

The logo for OMNIA EXATAS features the word "OMNIA" in large, bold, yellow 3D-style letters. The letter "I" is replaced by a vertical yellow bar containing the word "EXATAS" written vertically in black capital letters.

Faculdades Adamantinenses Integradas (FAI)

[www.fai.com.br](http://www.fai.com.br)

LIMA, Fernando Parra dos Anjos; MINUSSI, Carlos Roberto. Usando uma rede neural de kohonen para reconhecimento de padrões de som. Omnia Exatas, v.4, n.2, p.19-30, 2011.

# **USANDO UMA REDE NEURAL DE KOHONEN PARA RECONHECIMENTO DE PADRÕES DE SOM**

*USING A KOHONEN NEURAL NETWORK FOR PATTERN RECOGNITION SOUND*

**Fernando Parra dos Anjos Lima**

Mestrando em Engenharia Elétrica – FEIS – UNESP – Ilha Solteira.

**Carlos Roberto Minussi**

Professor Titular – DEE – FEIS – UNESP – Ilha Solteira.

## **RESUMO**

O objetivo deste artigo é utilizar uma rede neural artificial para reconhecimento de padrões de som. Em específico, tons DTMF (Dual Tone Multi Frequency), (MEDEIROS, 2006), ou melhor, os sinais sonoros do teclado telefônico. O tipo de rede neural utilizada para realizar esta tarefa é a rede neural de Kohonen, ou como também são conhecidos, os mapas auto-organizáveis de Kohonen. (KOHONEN, 1988 – 1990). Neste trabalho pretende-se realizar vários tipos de testes com a rede neural de Kohonen, fazendo modificações na quantidade de neurônios presentes na grade de reconhecimento, número de iterações de treinamento, e principalmente os dados para treinamento e teste da rede. Assim após realizar todos os tipos de teste com a metodologia proposta pode-se analisar os resultados a fim de qualificar e classificar o rendimento deste sistema inteligente em relação à tarefa realizada. Lembrando que como o trabalho é realizado com som, inicialmente é feito um pré-processamento no arquivo de som, que armazena uma série temporal que representa um tom ou uma sequência de tons DTMF. Nesta série temporal aplica-se a transformada de Fourier para obter a resposta em frequência, para somente assim apresentar a rede neural de Kohonen.

**Palavras-chaves:** Redes Neurais artificiais, Kohonen, Reconhecimento de padrões, Tons DTMF.

## **ABSTRACT**

The aim of this paper is to use an artificial neural network for pattern recognition of sound. In specific DTMF (Dual Tone Multi Frequency), (MEDEIROS, 2006), or better, the sound signals from the telephone keypad. The type of neural network used for this task is the Kohonen neural network, or they are also known, self-organizing maps of Kohonen. In this work we intend to perform various tests with the Kohonen neural network, making changes in the amount of neurons found in the grid of recognition, number of iterations of training, and especially the training and test data to the network. (KOHONEN, 1988 – 1990). So after doing all kinds of test with the proposed methodology can analyze the results in order to qualify and classify the performance of this intelligent system to the task performed. Remember that work is done with sound, is initially made a pre-processing the sound file, which stores a time series representing a tone or a DTMF tone sequence. This time series

applies the Fourier transform to obtain the frequency response, the only way to present the Kohonen neural network.

**Key-words:** Artificial Neural Networks, Kohonen, pattern recognition, DTMF tones.

## INTRODUÇÃO

Redes neurais artificiais (RNAs) (HAYKIN, 1994) são conceitos baseados no mecanismo de funcionamento do cérebro emulados em software e/ou hardware. São capazes de adquirir conhecimento, ou seja, apreenderem com a experiência. As RNAs são modeladas através de várias unidades de neurônios ou unidades de processamento, dispostos em fileiras ou camadas, formando um arranjo de interligações, estas interligações são conexões constituídas por pesos, ou sinapses, que são o ponto de calibração das unidades de processamento. (BRAGA, 2007), (HAYKIN, 1994), (JAIN, 1996). O processo de calibração dos pesos é chamado de treinamento ou aprendizado, o mesmo deve ser elaborado no modo off-line do sistema, ou como pode-se dizer, antes que o sistema entre em funcionamento. Uma vez ajustada, empregando-se algum tipo de algoritmo de adaptação dos referidos pesos, a rede, a princípio, estará apta para efetuar diagnósticos e emitir um grau de precisão satisfatório no reconhecimento de padrões não utilizados no seu processo de aprendizado. (HAYKIN, 1994), (KARTALOPOULOS, 1988).

Levando em consideração a importância das análises de rotinas em tempo real, este artigo visa colocar em prática o conhecimento obtido em um estudo embasado nos conceitos e arquiteturas neurais artificiais e suas aplicações de modo geral e, em específico, o reconhecimento de padrões de som. Será utilizada a plataforma e linguagem de programação Borland Delphi 7.0 para implementação de um sistema para reconhecimento de som com redes neurais artificiais. Serão utilizados alguns componentes que simulam a arquitetura e toda a técnica da rede SOM (Self Organizing Maps), ou mais conhecidos como mapas auto-organizáveis propostos por Kohonen. (KOHONEN, 1988 – 1990), (NASCIMENTO, 2004).

Na rede de reconhecimento de som, utiliza-se o conceito de tons DTMF (Dual Tone Multi Frequency) para gerar amostras de som para utilização no processo de treinamento e reconhecimento, ou seja, em questão, têm-se tons sonoros reproduzidos a partir de um teclado de telefone.

Toda Informação que será apresentada para a rede neural deve passar por um pré-processamento, pois o sistema faz praticamente o reconhecimento de frequências, os arquivos apresentados para a rede neural são constituídos de sinais sonoros, que na verdade estão dispostos em uma série temporal. O método utilizado para obter a resposta em frequência de cada uma das séries temporais, é a transformada de Fourier. (MEDEIROS, 2006).

## MATERIAL E MÉTODOS

Inicialmente faz-se uma abordagem no conceito da rede neural de Kohonen e seu algoritmo de treinamento.

### **O modelo de Kohonen e seu algoritmo**

A rede de Kohonen é composta por um reticulado (grade) bidimensional e um conjunto de vetores pesos, fixados inicialmente em valores aleatórios entre 0 e 1,  $w_j = [w_{1j}, w_{2j} \dots w_{nj}]^T=1$ , Assim cada neurônio sobre a grade esta conectado a entrada da rede, através do vetor de entrada conforme a figura 1. (KOHONEN, 1988 – 1990).

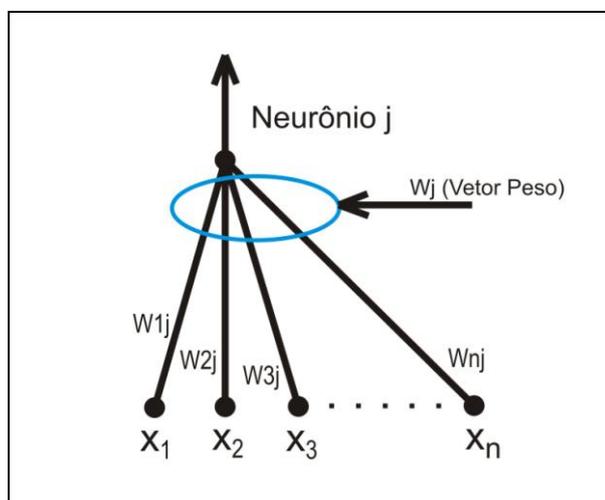


Figura 1: Conexão dos vetores de entrada com a rede de Kohonen

Deste modo a saída  $y_j$  (atividade do  $j$ -ésimo neurônio) pode ser calculada da seguinte forma:

$$y_j = \langle W_j, x \rangle \quad (1)$$

Sendo:

$y_j$  = Atividade (saída) do  $j$ -ésimo neurônio sobre a grade de Kohonen.

$$W_j = [w_{1j}, w_{2j} \dots w_{nj}]^T$$

$$X = [x_1, x_2 \dots x_n]^T$$

O treinamento é um processo através do qual os parâmetros de uma rede neural artificial são ajustados utilizando uma forma contínua de estímulos. No caso do treinamento por competição, dado um estímulo à rede (vetor de entrada), as unidades de saída disputam entre si para serem ativadas. A unidade vencedora terá seus pesos atualizados no treinamento. O valor atualizado dos pesos se dá através da seguinte fórmula: (KOHONEN, 1988 – 1990).

$$W_{h+1} = W_h + \alpha(x - W_h) \quad (2)$$

Sendo:

$\alpha$  = Taxa de treinamento ( $0 \leq \alpha \leq 1$ )

H = índice de atualização.

O neurônio vencedor é escolhido como sendo aquele que apresentar maior atividade ( $y$ ) sobre a grade de Kohonen, ou seja:

$$NV = \max_i \{ \langle W_i, x \rangle \} \quad (3)$$

$$NV = \min_i \{ \|x - W_i\| \} \quad (4)$$

Sendo: NV = neurônio vencedor

$\| \cdot \|$  = norma Euclidiana.

Ao neurônio vencedor atribui-se índice  $k$ . Usando-se, então a regra de adaptação, Kohonen propôs que o ajuste da rede neural em que os pesos são ajustados considerando-se vizinhança em torno do neurônio vencedor, como mostrado na figura 2. (KOHONEN, 1988 – 1990).

Sendo:

$NC_k(S_i) =$  Vizinhança do neurônio vencedor  $k$ , referente a região  $S_i$ , em que  $S_1 \subset S_2 \subset S_3$

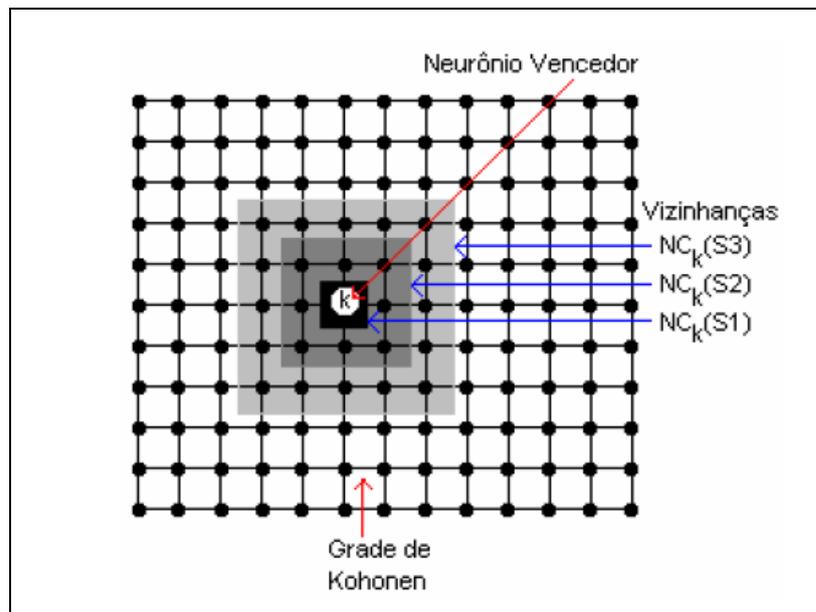


Figura 2: Vizinhanças do neurônio vencedor

Então de forma conceitual apresentam-se os passos utilizados para o treinamento da rede de Kohonen utilizada neste artigo: (KOHONEN, 1988 – 1990).

**Passo I:** Inicie os pesos da rede de Kohonen entre a entrada e a grade. Estes pesos podem ser gerados randomicamente com valores compreendidos entre 0 e 1.

**Passo II:** Normalize os vetores de pesos.

**Passo III:** Normalize todos os vetores padrões de entrada.

**Passo IV:** Apresente um novo vetor padrão de entrada.

**Passo V:** Calcule a distância ou o produto interno para todos os neurônios sobre a grade de Kohonen.

**Passo VI:** Encontre o neurônio vencedor (Equação (3) ou 4). Designar o neurônio vencedor pelo índice  $k$ .

**Passo VII:** Adaptar os vetores pesos do neurônio vencedor e dos demais neurônios contidos na vizinhança escolhida (vide Figura 2).

$$W_i(h+1) = w_i(h) + \alpha(x - w_i(h)) \quad (5)$$

$i \in N_k$ .

**Passo VIII:** Renormalize todos os vetores de pesos adaptados no passo VII. Este procedimento se faz necessário tendo em vista que o vetor peso, após a adaptação, não possui comprimento unitário, sendo, portanto, necessário torná-lo (com comprimento unitário).

**Passo IX:** Retornar ao Passo 4. Este procedimento deve ser repetido, considerando-se um número fixo de iterações, ou até que as variações (módulos) dos pesos sejam inferiores a uma certa tolerância preestabelecida.

## Os tons DTMF (Dual Tone Multi Frequency)

Os tons DTMF (MEDEIROS, 2006) são pares de frequências montados para identificação de teclas discadas no teclado telefônico. Esses pares estão definidos conforme é mostrado na Tabela 1.

Tabela 1: Pares de frequência dos Tons DTMF.

Frequência (Hz)	1209	1336	1477	1633
697	1	2	3	A
770	4	5	6	B
852	7	8	9	C
941	*	0	#	D

Assim, em cada tecla estão presentes duas frequências. Para a tecla 1, por exemplo, existem a frequência mais baixa, de 697Hz, e a mais alta, de 1209Hz.

## Funcionamento do sistema

O sistema inteligente implementado para este trabalho segue a metodologia ilustrada na figura 3 a seguir.

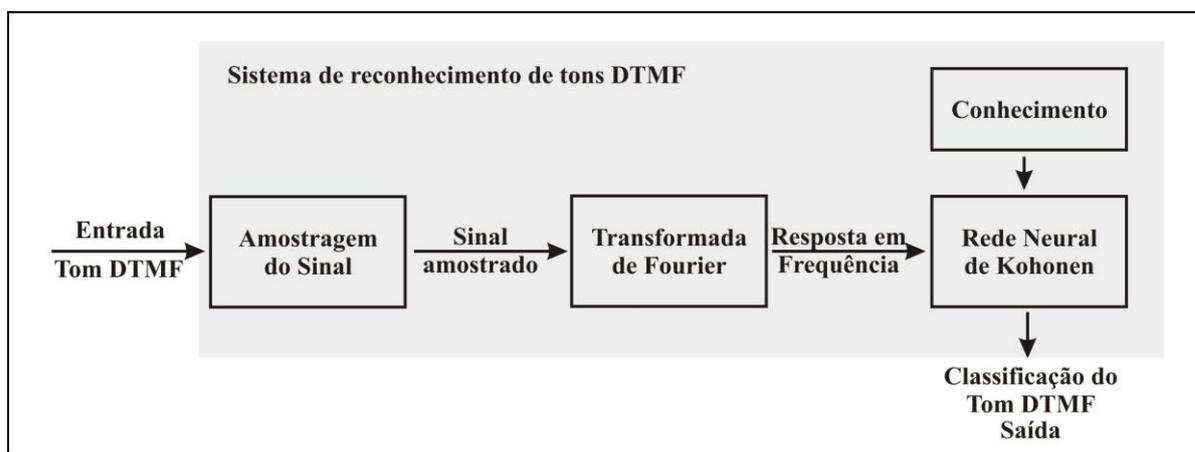


Figura 3: Funcionamento do sistema inteligente.

Como se podem observar no diagrama apresentado na figura 3, o arquivo com o tom DTMF é à entrada do sistema, e a saída do sistema é a classificação deste tom DTMF em relação ao som que ele representa em um teclado telefônico.

De forma breve e clara faz-se uma descrição do funcionamento do sistema.

Como se têm uma série temporal (sinal analógico) armazenada em um arquivo WAVE na entrada do sistema, deve-se aplicar um pré-processamento na informação para obter a representação em frequência daquele sinal. Para isto utilizam-se alguns conceitos de processamento de sinais digitais, inicialmente faz-se a amostragem do sinal, e após ter um sinal amostrado aplica-se a transformada de Fourier, método que permite identificar as frequências fundamentais presentes em uma série temporal. (MEDEIROS, 2006).

De maneira ilustrativa apresenta-se a seguir as figuras 4 e 5, onde são apresentados dois sinais, um que representa uma série temporal (sinal analógico) antes de qualquer

processamento, e outro que representa a resposta em frequência obtida após a aplicação da transformada de Fourier. (MEDEIROS, 2006).

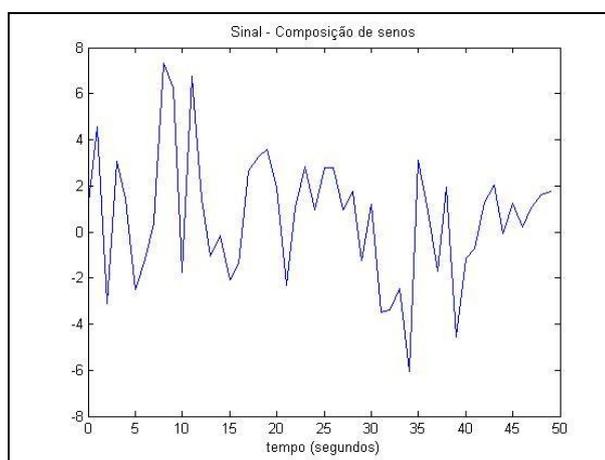


Figura 3: Sinal composto por uma soma de senos de 50 e 120 Hz.

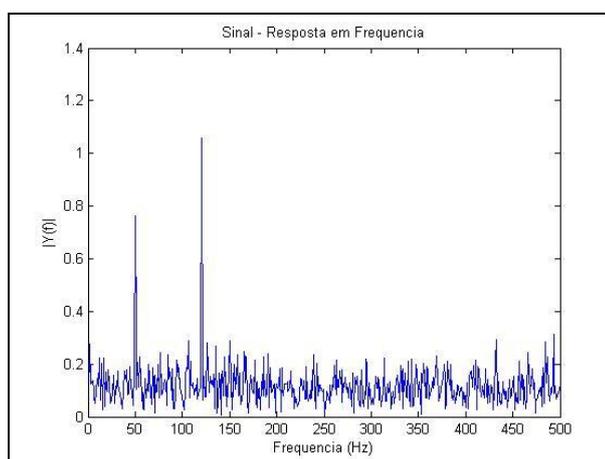


Figura 4: Resposta em frequência do sinal anterior.

Nota-se que na figura 3 tem-se um sinal qualquer, que é composto por uma soma de senos, um de 50 Hz e outro de 120 Hz. Na figura 4 pode-se observar, de forma clara, que se tem o espectro de frequência da transformada rápida de Fourier onde são identificados dois picos que são as frequências fundamentais deste sinal, um dos picos representa frequência de 50Hz e o outro a frequência de 120Hz, assim ficando bem claro que a resposta em frequência do sinal trás a identificação das frequências fundamentais presentes na série temporal analisada. A transformada de Fourier implementada neste trabalho é uma variante do algoritmo da transformada discreta de Fourier, assim torna-se mais rápido o processo denominado transformada rápida de Fourier (*Fast Fourier Transform*, ou FFT). (MEDEIROS, 2006). Para a rede neural de Kohonen, mais conhecida como rede SOM, utiliza-se uma rede com duas camadas. Uma camada de entrada e outra camada de saída, contendo 30 neurônios na entrada e 10 na saída, um para cada dígito do teclado, respectivamente.

Para realizar o treinamento deste modelo emprega-se o algoritmo LVQ (*Linear Vector Quantization*) ou como é mais conhecido, algoritmo do neurônio vencedor. (KOHONEN, 1988 – 1990). Assim o neurônio ou região que representa o som que esta sendo informado para a rede será o vencedor para o momento, e assim é possível reconhecer e classificar os

Tons DTMF. Para este projeto utiliza-se a vizinhança do tipo retangular como descrito no algoritmo do neurônio vencedor. (KOHONEN, 1988 – 1990).

A figura 5 a seguir ilustra um exemplo de uma rede neural de Kohonen com seus representativos clusters de representação de padrões.

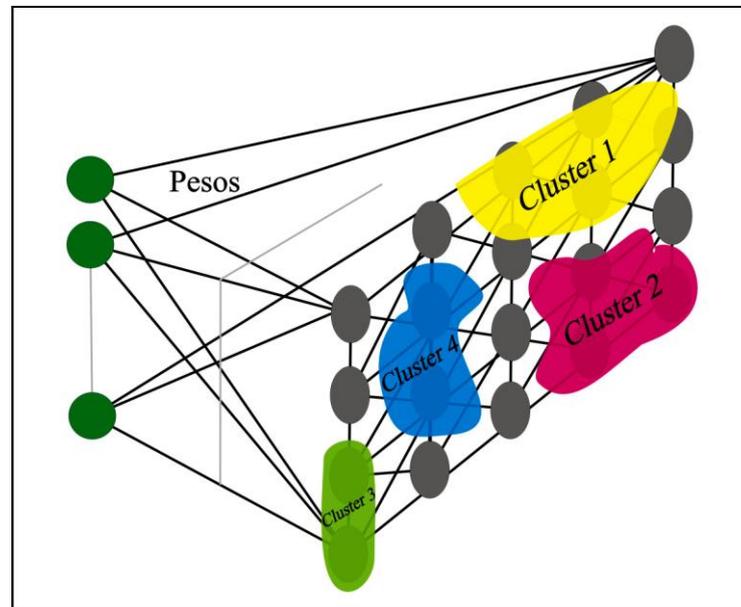


Figura 5: Ilustração de uma rede neural de Kohonen.

Cada cluster pode representar um tipo de padrão, assim a rede consegue identificar os padrões, durante o teste quando o neurônio vencedor pertencer à região em questão, evidentemente representa o padrão do cluster em questão e assim é feito o reconhecimento e classificação.

Através da metodologia apresentada são propostos vários testes para investigar o desempenho deste sistema.

### Testes propostos neste sistema

Para este sistema foram propostos os testes descritos a seguir:

- Teste I: Testar a rede neural de Kohonen com padrões do mesmo tipo que os padrões utilizados no seu processo de treinamento. (LIMA, 2010 – 2011).
- Teste II: Testar a rede neural de Kohonen com padrões diferentes dos padrões utilizados no seu processo de treinamento, ou seja, desconhecidos pelo sistema. (LIMA, 2010 – 2011).
- Teste III: Aumentar de 100 para 200 iterações a fase de treinamento e realizar o teste I e II novamente.
- Teste IV: Aumentar de 30 para 50 o número de neurônios presentes na grade de reconhecimento e realizar o teste I e II novamente.
- Teste V: Aumentar de 100 para 200 iterações a fase de treinamento, aumentar de 30 para 50 o número de neurônios presentes na grade de reconhecimento e realizar o teste I e II novamente.

Para todos os testes mencionados acima são utilizados os seguintes parâmetros:

- Número de iterações de treinamento: 100 iterações de treinamento. (exceto o Teste III).
- Conjunto de treinamento: Um arquivo WAVE contendo uma sequência com todos os tons DTMF, o arquivo “0123456789.wav”. A origem deste arquivo é advinda de um sintetizador de tons DTMF.
- Conjunto de testes: Para testar este sistema foram utilizados dois tipos de padrões, os padrões gerados em sintetizador, e os padrões gravados a partir de um teclado telefônico convencional. A tabela a seguir mostra o nome dos arquivos utilizados.

Tabela 2: Conjuntos de testes do sistema.

Conjunto 1: Tons Sintetizados	Conjunto 2: Tons Gravados
“357851.wav”	“421589.wav”
“784154.wav”	“852047.wav”
“985234.wav”	“975789.wav”
“124578.wav”	“971309.wav”
“369857.wav”	“654321.wav”
“478952.wav”	“423569.wav”
“123456.wav”	“874169.wav”
“457896.wav”	“784126.wav”
“25.wav”	“43.wav”
“8.wav”	“3.wav”

## RESULTADOS E DISCUSSÃO

Os resultados obtidos nos testes propostos para este artigo estão descritos a seguir:

**Teste I:** Foram testados 10 arquivos diferentes, todos eles gerados em um sintetizador, e tendo o mesmo tipo de padrão que foi utilizado no treinamento da rede neural, ou seja, o conjunto 1 de testes.

O resultado para este teste foi de 9 acertos e 1 erros (90% de acerto e 10% de erro).

**Teste II:** Foram testados 10 arquivos diferentes, todos eles gravados a partir de um teclado telefônico convencional, e todos sendo de um tipo de padrão que não foi utilizado no treinamento da rede neural, ou seja, o conjunto 2 de testes.

O resultado para este teste foi de 7 acertos e 3 erros (70% de acerto e 30% de erro).

**Teste III:** Após aumentar o número de iterações de treinamento foi testado o conjunto 1 de testes e o resultado para este teste foi de 10 acertos e 0 erros (100% de acerto e 0% de erro). Também utilizando o treinamento realizado foi testado o conjunto 2 de testes e o resultado para este teste foi de 9 acertos e 1 erros (90% de acerto e 10% de erro).

**Teste IV:** Após aumentar o número de neurônios na grade de reconhecimento e realizar o treinamento foi testado o conjunto 1 de testes e o resultado para este teste foi de 10 acertos e 0 erros (100% de acerto e 0% de erro). Também utilizando o treinamento realizado foi testado o conjunto 2 de testes e o resultado para este teste foi de 10 acertos e 0 erros (100% de acerto e 0% de erro).

**Teste V:** Após aumentar o número de iterações de treinamento, o neurônios na grade de reconhecimento e realizar o treinamento foi testado o conjunto 1 de testes e o resultado para este teste foi de 10 acertos e 0 erros (100% de acerto e 0% de erro). Também utilizando o treinamento realizado foi testado o conjunto 2 de testes e o resultado para este teste foi de 10 acertos e 0 erros (100% de acerto e 0% de erro).

Os resultados anteriormente obtidos estão ilustrados na figura 6 a seguir:

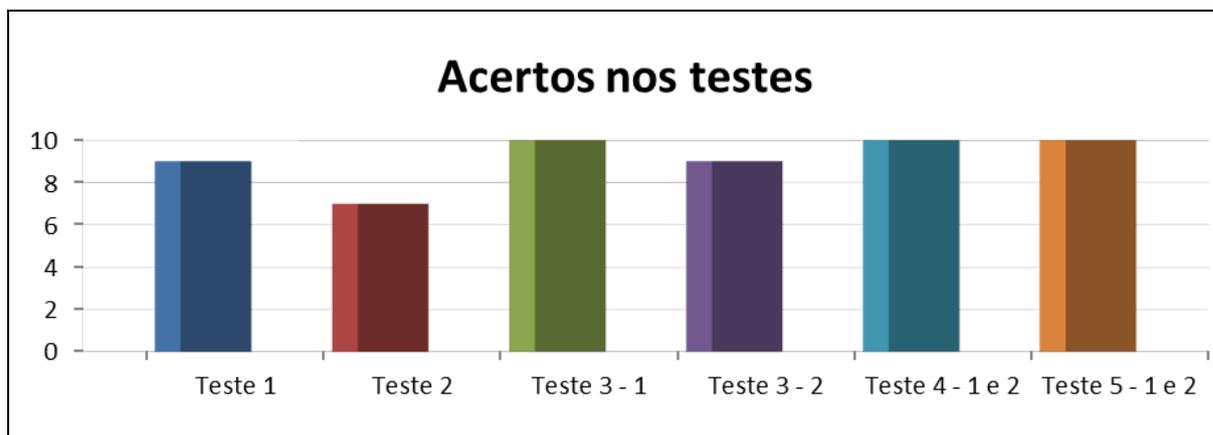


Figura 6: Resultados dos testes propostos.

Através dos resultados obtidos neste trabalho observa-se que a rede neural de Kohonen quando treinada com um bom som, tem capacidade de reconhecer com bastante eficiência sons que se diferem do padrão conhecido em seu treinamento, e de uma maneira eficaz a rede se adapta aos novos padrões e consegue realizar o reconhecimento da maioria das amostras. Nos testes I e II pode ser observado isto de maneira clara, onde existe um acerto de 9 amostras no teste I e 7 no teste II, as amostras do teste I são geradas em um sintetizador, portanto são puras, já as amostras do teste II são gravadas a partir de um telefone convencional, assim sendo sujeita a ruídos e interferências presentes do ambiente.

O aumento do número de iterações de treinamento melhora o desempenho do sistema para os dois conjuntos de teste, isso porque à medida que se aumenta o número de iterações de treinamento, melhora-se a auto-organização da rede neural de Kohonen, e em consequência disto o conhecimento é melhor absorvido pelo sistema, e o desempenho melhora.

O aumento do número de neurônios na grade de reconhecimento da rede de Kohonen proporciona uma melhora significativa no desempenho do sistema, pois os clusters de representação de cada tom DTMF serão maiores, e assim é possível que a rede extraia mais características importantes para a composição de cada padrão de representação dos tons.

No teste V observa-se a união do teste III e IV, ou melhor dizendo aumentando a segurança e confiabilidade da rede neural de Kohonen ao máximo, pois tanto o aumento de iterações de treinamento como o aumento de número de neurônios na grade de reconhecimento trazem melhorias significativas para o desempenho do sistema.

Assim observando todas as observações anteriores aconselha-se que no caso de usar este sistema para fins comerciais, ou outras situações, seja utilizada a configuração do teste V, pois a mesma elimina aproximadamente todas as chances de erro do sistema, tornando o mesmo seguro, flexível e principalmente adaptativo.

## CONCLUSÕES

Como citado por alguns autores, às redes neurais artificiais, podem ter um desempenho inicial insatisfatório, porém podem melhorar o seu desempenho, através do processo de treinamento, tentando proporcionar o menor erro possível. A auto-organização da rede neural de Kohonen se dá à medida que se aumenta o número de iterações de treinamento.

Na fase de testes deste sistema foi possível visualizar que o sistema assume um comportamento preciso e seguro em relação à tarefa a ser desempenhada. Isto pode ser notado quando o sistema acerta a maioria das amostras de tons DTMF que são de padrões desconhecidos pelo sistema. Com os novos testes realizados (Teste III, IV, V) foi possível identificar melhorias para o sistema, e acrescentar inovações na metodologia inicial apresentada.

Então para finalizar pode-se concluir que as redes neurais de Kohonen tem um desempenho muito bom quando são aplicadas ao reconhecimento de sons, e em específico os tons DTMF, e pode-se observar em todos resultados obtidos que este modelo é muito flexível e adaptativo tornando o processo de reconhecimento e classificação de tons DTMF muito seguro e confiável.

## AGRADECIMENTOS

Agradecemos primeiramente a Deus, as nossas famílias, e por fim um especial agradecimento a CAPES pelo apoio (concessão de bolsa de Mestrado).

## REFERÊNCIAS

BRAGA, A. P.; CARVALHO, A. L.; LUDERMIR, T. B. **Redes Neurais Artificiais - Teorias e Aplicações**, LTC, 2nd Edição, Rio de Janeiro, Brasil, 2007.

LIMA, F. P. A.; SILVA, J. C.; ESTEVAM, G. P.; MINUSSI, C. R. **Redes neurais artificiais aplicadas no reconhecimento de tons DTMF**. *Omnia Exatas*, v.3, n.1, p.79-89, 2010.

LIMA, F. P. A.; MINUSSI, C. R. **Usando uma rede neural de Kohonen para reconhecimento de padrões de som**. *Anais X conferência Brasileira de Dinâmica, Controle e Aplicações – Águas de Lindoia – SP – 2011*.

HAYKIN, S. **Neural Networks: A Comprehensive Foundation**, Prentice-Hall - Upper Saddle River - New Jersey – USA – 1994.

JAIN, A. K.; MOHIUDDIN, K. M. **Artificial Neural Networks: A Tutorial**, *IEEE Computer* – Vol. 29 – Nº. 3 – pp. 31-44 – March 1996.

KARTALOPOULOS, S. V. **Understanding Neural Networks and Fuzzy Logic: Basic Concepts and Applications**, IEEE Press – Piscataway – NJ – USA – 1996.

KOHONEN, T. **Self-Organization And Associative Memory**, Springer-Verlag, 2nd Edition, Berlin, Germany, 1988.

KOHONEN, T. **The Self-Organizing Map**, Proceedings of IEEE, September 1990, pp. 1464-1480.

MEDEIROS, L. F. **Redes Neurais em Delphi**, Visual Books – 2nd Edição – Florianópolis – Brasília – 2006.

NASCIMENTO J. C. L. **Inteligência Artificial em controle e automação**, Edgard Blucher – 1º Edição – São Paulo – Brasil – 2004.