



Estudo Comparativo entre Redes Neurais Artificiais Aplicadas a Reconhecimento Facial

Renan Bessa¹, Elder dos Santos Teixeira², Guilherme Costa Holanda³, Auzuir Ripardo de Alexandria⁴

¹Graduando em Engenharia Mecatrônica – IFCE. Bolsista Funcap/IFCE. e-mail: renan@ifce.edu.br

²Mestrando em Engenharia Mecatrônica – IFCE. Professor do IFCE. e-mail: elderteixeira@ifce.edu.br

³Graduando em Engenharia Mecatrônica – IFCE. Bolsista CNPq. e-mail: guilherm3@gmail.com

⁴Doutorado em Engenharia de Teleinformática - UFC. Professor do IFCE. e-mail: auzuir@ifce.edu.br

Resumo: Um sistema de reconhecimento facial pode ser utilizado em diferentes aplicações, como controle de acesso e segurança, monitoramento de pessoas, reconhecimento de fichas criminais, robôs móveis para monitoramento de espaço, monitoramento do fluxo de pessoas em um determinado ambiente. Neste artigo, é realizado um estudo comparativo entre Redes Neurais Artificiais (RNAs) para a tarefa de reconhecimento de faces. As RNAs estudadas foram a *Multilayer Perceptron* (MLP), *Self-Organizing Map* (SOM) e *Neural-Gas*, que, juntamente com o método de Análise de Componentes Principais, obtiveram resultados acima de 80% de acerto médio de classificação dos indivíduos de um banco de faces amplamente utilizado na literatura.

Palavras-chave: Redes Neurais Artificiais, MLP, SOM, *Neural-Gas*, PCA, Reconhecimento Facial.

1. INTRODUÇÃO

A visão é um dos sentidos mais importantes para o ser humano, bastante útil para a sua localização, aprendizado, segurança. A visão é o principal recurso utilizado pelo homem para conhecer e reconhecer tudo que está ao seu redor. Nesse âmbito, é imprescindível a ideia de se criar sistemas automáticos para executar essas tarefas. Neste trabalho, foi escolhido o estudo de sistemas para o reconhecimento de faces humanas. Dentre os inúmeros exemplos de aplicação, podem-se citar como motivações desta área de pesquisa: controle de acesso e segurança, identificação pessoal para bancos, passaporte e fichas criminais, controlar o acesso de pessoas em um determinado ambiente e monitorar multidões em locais públicos. Uma das técnicas encontradas na literatura para o reconhecimento facial são as Redes Neurais Artificiais (RNAs). Existem inúmeros modelos de redes neurais e o presente trabalho tem o objetivo de realizar um estudo comparativo de três delas: a Rede Neural *Multilayer Perceptron* (MLP), Rede Neural *Self-Organizing Map* (SOM) e a *Neural-Gas*.

Uma Rede Neural Artificial é um modelo matemático inspirado na rede neural biológica em que os neurônios adquirem conhecimento através das experiências. Esse modelo consiste de elementos de processamento, chamados neurônios, e ligações entre eles dotadas de informação, que são chamadas de pesos sinápticos, constituindo a estrutura neural. A aprendizagem da rede neural ocorre na etapa de treinamento, pois os pesos das ligações podem ser modificados. No treino, a rede neural é estimulada pelo ambiente e sofre modificações resultantes desses estímulos, respondendo de maneira nova ao ambiente. Apesar da semelhança com o cérebro humano, essas redes não são feitas simplesmente para modelá-lo e sim para serem modelos úteis para a resolução de problemas de engenharia (KASABOV, 1998; HAYKIN, 2001).

O objetivo deste trabalho é a realização de um estudo comparativo entre os três modelos de redes neurais citados anteriormente. Os algoritmos propostos e os resultados foram implementados e obtidos utilizando o ambiente de desenvolvimento *MATLAB*®. Também foram usadas funções nativas do mesmo.

Nesse trabalho, são usados dois tipos de método de aprendizagem. A Rede Neural MLP possui o aprendizado supervisionado, em que é fornecido à rede um conjunto de dados de treino formado de dados de entradas com suas respectivas saídas desejadas. O objetivo da aprendizagem supervisionada é modificar os pesos tentando minimizar a diferença entre resposta desejada e a que é produzida pela rede. As redes SOM e *Neural-Gas* possuem o aprendizado não-supervisionado, em que o objetivo é descobrir padrões ou características nos dados de entrada sem ajuda externa (ENGELBRECHT, 2007).



2. FUNDAMENTAÇÃO TEÓRICA

2.1 ANÁLISE DE COMPONENTES PRINCIPAIS

A análise de componente principal (PCA) é um método multivariável que tem como objetivo a redução de dimensionalidade de um conjunto de dados, mantendo a maior parte da variância das amostras. Isto é possível por meio de uma transformação linear para um novo conjunto de coordenadas e que são ordenadas, levando em conta a variância dos dados. Logo, as últimas componentes têm pouca representatividade do conjunto de dados e podem ser desconsideradas (JOLLIFFE, 2002).

Monteiro et al. (2006) propôs que as imagens, representadas por uma matriz 243×320 pixels cada, fossem reorganizadas de forma que cada coluna fosse posta uma abaixo da outra, formando um vetor coluna de 77760×1 pixels. Esse processo é feito com todas as 165 imagens do banco de dados Yale e, ao final, concatena-se todos os 165 vetores, produzindo matriz de 77760×165 pixels. Após a aplicação do PCA nessa nova matriz, constata-se experimentalmente que as 26 primeiras componentes principais explicam 90% da variância total do conjunto de dados, tornando a matriz 26×165 .

2.2 REDE NEURAL MLP

A rede MLP é uma das mais importantes topologias de redes neurais. Ela é constituída de unidades sensoriais, formando uma camada de entrada, representando o local de apresentação dos dados de entrada, uma ou mais camadas ocultas de neurônios e uma camada de neurônios de saída, onde o número de neurônios desta camada é escolhido pela codificação das classes proposta pelo programador (HAYKIN, 2001).

Os perceptrons de multicamadas são treinados por meio de um algoritmo de retropropagação do erro (*back-propagation*). Esse algoritmo é fundamentado pela regra de aprendizagem por correção de erro. A aprendizagem dessa rede possui dois passos fundamentais: primeiramente, a propagação das entradas por toda a rede até que a camada de saída apresente um valor, e no segundo momento, é calculado o erro em relação ao valor desejado. Esse erro é retropropagado para as camadas no sentido da saída para a entrada, ajustando os pesos sinápticos dos neurônios de acordo com a relação de aprendizagem por correção de erro (HAYKIN, 2001).

Um neurônio artificial i é ilustrado pela Figura 1. Os dados de entrada desse neurônio são representados pelo vetor

$$\mathbf{x} = x_1, x_2, \dots, x_p \quad (\text{Eq. 02})$$

e cada entrada é associada a um peso representado pelo vetor

$$\mathbf{w}_i = w_{i1}, w_{i2}, \dots, w_{ip} \quad (\text{Eq. 03})$$

O valor da saída do neurônio é influenciado por um limiar θ , também chamado de *bias*. Normalmente, é adicionado no vetor de entrada \mathbf{x} o valor -1 (ou 1), tornando-se o x_0 de todas as entradas do neurônio. O *bias* é adicionado no vetor de pesos do neurônio \mathbf{w}_i , tornando-se o w_{i0} (ENGELBRECHT, 2007).

O valor do sinal de entrada líquido u_i é calculado pelo somatório dos dados de entrada ponderados pelos pesos sinápticos:

$$u_i = \sum_{p=0}^n (x_p * w_{ip}). \quad (\text{Eq. 04})$$

O valor de u_i é restringido por uma função de ativação. A saída de um neurônio é limitada ao intervalo de $[0,1]$ ou $[-1,1]$, que é o caso da tangente hiperbólica (ENGELBRECHT, 2007). Esta é a função de ativação escolhida para este trabalho e sua função é dada por

$$\text{tanh } u_i = \frac{1 - \exp(-u_i)}{1 + \exp(-u_i)}. \quad (\text{Eq. 05})$$

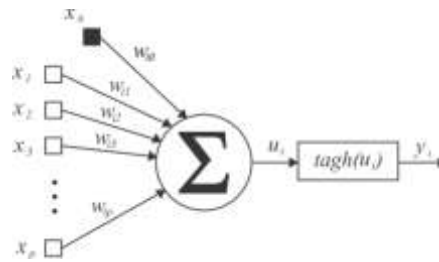


Figura 1 – Um modelo de neurônio artificial.

Em um problema de classificação de M classes, em que M é o número total de classes, são necessárias M saídas para representar todas as decisões de classificações possíveis (HAYKIN, 2001).

A MLP implementada nesse trabalho possui 15 saídas, pois existem 15 classes de indivíduos. A codificação foi feita por valores 1 e -1 como, por exemplo, para representar o indivíduo da classe 1, tem-se

$$\mathbf{C}_1 = (1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1). \quad (\text{Eq. 06})$$

Quando o neurônio está localizado na camada de saída, como, por exemplo, o neurônio à direita na Figura 2, ele possui um valor desejado d particular dado pela representação das classes. Por esse motivo, podem-se atualizar os pesos \mathbf{m}_k do neurônio k diretamente pela regra de correlação do erro, onde o erro instantâneo do neurônio i na interação n é calculado por

$$e_k n = d_k n - z_k(n). \quad (\text{Eq. 07})$$

A regra de aprendizagem para o neurônio da camada de saída é expressa por

$$\mathbf{m}_k n + 1 = \mathbf{m}_k n + \alpha * e_k n * \frac{1}{2} * (1 - z_k^2 n) * \mathbf{y}_i(n), \quad (\text{Eq. 08})$$

em que α é a taxa de aprendizagem ou passo de adaptação, assumindo valor entre 0 e 1, e $\frac{1}{2} * (1 - z_k^2 n)$ é a derivada primeira da função tangente hiperbólica.

No caso do neurônio localizado na camada oculta, não existe valor desejado particular. O algoritmo *back-propagation* propõe que o sinal de erro para um neurônio oculto seja constituído dos sinais de erro de todos os neurônios da camada de saída que estejam conectados diretamente a ele (HAYKIN, 2001).

A regra de aprendizagem para o neurônio da camada de entrada é expressa por

$$\mathbf{w}_i n + 1 = \mathbf{w}_i n + \alpha * \frac{1}{2} * (1 - y_i^2 n) * \mathbf{x} n * \sum_{k=1}^M m_{ki} * e_k n * \frac{1}{2} * (1 - z_k^2 n), \quad (\text{Eq. 09})$$

em que M é o número de neurônios de saída.

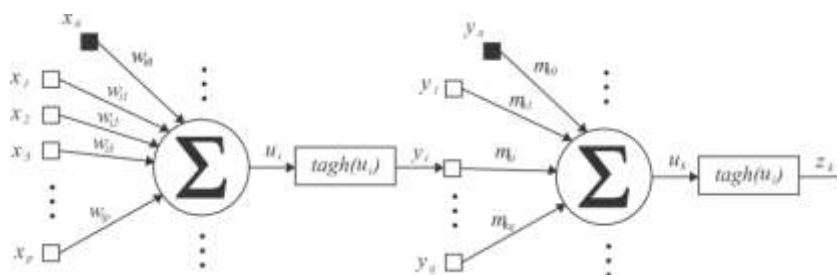


Figura 2 – Rede MLP.



2.3 REDE NEURAL SOM

A rede SOM, proposta por Teuvo Kohonen, possui um algoritmo não supervisionado para o seu treinamento. O objetivo deste mapa auto-organizável é criar um arranjo para os padrões de entrada em um mapa discreto uni ou bidimensional de uma maneira a se obter uma topologia ordenada. Basicamente, o algoritmo consiste em apresentar um vetor de entrada \mathbf{x} a todos os neurônios da rede, encontrar um neurônio vencedor e modificar os pesos do neurônio vencedor e dos neurônios que formam sua vizinhança naquele momento (HAYKIN, 2001; KASABOV, 1998).

O neurônio vencedor é o que melhor representa o vetor de entrada. Usando a distância euclidiana podemos encontrar o neurônio com pesos \mathbf{w}_i que melhor case com o vetor de entrada \mathbf{x} . O índice do neurônio vencedor é dado por

$$i = \arg \min_j \|\mathbf{x} - \mathbf{w}_j\|, \quad j = 1, 2, \dots, l, \quad (\text{Eq. 10})$$

em que l é o número de neurônios da rede.

Os neurônios podem ser dispostos em uma matriz de vizinhança ilustrada na Figura 3. Quando a rede SOM é bidimensional, o índice i de \mathbf{x} é um par ordenado (linha, coluna) que localiza o neurônio na matriz de vizinhança. Podemos assumir que a vizinhança topológica $h_{j,i}$ é uma função unimodal da distância $d_{j,i}$, em que $d_{j,i}$ é a distância euclidiana entre o neurônio vencedor e um certo vizinho, sendo expressa por

$$h_{j,i} = \exp \left(-\frac{d_{j,i}^2}{2\sigma^2} \right), \quad (\text{Eq. 11})$$

em que σ é a largura efetiva da vizinhança e é responsável pela diminuição da vizinhança topológica com o tempo. Normalmente, σ possui um decaimento exponencial ao longo das iterações (HAYKIN, 2001).

A regra de atualização dos pesos dos neurônios da rede é expressa por

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \alpha(n) * h_{j,i}(n) * \mathbf{x}(n) - \mathbf{w}_j(n), \quad (\text{Eq. 12})$$

em que a taxa de aprendizado α deve ser variável com as interações, possuindo um decaimento.

Após o treinamento, os neurônios devem ser rotulados. O conjunto de treino é novamente apresentado à rede e cada neurônio vencedor guarda o rótulo da classe pertencente à entrada apresentada. Ao fim desta fase, o neurônio é rotulado pela classe que mais vezes ele foi o vencedor. Em caso de empate de número de classe, a escolha do rótulo é feita aleatoriamente. Caso o neurônio não tenha sido apontado como vencedor em nenhum momento, ele é descartado (MONTEIRO, 2006).

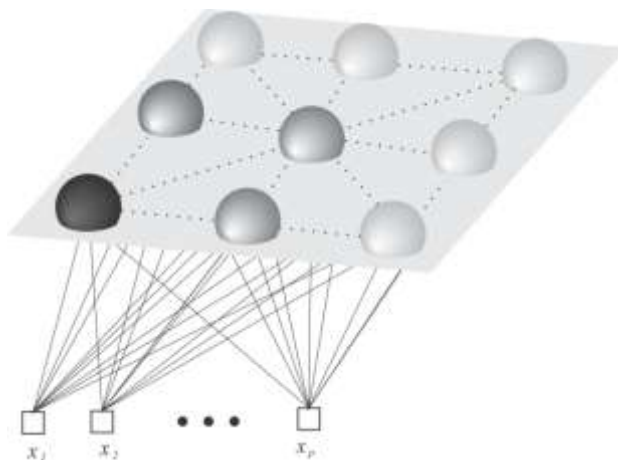


Figura 3 – Rede SOM.



2.4 REDE NEURAL-GAS

A rede *Neural-Gas* tem os pesos sinápticos dos seus neurônios adaptados independente de qualquer arranjo topológico das unidades neurais dentro da rede neural, diferenciando da rede SOM que possui uma estrutura topológica previamente determinada. As mudanças nos pesos sinápticos, na rede *Neural-Gas* são determinadas pelas distâncias relativas entre as unidades neurais dentro do espaço de entrada (MARTINETZ et al., 1993).

A cada vetor de entrada x é construído um ranking de vizinhança em que o vetor com a menor distância euclidiana é o primeiro, o segundo de menor distância é o segundo e assim para todos os neurônios da rede. Após a formação instantânea da vizinhança relacionada à entrada x , os neurônios têm seus pesos atualizados da mesma forma que uma rede SOM unidimensional. Ao término do treinamento, a rotulação dos neurônios utilizada foi a mesma da rede SOM proposta por Monteiro et al (2006).

3. Materiais e Métodos

Para o estudo comparativo das redes neurais no reconhecimento facial é utilizado o banco de faces da Universidade Yale, composto por 165 imagens de 15 indivíduos em várias feições (triste, alegre, sério), iluminados em diferentes direções e portando ou não acessórios (óculos) (BELHUMEUR et al., 1997).

Os passos deste trabalho são expostos na Figura 4.

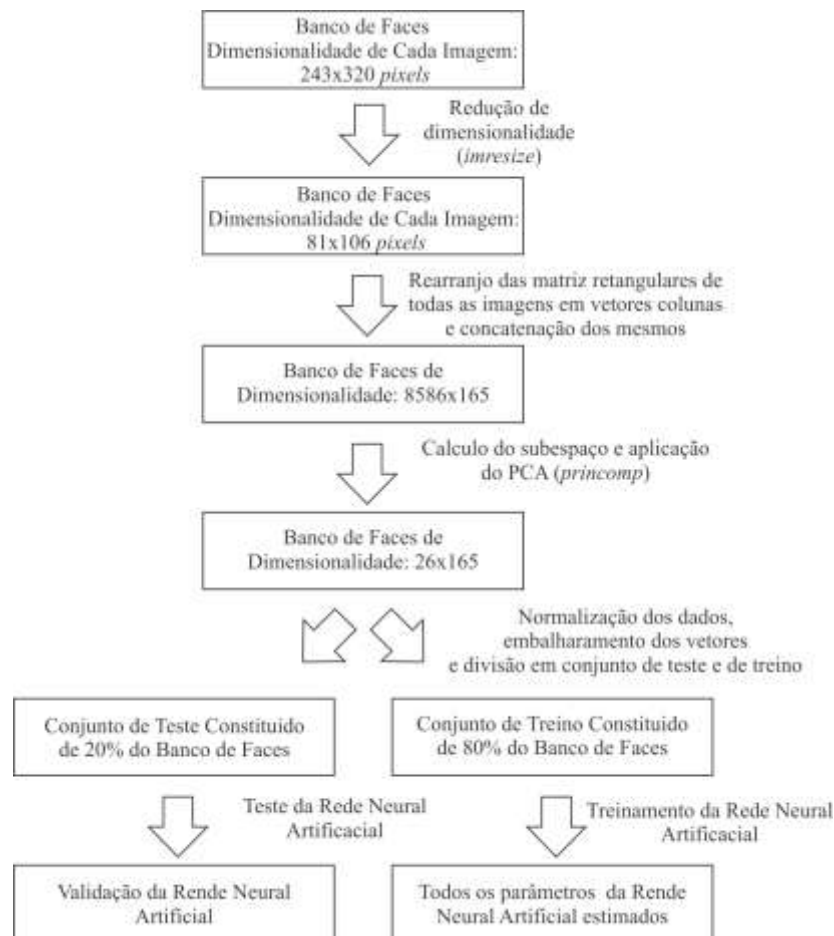


Figura 4 – Metodologia do Trabalho.



No trabalho, utiliza-se a função *princomp* do ambiente *MATLAB*®, que retorna o subespaço das componentes principais, porém, o *MATLAB*® apresentou um “estouro de memória” de uma variável quando apresentada a matriz 77760x165 pixels. Este problema é solucionado utilizando a função *imresize*, que transforma as imagens originais em imagens de 81x106 pixels. A função *princomp* também retorna a porcentagem da variância dos dados de cada componente principal. Como no trabalho de Monteiro et al. (2006), as 26 primeiras componentes principais portam 90% da variância do conjunto de dados.

Após a aplicação do método PCA no conjunto de dados, cada imagem é representada por um vetor de dimensão 1x26. As 165 imagens são divididas aleatoriamente em grupos para treinamento e para teste, sendo 80% dos vetores para treinamento e 20% para teste. O grupo de treinamento é responsável pela parametrização da rede neural e o grupo de teste valida a eficiência do classificador.

3. RESULTADOS E DISCUSSÃO

Após vários testes de parâmetros, de números de neurônios e de épocas de treinamento, foram obtidos resultados que são apresentados nesse tópico.

A evolução da porcentagem média de acertos da Rede Neural MLP pelo número de neurônios em sua única camada oculta é ilustrada na Figura 4. Para o cálculo da média de acerto, foram realizados 30 vezes o experimento, com os valores de taxa de aprendizagem e número de épocas que produzem o melhor acerto. É observado que a partir de 10 neurônios na camada oculta não ocorre um acréscimo efetivo na porcentagem de acerto médio pelo classificador.

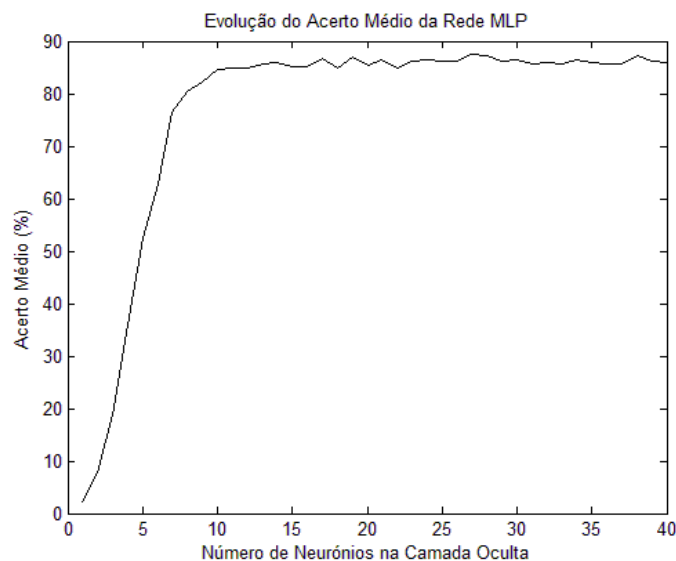


Figura 4 – Evolução do acerto da rede MLP.

A evolução da porcentagem média de acertos por número de neurônios da Rede Neural SOM e *Neural-Gas* são ilustrados nas Figuras 5 e 6, respectivamente. Opta-se em utilizar matrizes de vizinhança quadradas para a rede SOM e o mesmo número de neurônios para a rede *Neural-Gas*. Também para o cálculo da média de acerto, são realizados 30 vezes o experimento com os valores de taxa de aprendizagem e número de épocas que produzem o melhor acerto. As porcentagens de acerto médio da rede SOM e *Neural-Gas* apresentam-se estáveis a partir de 64 neurônios.

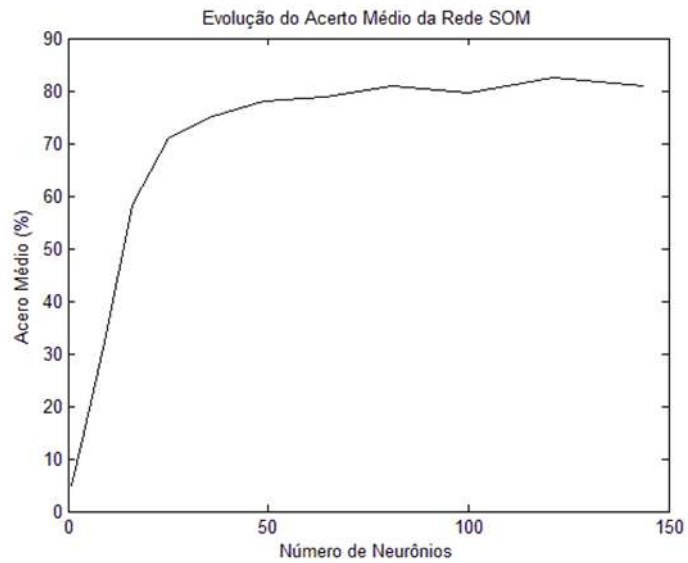


Figura 5 – Evolução do acerto da rede SOM.

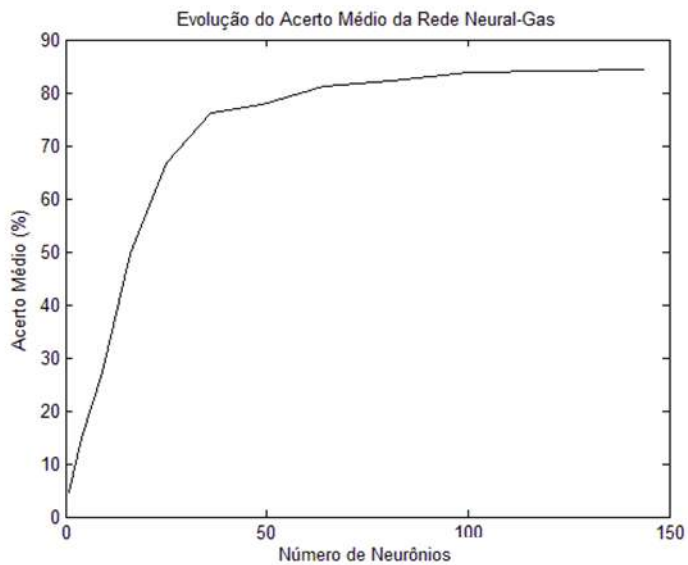


Figura 6 – Evolução do acerto da rede *Neural-Gas*.

Os melhores resultados obtidos por cada classificador são apresentados na Figura 7.

Resultados (%)			
Rede Neural	Média	Máximo	Mínimo
MLP	87,47	100	75,75
SOM	82,52	96,96	69,69
<i>Neural-Gas</i>	84,24	96,96	72,72

Figura 7 – Tabela de Resultados.



4. CONCLUSÕES

São obtidos resultados de acertos médios das RNAs na classificação dos indivíduos do banco de dados YALE acima de 80%.

O método de Análise de Componentes Principais se mostrou bastante eficiente na extração de atributos das imagens.

A Rede Neural MLP apresenta melhores resultados do que as demais redes estudadas nesta pesquisa.

Pode-se sugerir como trabalhos futuros: estudos comparativos de outros modelos de redes neurais supervisionadas, não-supervisionadas e as chamadas redes neurais crescentes. A implementação de um algoritmo de PCA que suporte imagem com maior resolução.

REFERÊNCIAS

BELHUMEUR, P.; HESPANHA, J.; KRIEGMAN, D. **Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection**, IEEE Transactions on Pattern Analysis and Machine Intelligence, Julho 1997, pp. 711-720

ENGELBRECHT, A.P. **Computational Intelligence: An Introduction**. 2. ed. Bookman, 2001.

HAYKIN, S. **Redes Neurais: Princípios e Prática**. 2. ed. Wiley, 2007.

JOLLIFFE, I.T. **Principal Component Analysis**. 2. ed. Springer, 2002.

KASABOV, N.K. **Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering**. A Bradford Book The MIT Press Cambridge, 1998.

MARTINETZ, T.M.; BERKOVICH, S.G.; SHULTEN, K.J. **“Neural-Gas” Network for Vector Quantization and its Application to Time-Series Prediction**. IEEE Transactions on Neural Networks, Vol. 4, No. 4, Julho 1993.

MONTEIRO, I.Q.; QUEIROZ, S.D.; CARNEIRO, A.T.; Souza, L.G.; BARRETO, G.A. **Face recognition independent of facial expression through SOM-based classifiers**. IEEE International Conference on Computer Vision Systems, 2006.