

Redes Neurais Artificiais Invariantes à Rotação de 10° e a Translação Utilizada no Reconhecimento de Figuras Geométricas

Aderlon Marcelino Queiroz¹, Edna Lúcia Flôres²

¹Departamento de Ciência da Computação - UNICEUB
70.790-075 – Brasília – DF

²Faculdade de Engenharia Elétrica – UFU
38.400-902 – Uberlândia – MG

E-mails: amqueiroz@bol.com.br, edna@ufu.br

Abstract

This paper deals with the recognition of geometric figures through the use of Neural Networks Artificial (NNA's). Before the geometric figures are trained and can be recognized by the NNA, in order to archive a satisfactory result, these figures are submitted to a translation and rotation of 10° by an invariant network. This invariant network is constituted of a single neuron ADALINE, which has its synaptic weights translated and rotated. This article is composed of five main aspects, firstly is done a general revision on the works that use NNA's and shown pattern recognition techniques for this purpose. The second item, shows the method used to make the ADALINE neuron invariant to a translation and rotation of 10° . The third aspect, deals briefly with the NNA training and recognition algorithms. The fourth item, shows the results, and finally the fifth presents the conclusion on this work.

1. Introdução

Vários centros de pesquisas e Universidades estão tentando desenvolver técnicas de reconhecimento que tenham um nível de desempenho praticamente igual ao do ser humano. O principal motivo do desenvolvimento de sistemas de reconhecimento são as aplicações científicas e econômicas, que tais sistemas apresentam. Entre estes sistemas pode-se citar, os sistemas de: reconhecimento de código de endereçamento postal, leitura automática de cheques bancários, tarifação telefônica, leitura automática para cegos, reconhecimento de figuras geométricas, etc. A obtenção de um software que consiga fazer o papel do homem pode representar uma grande economia de tempo e um bom ganho para as empresas.

Os métodos de reconhecimento de padrões convencionais requerem que os dados de entrada apresentem posição, orientação e tamanho similares a um determinado padrão. Na prática, porém, os dados de entrada podem apresentar diferenças de tamanho, posição ou rotação em relação a um padrão pré-determinado. Além disso, os dados de entrada podem ser corrompidos por ruídos. Os métodos convencionais

de reconhecimento, o casamento de padrões, por exemplo, não trabalham bem sob essas condições.

A mudança de escala ocorre quando as dimensões do padrão utilizado no treinamento são diferentes das dimensões do padrão utilizado no reconhecimento. A translação ocorre quando o padrão utilizado no treinamento encontra-se localizado em uma posição distinta à posição do padrão utilizado no reconhecimento. A rotação ocorre quando os padrões utilizados no treinamento e no reconhecimento apresentam inclinações diferentes. Com o objetivo de minimizar essas diferenças, a RNA deve ser insensível à mudança de escala, translação e rotação.

WIDROW et al [1] desenvolveram um sistema para o reconhecimento de padrões utilizando uma rede invariante constituída de um único neurônio ADALINE e uma RNA classificadora. A rede invariante é treinada ou projetada para produzir um conjunto de saídas que é insensível à translação e à rotação de 90° do padrão de entrada.

FUKUMI et al [2] utilizaram o mesmo princípio de uma rede invariante composta de um único neurônio ADALINE e uma RNA classificadora de padrões desenvolvido por Widrow para tornar a rede invariante somente à rotação.

Um dos algoritmos mais importantes na área de reconhecimento de padrões é o “Neocognitron”, desenvolvido por FUKUSHIMA et al [3,4]. Este sistema de reconhecimento possui múltiplas camadas de neurônios que são agrupadas em células de neurônios.

2. Rede Invariante à Translação e a Rotação de 10°

2.1. Translação

2.1.1. Translação cima-baixo

O princípio da translação cima-baixo do padrão de treinamento consiste em realizar um movimento para buscar a melhor situação para o reconhecimento. O movimento de busca é realizado no conjunto de pesos sinápticos de um único neurônio ADALINE.

A translação cima-baixo consiste em promover no conjunto de pesos sinápticos do neurônio ADALINE, um movimento de rotação com rolagem no sentido das linhas. A rotação com rolagem no sentido das linhas consiste em fazer com que os elementos de uma determinada linha sejam transladados para a linha inferior, com exceção dos elementos da última linha que serão transladados para a primeira linha da matriz. Optou-se por utilizar a rotação por rolagem no sentido das linhas iniciando em cima e indo para baixo. Pode-se realizar a rotação no sentido contrário a esse, contudo, os resultados serão diferentes. A figura 1 ilustra a rotação com rolagem no sentido das linhas realizada em uma matriz de dimensão 3 x 3.

1	2	3	7	8	9
4	5	6	1	2	3
7	8	9	4	5	6
Antes da rotação com rolagem no sentido das linhas			Após a rotação com rolagem no sentido das linhas		

Figura 1: Representação matricial da rotação com rolagem no sentido das linhas.

Pode-se verificar na figura 1, que os elementos da primeira linha foram deslocados para segunda linha, os elementos da segunda linha ocuparam a terceira linha e os elementos da terceira linha foram deslocados para primeira linha. Pode-se observar também nessa figura, que não ocorreu troca de elementos entre as colunas.

WIDROW et al [1] desenvolveu a operação de rotação da matriz de pesos sinápticos de um único neurônio ADALINE para a obtenção de uma rede invariante à translação cima-baixo. A figura 2 ilustra a representação esquemática da translação cima-baixo dos pesos sinápticos de um único neurônio ADALINE idealizado por Widrow.

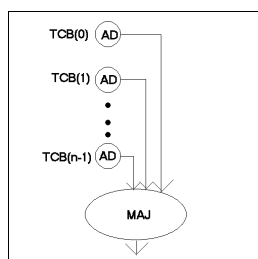


Figura 2: Representação esquemática da translação cima-baixo.

Na figura 2, o rótulo AD representa o neurônio ADALINE, o rótulo MAJ a função MAJ (“MAJority vote-taker”, Maioria dos Votos Dados) realiza a soma dos resultados obtidos pelos vários elementos AD, o rótulo TCB(x) (Translação Cima-Baixo) representa a rotação com rolagem no sentido das linhas promovida no conjunto de pesos sinápticos do único neurônio ADALINE, e x é a quantidade de linhas que foram

rotacionadas na matriz de pesos sinápticos do neurônio ADALINE.

2.1.2. Translação esquerda-direita

O princípio da translação esquerda-direita do padrão de treinamento consiste em realizar um movimento para buscar a melhor localização para o reconhecimento. O movimento de busca é realizado no conjunto de pesos sinápticos de um único neurônio ADALINE.

A translação esquerda-direita consiste em promover no conjunto de pesos sinápticos, do neurônio ADALINE, um movimento de rotação com rolagem no sentido das colunas. A rotação com rolagem no sentido das colunas consiste em fazer com que os elementos de uma determinada coluna sejam transladados para a coluna do lado direito, com exceção dos elementos da última coluna que são transladados para a primeira coluna da matriz. Optou-se por utilizar a rotação por rolagem no sentido das colunas iniciando à esquerda e indo para a direita. Pode-se realizar a rotação no sentido contrário a esse, contudo, os resultados da rotação serão diferentes. A figura 3 ilustra a rotação com rolagem no sentido das colunas realizada em uma matriz 3 x 3.

1	2	3	3	1	2
4	5	6	6	4	5
7	8	9	9	7	8
Antes da rotação por rolagem no sentido das colunas			Após a rotação por rolagem no sentido das colunas		

Figura 3: Representação matricial da rotação com rolagem no sentido das colunas.

Na figura 3, pode-se observar que os elementos da primeira coluna ocuparam a segunda coluna, os elementos da segunda coluna foram para a terceira coluna, e os elementos da terceira coluna ocuparam o lugar da primeira coluna que estava vazia. Deve-se observar também que não ocorreu troca de elementos entre as linhas e não houve deslocamento dos elementos da matriz no sentido das linhas.

WIDROW et al [1] desenvolveram um esquema de rotação da matriz de pesos sinápticos de um único neurônio ADALINE para obtenção de uma rede invariante à translação esquerda-direita. A figura 4 ilustra a representação esquemática da translação esquerda-direita dos pesos sinápticos deste neurônio.

Na figura 7, pode-se verificar que os elementos da primeira linha foram transpostos para a terceira coluna, os elementos da segunda linha foram transpostos para a segunda coluna e os elementos da última linha foram transpostos para a primeira coluna. Pode-se verificar também nessa figura que quando foi realizado a transposição dos elementos da matriz, a ordem dentro da linha foi mantida dentro da coluna.

WIDROW et al [1] desenvolveram a operação de transposição no sentido horário na matriz de pesos sinápticos de um único neurônio ADALINE seguindo o mesmo procedimento mostrado nas subseções anteriores, com o objetivo de obter uma rede invariante à rotação de 90° . A figura 8 ilustra o esquema da laje que torna a RNA invariante à rotação de 90° .

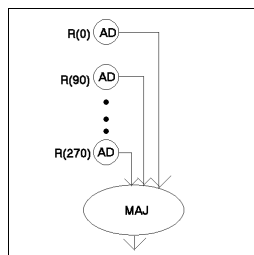


Figura 8: Representação esquemática da rotação de 90° .

Os rótulos mostrados na figura 8 são os mesmos mostrados nas figuras 2, 4 e 6, com exceção do rótulo $R(z)$, que representa a transposição no sentido horário do conjunto de pesos sinápticos do neurônio único ADALINE e z representa o ângulo de transposição.

2.2.2. Rotação de 10°

Na seção 2.2.1, foi explicada a teoria da rede invariante à rotação de 90° . Contudo, pode-se verificar a limitação que tal rede invariante apresenta, pois o intervalo de rotação da figura geométrica é muito grande, o que limita em muito a ação dessa rede. Para resolver este problema, foi utilizado neste artigo o princípio exposto por FUKUMI et al [2], onde a rede invariante à rotação de 90° é substituída por uma rede invariante à rotação para qualquer ângulo.

O trabalho de Fukumi trata apenas de uma única rede invariante à rotação para qualquer ângulo. O trabalho de WIDROW et al [1] utilizaram uma matriz quadrada, que limita o mecanismo da rede invariante à rotação. A solução nesse artigo foi utilizar as matrizes circulares mostradas por Fukumi em seu trabalho para obter uma rede invariante à rotação para qualquer ângulo.

A matriz circular é uma matriz diferente da matriz comum, pois a matriz comum é composta de linhas e colunas, a matriz circular é composta de raios e segmentos, tendo os contornos arredondados, e por isto é ideal para construir uma rede invariante à rotação QUEIROZ [6, 7]. Para obter uma rede invariante à rotação utilizando as matrizes circulares, é necessário

primeiro transformar a matriz comum em circular, essa transformação é chamada de circularização. Então é realizado a transposição no sentido horário e por fim a matriz circular é transformada em matriz comum, essa transformação é chamada de quadratização QUEIROZ [6, 7].

Pode-se observar que as redes invariantes à rotação propostas por WIDROW et al [1] e FUKUMI et al [2] são semelhantes a princípio, contudo, há uma diferença substancial. Com base na quantidade de operações de rotação que ambas podem fazer. Na rede invariante de Widrow sempre são 4 operações, já na rede invariante de Fukumi são 36 operações, se a ordem da matriz quadrada for igual a 10.

2.3. Translação e rotação de 10°

Este artigo utiliza os princípios desenvolvidos por WIDROW et al [1] e FUKUMI et al [2] para criar uma rede invariante à translação e rotação de 10° , ilustrada na figura 10.

Na figura 10 pode-se observar que a rede invariante é composta de um conjunto de lajes, cada laje recebe a matriz de pesos sinápticos do único neurônio ADALINE rotacionado pelo mecanismo de Fukumi e a translação completa conforme o método utilizado por Widrow.

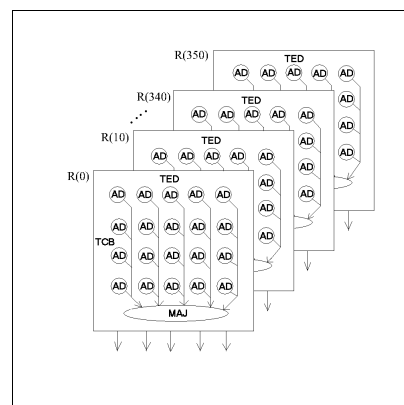


Figura 10 Representação esquemática da rede invariante à translação e rotação de 10° .

Na figura 10 não são mostrados os processos de circularização e quadratização; deve-se subentender que para cada laje de rotação, o conjunto de pesos sinápticos do único neurônio ADALINE é transformado em uma matriz circular, rotacionado por um ângulo, transformado em uma matriz quadrada, para então ser aplicado o processo de translação completa.

3. Redes Neurais Artificiais

Neste trabalho foram utilizadas quatro RNAs diferentes. A primeira RNA não possui camada oculta e foi treinada pela regra Delta, a segunda não possui

camada oculta e é treinada pela regra MADALINE II, a terceira RNA possui uma camada oculta e é treinada pela regra Back-Propagation, e a última RNA possui duas camadas ocultas e é treinada pela regra Back-Propagation.

A primeira RNA utiliza a regra delta para correção dos pesos sinápticos do neurônio ADALINE mostrada na eq. (1), onde w é o conjunto de pesos sinápticos da camada de saída. Onde o parâmetro η é a constante de aprendizagem do neurônio x^d é o vetor do padrão que está sendo treinado e y^d é o vetor desejado como resposta na saída da RNA

$$w(k+1) = w(k) + \eta \sum_{l=1}^L \delta_l x_l^d \quad (1)$$

onde: $\eta' = 2\eta$; $\delta_l = y_l^d - y_l$.

A segunda RNA utiliza a regra MADALINE II para correção dos pesos sinápticos do neurônio ADALINE. A eq. (2) mostra esta regra para o treinamento das RNAs.

$$w(k+1) = w(k) + \sum_{l=1}^L \frac{\eta \delta_l x_l^d}{\sqrt{\sum_{i=1, n} x_{i,l}^2}} \quad (2)$$

Onde no denominador da eq. (2) é colocado o módulo do vetor de entrada do padrão a ser reconhecido.

As duas últimas RNAs são treinadas pelo mesmo algoritmo de correção dos pesos sinápticos, contudo, existem algumas alterações para quando a RNA possui mais de uma camada oculta. A eq. (3) é o mecanismo de correção na camada de saída e a eq. (4) na camada oculta.

$$W_{ij}(n+1) = W_{ij}(n) + \eta' \delta_i^1 Z_j \quad (3)$$

$$V_{kl}(n+1) = V_{kl}(n) + \eta' X_l^d Z_k (1 - Z_k) \sum_{m=1, N_S} \delta_m^1 W_{mk} \quad (4)$$

onde: $\eta' = 2\eta$; $\delta_i^1 = (Y_i^d - Y_i) f'(\zeta_i) = (Y_i^d - Y_i) Y_i (1 - Y_i)$.

4. Resultados

Os resultados obtidos utilizando-se as quatro RNA's citadas no item 3 foram encontrados após elas terem sido treinadas para todas as possibilidades de alterações de translação e rotação. Nos testes realizados nas figuras geométricas ocorreram alterações de deslocamento e inclinação. O ruído adicionado as figuras geométricas possui distribuição uniforme, variando de 10 % a 50 %. As figuras geométricas que

foram treinadas possuíam uma porcentagem de 5 % de ruído adicionado. No reconhecimento, as figuras geométricas eram transladas e rotacionadas por valores aleatórios, para finalmente serem aplicadas as RNA's para serem reconhecidas.

As figuras 11 a 14 ilustram os resultados obtidos para cada um dos quatro tipos de RNA's quando elas são treinadas por 25 figuras geométricas e são utilizadas e a mesma quantidade de figuras para serem reconhecidas pelas redes.

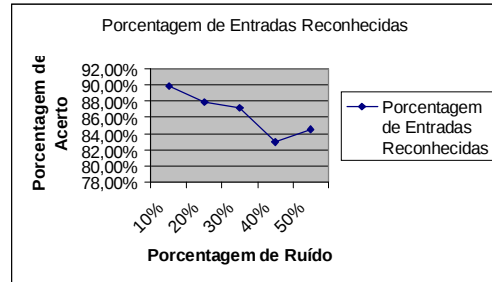


Figura 11: Resultados obtidos para uma RNA sem camada oculta, treinada pela regra Delta

Os valores obtidos na figura 11 para a RNA sem camada oculta, treinada pela regra delta diminuíram com o aumento da porcentagem de ruído adicionado as figuras geométricas, com exceção do resultado obtido para a porcentagem de ruído adicionado às figuras geométricas de 50 % que foi maior do que a porcentagem de acerto quando foi adicionado a porcentagem de ruído de 40 %; isto ocorreu devido a característica do ruído ser aleatório e o comportamento não linear dessa RNA.

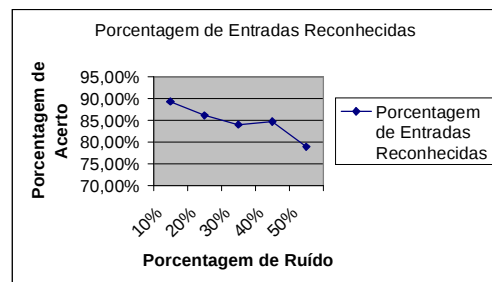


Figura 12: Resultados obtidos para uma RNA sem camada oculta, treinada pela regra MADALINE II

De forma semelhante ao teste realizado anteriormente, os resultados obtidos na figura 12 para a RNA sem camada oculta, treinada pela regra MADALINE II decresceram com o aumento da porcentagem de ruído adicionado as figuras geométricas. Com exceção da porcentagem de acerto obtida quando a porcentagem de ruído adicionado era de 40 % o resultado obtido foi maior do que quando a porcentagem de ruído adicionado as figuras geométricas foi de 30 %. Isto deve-se ao fato da natureza do ruído adicionado às figuras geométricas e a

não linearidade da rede neural treinada por esse método.

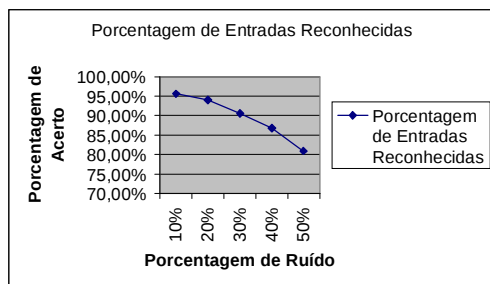


Figura 13: Resultados obtidos para uma RNA com 1 camada oculta, treinada pela regra “Back-Propagation”

Na figura 13, pode-se observar que a porcentagem de acerto diminuiu com o aumento da porcentagem de ruído adicionado às figuras geométricas, para a RNA com 1 camada oculta treinada pela regra “Back-Propagation”.

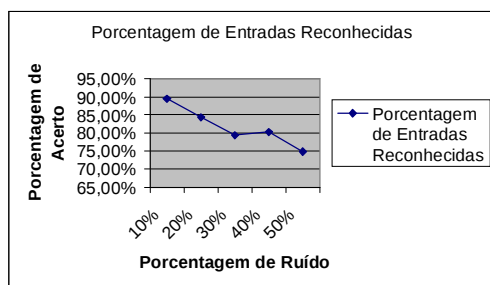


Figura 14: Resultados obtidos para uma RNA com 2 camadas ocultas, treinada pela regra “Back-Propagation”

Na figura 14, pode-se verificar que os resultados obtidos diminuíram com o aumento da porcentagem de ruído acrescentado às figuras geométricas, para a RNA com 2 camadas ocultas, treinada pela regra “Back-Propagation”. Contudo, verifica-se também que o resultado obtido para a porcentagem de acerto quando foi adicionado a porcentagem de ruído de 40 % foi maior do que o resultado obtido para a porcentagem de ruído de 30 %. O que não era o esperado, pois com uma quantidade maior de camadas ocultas os resultados tornaram-se mais estáveis, porém, como o comportamento da RNA não é linear e o ruído aleatório influí também nesse resultados.

Pode-se verificar que os resultados obtidos nas figuras 11, 12, 13 e 14 ficaram acima do 70 % de acerto. A RNA que obteve os melhores resultados foi a que possui 1 camada oculta, treinada pela regra “Back-Propagation”, com valores mais próximos entre si.

5. Conclusão

Este artigo mostrou os algoritmos para tornar uma RNA invariante à translação e rotação de 10°. Primeiro foi feito um resumo dos vários métodos para solucionar

o problema de invariância, em seguida foram expostos os mecanismos invariantes à translação e à rotação. Depois foram explicados alguns métodos para a correção dos pesos sinápticos e finalmente, foram mostrados os resultados obtidos nos testes realizados.

Os resultados obtidos utilizando-se os quatro tipos de RNA's foram maiores do que 70 %. A RNA que obteve os melhores resultados foi a que possui 1 camada oculta, treinada pela regra “Back-Propagation”, com valores mais próximos entre si.

Referências

- [1] WIDROW, B., et al. Layered Neural Nets for Pattern Recognition. IEEE Transaction on Acoustics, Speech, and Signal Processing, v. 36, n. 7, p. 1109-1118, jul, 1988.
- [2] FUKUMI, M., et al. Rotation-Invariant Neural Pattern Recognition System with Application to Coin Recognition. IEEE Transaction on Neural Networks, v. 3, n. 2, p. 272-279, mar. 1992.
- [3] FUKUSHIMA, K. e MIYAKE, S. NEOCOGNITRON: A New Algorithm for Pattern Recognition Tolerant of Deformations and Shifts in Position. Pattern Recognition, v. 15, n. 6, p. 455-469, 1982.
- [4] FUKUSHIMA, K., et al. NEOCOGNITRON: A Neural Networks Model for a Mechanism of Visual Pattern Recognition. IEEE Transaction on Systems, Man, and Cybernetics, v. SMC-13, n. 5, p. 826-834, set/out, 1983.
- [5] FLÔRES, E. L. et al. Reconhecimento de Dígitos Manuscritos Insensíveis a Mudança de Escala, Translação e Rotação Usando Redes Neurais. Revista Ciência & Engenharia. EDUFU. Uberlândia, MG. Ano 5, No 1 p. 83-89 jan/jun, 1995.
- [6] QUEIROZ, A. M. e FLÔRES, E. L. Reconhecimento de Figuras Geométricas Utilizando-se Redes Neurais Artificiais Invariantes à Rotação de 10o e a Translação. IV Simpósio Brasileiro de Automação Inteligente, Escola Politécnica da Universidade de São Paulo. São Paulo, SP. p. 586-591, set, 1999.
- [7] QUEIROZ, A. M. e FLÔRES, E. L. Reconhecimento de Figuras Geométricas Utilizando Redes Neurais Artificiais Invariantes à Translação e Rotação de 10o. Revista Ciência & Engenharia. Editora da Universidade Federal de Uberlândia – EDUFU. Uberlândia, MG. Ano 9, No 1. p. 115-122, jan/jun, 2000.
- [8] TEIXEIRA, E. P. Controle de Sistemas Não Lineares Através de Redes Neurais. Campinas: Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica, Departamento de Engenharia de Computação e Automação Industrial, 1991. 191p.