 8-27.