



Faculdades Adamantinenses Integradas (FAI)

www.fai.com.br

LIMA, Fernando Parra dos Anjos; SILVA, James Clauton; ESTEVAM, Giuliano Pierre; MINUSSI, Carlos Roberto. Reconhecimento de dígitos com uso de redes neurais artificiais. *Omnia Exatas*, v.3, n.2, p.29-39, 2010.

RECONHECIMENTO DE DÍGITOS COM USO DE REDES NEURAIIS ARTIFICIAIS

DIGITS RECOGNITION USING ARTIFICIAL NEURAL NETWORKS

Fernando Parra dos Anjos Lima

James Clauton da Silva

Graduandos de Engenharia da Computação – Unisaesiano - Araçatuba – SP

Giuliano Pierre Estevam

Professor Titular de Engenharia – Unisaesiano - Araçatuba – SP

Carlos Roberto Minussi

Prof. Titular do Dep. Engenharia Elétrica - UNESP – Ilha Solteira – SP

RESUMO

Este plano de pesquisa visa um estudo dos principais conceitos e arquiteturas das Redes Neurais Artificiais bem como o desenvolvimento prático destes conceitos. Para o desenvolvimento prático presente neste estudo utiliza-se um problema proposto em referencia bibliográfica, assim implementa-se uma aplicação na plataforma de desenvolvimento Borland Delphi 7 para reconhecimento e classificação de dígitos, proposta pelo autor Luciano Frontino de Medeiros, no livro “Redes Neurais em Delphi”. Para o desenvolvimento da aplicação utiliza-se uma rede neural do tipo MLP (Mult Layer Perceptron), através de um componente chamado TMLP, que se consiste na implementação do conceito do algoritmo MLP para uso em qualquer aplicação na plataforma Borland Delphi 7. Os dados que serão apresentados para a rede neural no processo de treinamento, e no processo de teste serão armazenados em uma tabela de banco de dados do tipo paradox, para facilitar a manipulação das imagens. Assim temos uma etapa antes de inicializar os procedimentos com a rede neural para cadastrar os padrões de imagens de entrada, e também para os padrões de imagens de teste. Então o aplicativo em questão será composto por duas etapas. A primeira etapa será a etapa de apresentar os dados de entrada (imagens de dígitos) para treinamento da rede neural, a fim de adquirir experiência e armazenar os padrões dos números para uma futura identificação. A segunda etapa consiste em executar um teste para verificação do aprendizado da rede neural, através da apresentação de imagens de dígitos diferentes dos padrões apresentados no processo de treinamento, e obter a identificação de cada número respectivamente. Devo lembrar que no processo de treinamento e no processo de teste será realizada uma varredura em todas as imagens cadastradas na tabela paradox, tanto para adquirir conhecimento sobre os padrões, como também para identificação dos padrões. Assim através da varredura será possível identificar e classificar cada dígito e fazer o processo de reconhecimento.

Palavras-chaves: Redes Neurais Artificiais, Reconhecimento de dígitos, identificação, classificação.

ABSTRACT

This research plan aims to study the main concepts and architectures of neural networks as well as the practical development of these concepts. For the practical development in this study we use a proposed problem in bibliographical references, so we implement an application development platform for Borland Delphi 7 digit recognition and classification proposed by the author Frontino Luciano de Medeiros, in his book "Neural Networks in Delphi. " To develop the application itself uses a neural network MLP (Multi Layer Perceptron) through a component called WLMT, which consists in implementing the concept of the MLP algorithm for use in any application on the platform Borland Delphi 7. Data to be presented to the neural network in the training process, and in the testing process will be stored in a database table of type paradox, to facilitate manipulation of images. So we have a step before starting the procedure with the neural network to record the patterns of input images, and also to the standards of test images. Then the application in question will consist of two stages. The first step is the step of presenting the input data (images of digits) for

training the neural network in order to gain experience and store the patterns of numbers for future identification. The second step is to run a test to verify the neural network learning, by presenting images of digits of different patterns in the training process, and obtain the identification number of each respectively. I remember that in the process of training and testing process will be a sweep in all the images registered in paradox table, both to acquire knowledge of the patterns, but also for identification of standards. Just scan through it will be possible to identify and classify each digit and make the recognition process.

Key-words: Artificial Neural Networks, digit recognition, identification, classification.

INTRODUÇÃO

Redes Neurais Artificiais são conceitos do mecanismo de funcionamento do cérebro aplicados em software e/ou hardware, por tanto são capazes de adquirir conhecimento, ou seja, apreenderem com a experiência. As redes são capazes de efetuar diagnósticos complexos, sendo assim as redes se modelam através de várias unidades de neurônios ou unidades de processamento, dispostos em fileiras ou camadas, formando um arranjo de interligações, tais interligações são conexões constituídas por pesos, ou sinapses, que são o ponto de calibração das unidades de processamento. 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 podemos dizer, antes que o sistema entre em funcionamento. Uma vez ajustada, empregando-se algum tipo de algoritmo de adaptação dos referidos pesos, a rede se denomina apta para efetuar diagnósticos e emitir um grau de precisão satisfatório no reconhecimento de padrões não utilizados no seu processo de aprendizado.

Este processo pode ser visto como uma possibilidade na viabilização de rotinas de análises em tempo real, que certamente é importante no diagnóstico de várias áreas do conhecimento humano, tais como: medicina, biologia, engenharia, ecônomia entre outras.

Levando em consideração a importância das análises de rotinas em tempo real, este plano de pesquisa visa realizar um estudo embasado nos conceitos e arquiteturas neurais artificiais e suas aplicações de modo geral, e em específico o reconhecimento de padrões. Utilizaremos a plataforma e linguagem de programação Borland Delphi 7 para implementação de um sistema para reconhecimento de dígitos com redes neurais artificiais. Na plataforma de programação Borland Delphi será usado um componente que simula a arquitetura e todos os conceitos da rede MLP (Mult Layer Perceptron). Assim podemos elaborar um sistema e utilizar a rede neural para realizar o processamento das informações.

No sistema a ser desenvolvido utilizaremos uma tabela de banco de dados paradox, que servirá para armazenar imagens de padrões de entrada, e imagens de padrões de testes, todas previamente cadastradas. Após esta etapa damos início ao processamento da rede neural com a fase de treinamento, nesta fase será executado uma varredura em todas as imagens de entrada cadastradas, a fim de adquirir conhecimento e guardar os padrões para realizar identificações no futuro. Realizado o treinamento da rede teremos um arquivo de conhecimento que será utilizado na fase de teste, este arquivo armazena informações sobre cada padrão treinado, e é de suma importância no processo de teste, pois com ele que serão identificados os padrões na fase de teste. Na fase de teste temos um procedimento igual ao processo de treinamento, porém somente realizamos a varredura para identificar e classificar os padrões que estão sendo apresentados, com base no arquivo de conhecimento obtido no treinamento.

A seguir vamos descrever este processo com mais detalhes.

MATERIAIS E MÉTODOS.

Para implementar o sistema utiliza-se a plataforma de desenvolvimento Borland Delphi 7, e um componente pronto que nos ajudará nas tarefas a serem desenvolvidas. Utilizaremos o componente a seguir:

- TMLP: O objeto TMLP é uma implementação do conceito de rede Mult Layer Perceptron em Object Pascal para que possamos utilizar, através de um componente instalado no Delphi 7, a rede neural em qualquer aplicativo.

Este componente nos permite executar qualquer função que contenha um conceito da rede MLP (Mult Layer Perceptron). Assim vamos à modelagem do aplicativo.

Definições dos parâmetros do aplicativo.

Para implementação utiliza-se uma rede neural do tipo MLP (Mult Layer Perceptron) e uma tabela de banco de dados paradox para armazenar as informações de entrada e teste.

Para modelarmos o sistema foram montadas duas matrizes de dígitos, cada uma contendo 100 dígitos de 0 a 9, totalizando assim, 200 dígitos. A finalidade aqui é utilizarmos 100 dígitos para a fase de treinamento e 100 dígitos para a fase de testes. As matrizes foram escritas manualmente e digitalizadas para utilizar-se no sistema.

As figuras 1 e 2 ilustram as matrizes montadas para o sistema:

1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0

Figura 1: Matriz de dígitos de treinamento.

1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0
1	2	3	4	5	6	7	8	9	0

Figura 2: Matriz de dígitos de teste.

Para o uso na rede e armazenamento na tabela paradox, estes caracteres foram extraídos da matriz na forma de bitmaps de tamanho 32 x 32 que serão associados aos dígitos respectivos no arquivo de treinamento e de teste.

Na entrada da rede temos uma imagem de amostra do dígito com 32 x 32 pixels, assim temos processos de convolução e subamostragem para geração dos Ocelos (pequenos olhos) que serão as entradas da rede. A figura 3 ilustra como será realizado o processo para obter os ocelos:

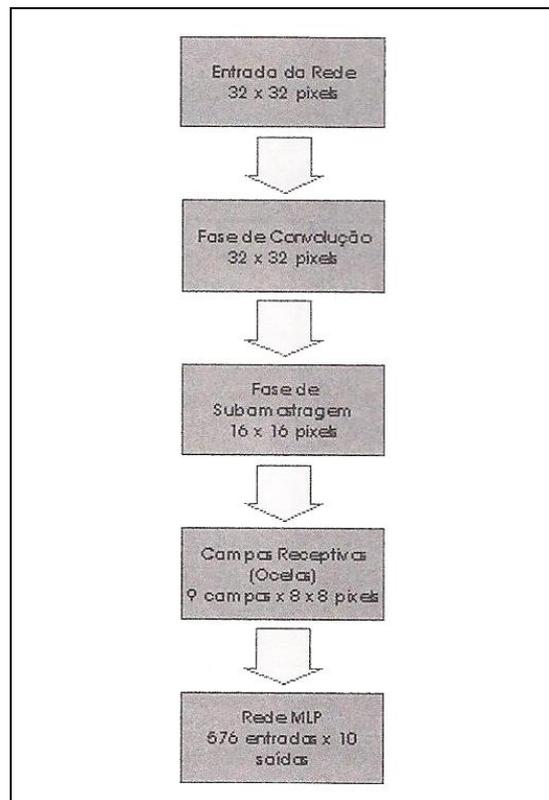


Figura 3: Manutenção do conjunto de treinamento.

Para obter os ocelos temos que passar por varias etapas, que são as seguintes: convolução (Suavização e eliminação de ruídos), subamostragem (redução da escala da imagem) e finalmente obtenção dos Ocelos (pequenos olhos).

Veja a figura 4, onde temos um exemplo de uma mascara de 3 x 3 pixels, utilizada pra convolução. Os valores dos pixels são ponderados (na figura os valores dos pesos são os valores de "a" a "i") e somados de forma a se obter um valor da cor para o pixel resultante.

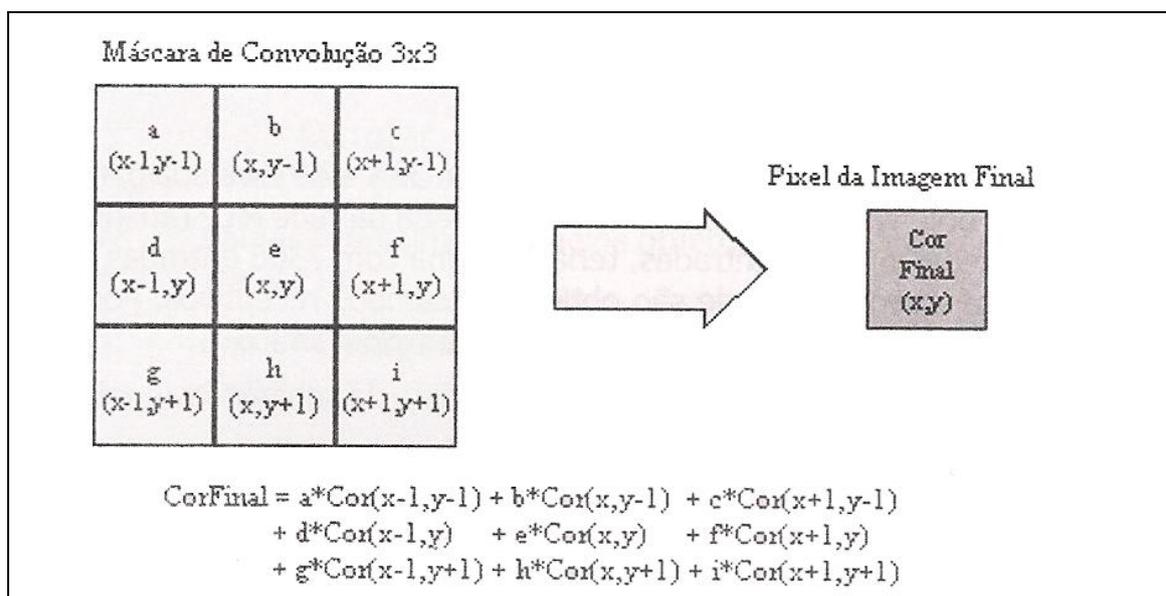


Figura 4: representação esquemática da convolução em pixels.

Em nosso sistema utilizamos uma convolução onde os valores de ponderação são unitários, o que se denomina filtro de média. Esse processo intensifica a forma do dígito e facilita a geração dos ocelos.

Na fase de subamostragem, a imagem convolvida 32×32 é escalada para 16×16 . Nesta fase o objetivo é reduzir o tamanho da entrada da rede MLP.

Os campos receptivos ou Ocelos são obtidos através da fase anterior como se fosse deslocada uma janela do tamanho 8×8 através da imagem, de forma a termos nove imagens (Figura 5) com certo grau de redundância proposital. Esta redundância aumenta as opções para classificação do dígito na rede MLP.

A Figura 5 ilustra estes processos realizados:

Como temos todo processamento da imagem realizado, vamos definir as características da rede neural artificial. Como temos 9 Ocelos de 8×8 pixels, temos 9 imagens de 64 pixels, sendo assim temos um total de 576 pixels de entrada para a rede. E temos 10 saídas, respectivamente uma para cada dígito.

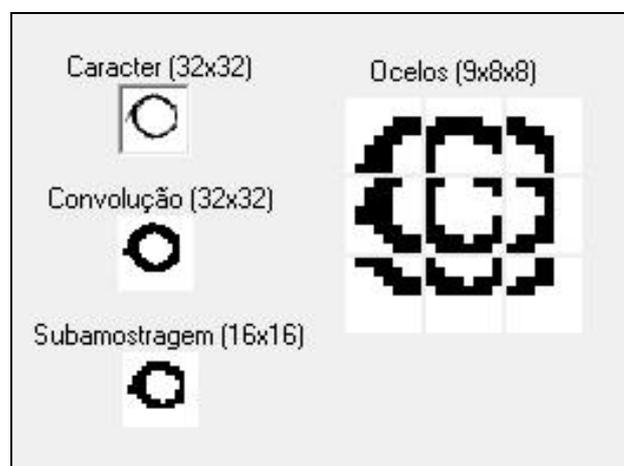


Figura 5: 9 Ocelos para o dígito zero.

Modelagem do aplicativo

Após definirmos todos os parâmetros dos conjuntos de treinamento e teste, e os parâmetros da rede MLP, vamos ao funcionamento do aplicativo.

Na entrada da rede temos 576 entradas, que foram obtidas no processamento da imagem antes do treinamento. Assim temos previamente cadastradas as 200 imagens, 100 de treinamento e 100 de teste. Agora podemos executar a rede MLP. O processo de treinamento se dá através de uma varredura em todas as 100 imagens de treinamento cadastradas na tabela paradox, para adquirir conhecimento sobre os padrões existentes em cada dígito apresentado, e assim é gerado e definido um padrão sobre cada tipo de dígito, após definido esse padrão ele é armazenado em um arquivo de conhecimento, o qual será utilizado posteriormente para identificar e classificar os dígitos no processo de teste.

O processo de teste se baseia no processo de treinamento, será realizada uma varredura nas 100 imagens dos dígitos de teste, para identificação e classificação dos dígitos, através do conhecimento obtido no processo de treinamento, que está armazenado no arquivo de conhecimento. Assim a cada padrão que a rede encontra que combina com seu padrão armazenado é obtido um par e é realizado o processo de reconhecimento.

Assim ao final da varredura temos a quantidade de acertos e erros realizados pela rede.

Telas do sistema

Inicialmente temos a tela principal, onde temos os menus para acessar as opções de configuração de parâmetros da rede MPL, e também as opções de processamento, como conjunto de entrada e teste, e treinamento e teste. A Figura 6 ilustra a tela principal do aplicativo os menus presentes nela.

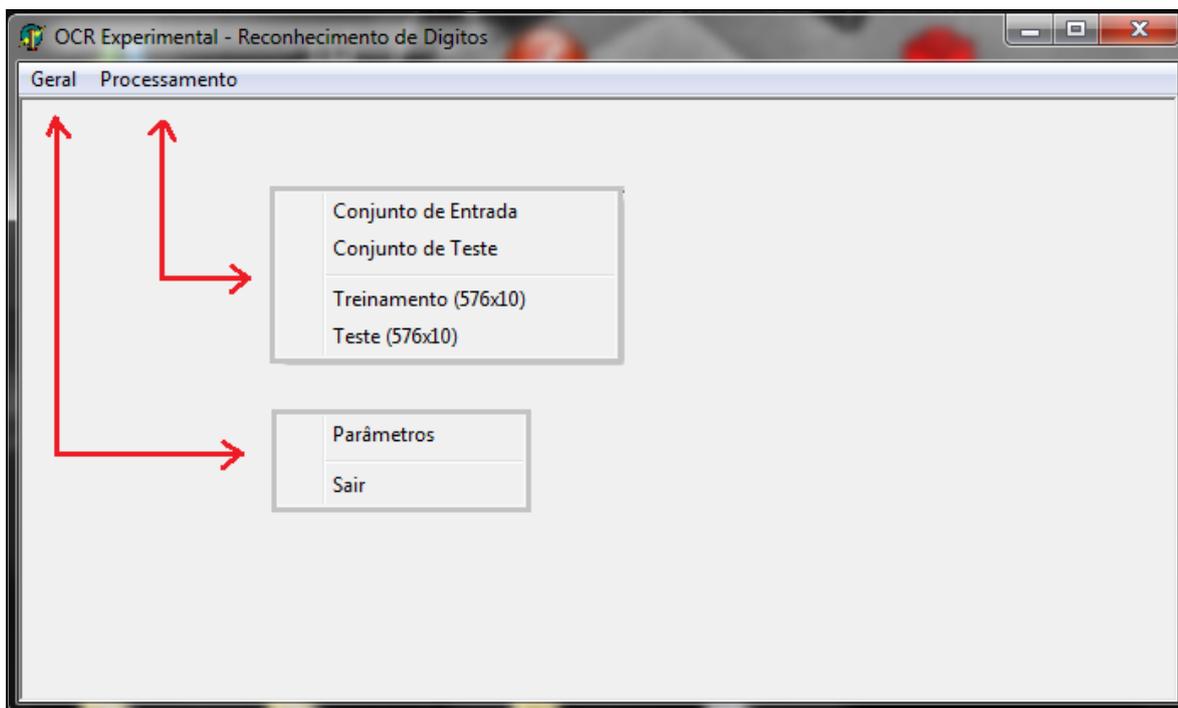


Figura 6: Tela principal e os menus.

Inicialmente na utilização do aplicativo temos que definir os parâmetros da rede neural artificial, tais como o número de épocas que será realizado no treinamento da rede, o número de neurônios na camada oculta, a taxa de aprendizagem na rede, a taxa de inércia, e também o nome do arquivo de conhecimento. A Figura 7 mostra a tela de configuração dos parâmetros.

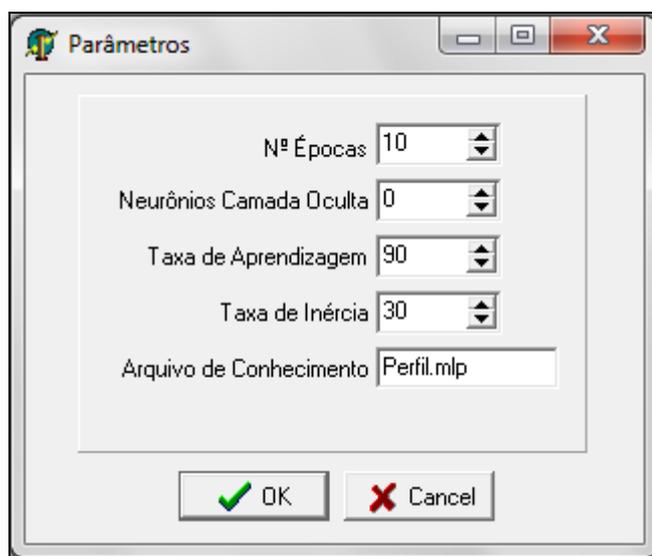


Figura 7: tela de parâmetros da rede MLP.

Após definirmos os parâmetros de configuração da rede temos que cadastrar as imagens no conjunto de treinamento e no conjunto de teste, estas imagens foram extraídas da matriz com o tamanho de 32 x 32 pixels, e estão prontas para serem cadastradas. As Figuras 8 e 9 ilustram as telas de cadastros dos conjuntos de treinamento e teste.

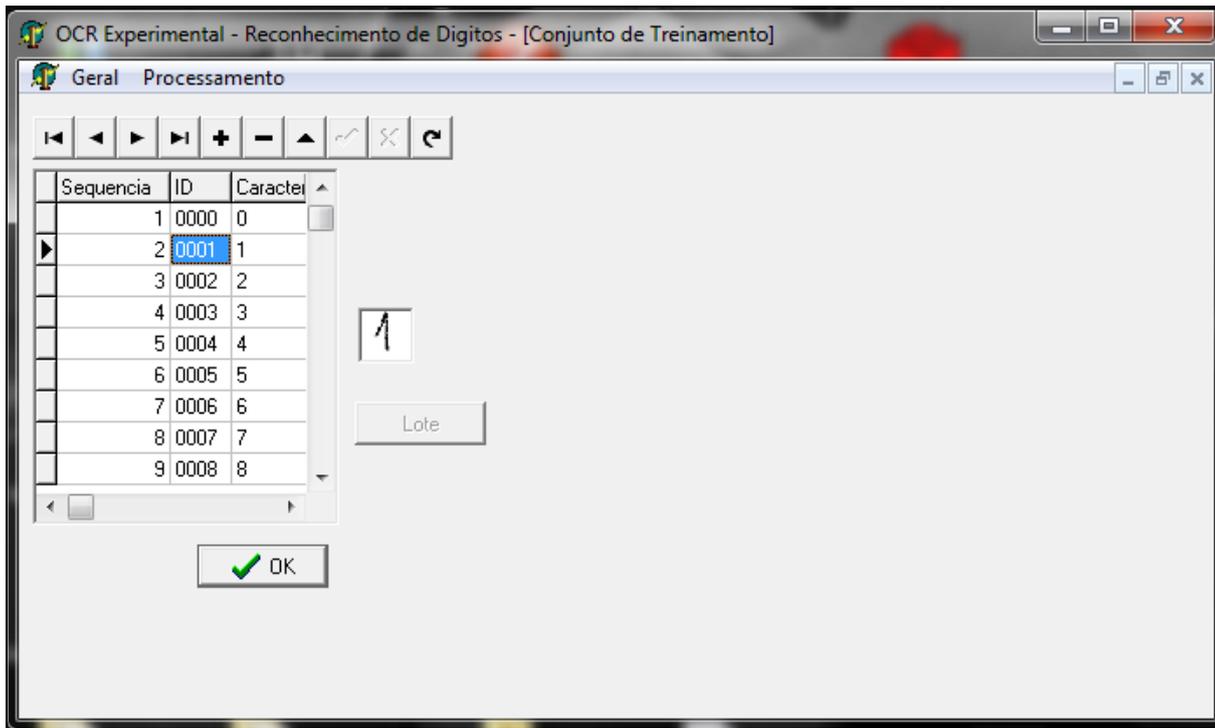


Figura 8: Tela de cadastro do conjunto de treinamento.

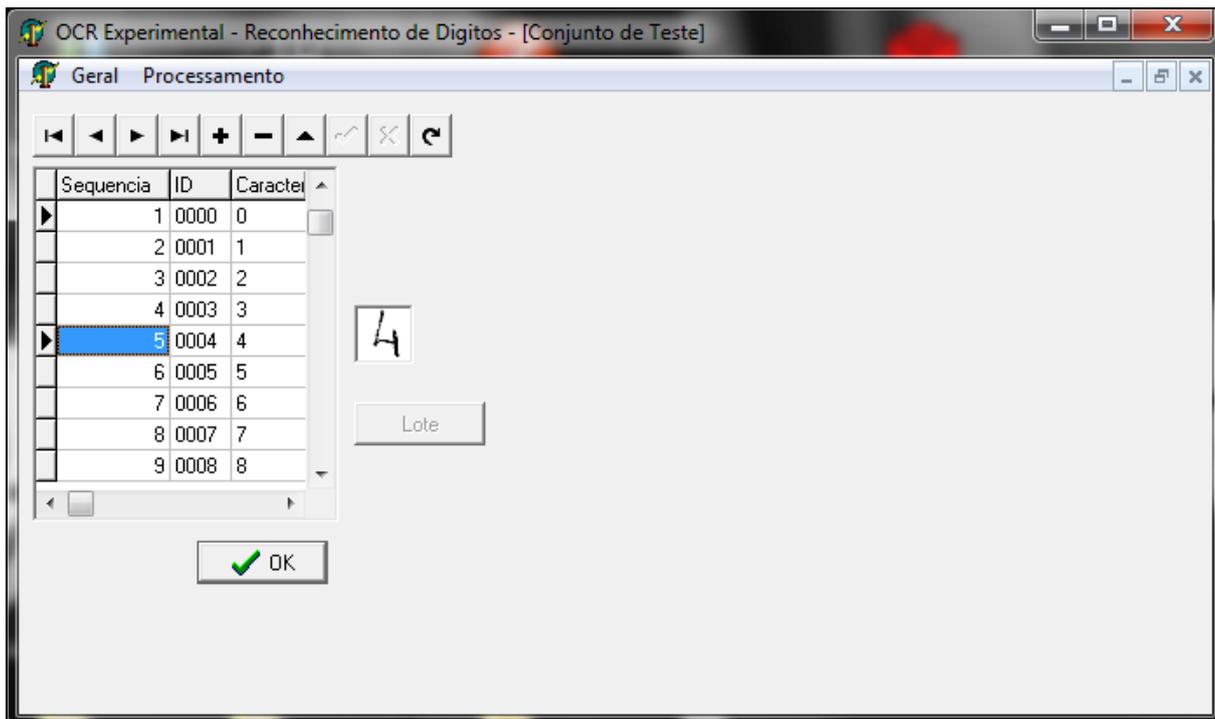


Figura 9: Tela de cadastro do conjunto de teste.

Depois de definidos os conjuntos de treinamento e teste, podemos executar a rede MLP e realizar o treinamento da mesma. A Figura 10 apresenta a tela de treinamento:

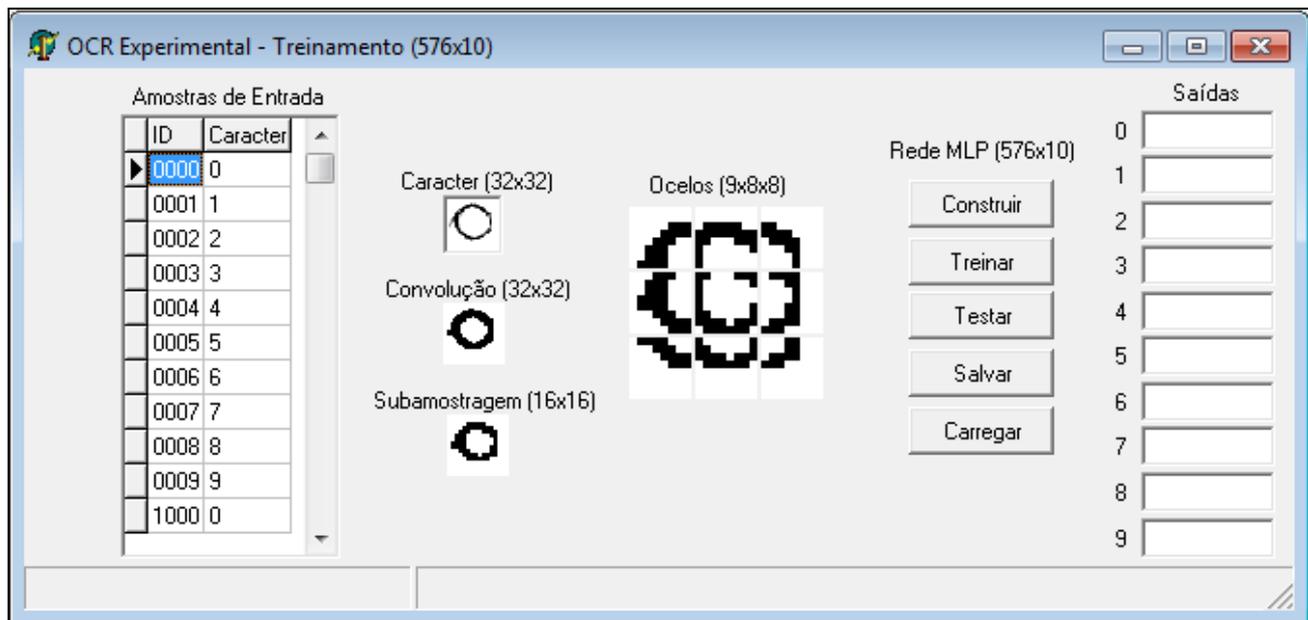


Figura 10: Tela de treinamento.

Como dito anteriormente o processo de treinamento irá realizar uma varredura em todas as amostras de entrada e armazenar o conhecimento obtido de cada amostra em um arquivo de conhecimento, para se utilizar posteriormente no processo de teste. Assim durante o processo de teste é carregado o conhecimento para se utilizar naquele momento. Na Figura 11 tem-se a imagem da tela de teste, que é praticamente idêntica a tela de treinamento:

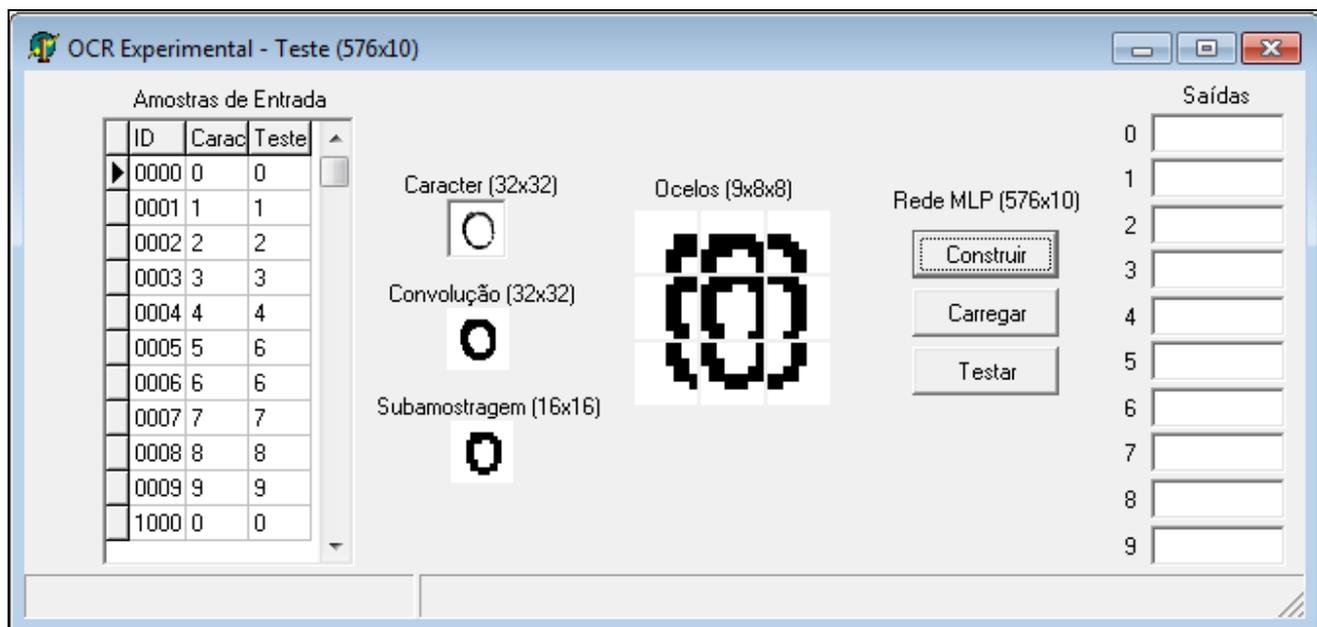


Figura 11: Tela de teste.

Esta tela que permite testar o conhecimento e desempenho da rede MLP.

RESULTADOS E DISCUSSÕES

Para este sistema propusemos uma análise com as 200 imagens, 100 no conjunto de treinamento e 100 no conjunto de teste, assim optamos por fazer dois testes diferentes, um em que temos um número de épocas baixo e outro em que temos um número de épocas grande.

Para o primeiro utilizamos um número de épocas igual a 10, sendo assim foi rápido o processo de treinamento, pois executou a varredura em todas as imagens apenas 10 vezes. Para o segundo utilizamos um número de épocas igual a 1000, sendo assim este processo de treinamento foi mais demorado, porque é necessário realizar a varredura em todas as imagens 1000 vezes.

Os resultados que obtivemos com estes dois testes confirmaram o que as bibliografias e teorias dizem a respeito do assunto. Portanto temos que quando mais se realiza um procedimento, dificilmente se erra no mesmo, e assim pelo contrario, quando menos se treina para realizar um procedimento, maior a chance de se errar. Nos testes da rede MLP obtivemos este resultado.

Para o primeiro treinamento, com 10 épocas, obtivemos 5 erros no reconhecimento e identificação das 100 imagens do conjunto de teste. Agora Para treinamento, com 1000 épocas, não obtivemos erros no reconhecimento e identificação das 100 imagens do conjunto de teste.

CONCLUSÃO

Obtivemos resultados esperados em nossos testes e podemos dizer que seja ela simples ou complexo o procedimento a ser realizado o desempenho se dará através do processo de adaptação e treinamento desta tarefa.

Assim a utilização do conceito de redes neurais artificiais no processo de identificação e classificação de dígitos se torna interessante, pois temos um processo adaptativo com maior precisão, que pode ter maior desempenho a partir de treinamentos mais longos e precisos realizados pela rede. Podemos dizer que o uso da rede neural da uma segurança maior ao processo de reconhecimento porque temos um procedimento de identificação adaptativo e variável.

Uma observação a ser feita em relação ao sistema é que quanto às imagens utilizadas no processo, se estas imagens são de baixa qualidade, elas podem tornar o processo de identificação mais fraco e prejudicar o desempenho. Agora se pelo contrario obtivermos imagens de alta qualidade, as mesmas iram favorecer o processo de identificação e assim confirmar o bom desempenho da rede MLP.

REFERÊNCIAS BIBLIOGRÁFICAS

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

CARPENTER, G. A.; GROSSBERG, S.; MARKUZON, N.; REYNOLDS, J. H.; ROSEN, D. B. **“Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supervised Learning of Analog Multidimensional Maps”**, **IEEE Transactions on Neural Networks**, September-92, Vol. 3, Number 5, pp. 698-713.

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

HOPFIELD, J.J. **“Neural Networks And Physical Systems With Emergent Collective Computational Abilities”**, **Proceedings of the National Academy of Sciences of the U.S.A.**, vol. 79, pp. 2554-2558, 1982.

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

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. de. **“Redes Neurais em Delphi”**, Visual Books, 2nd Edição, Florianópolis, Brasil, 2006.

NASCIMENTO JÚNIOR, C. L.. **“Inteligência Artificial em controle e automação”**, Edgard Blucher, 1nd Edição, São Paulo, Brasil, 2004.

WERBOS, P. J. **“Beyond Regression: New Tools For Prediction And Analysis in The Behavioral Sciences”**, PhD. Thesis, Harvard University, 1974.

WIDROW, B.; LEHR, M. A. **“30 Years of Adaptive Neural Networks: Perceptron, Madaline, and Backpropagation”**, **Proceedings of the IEEE**, pp. 1415-1442, Vol. 78, n^o 9, Sep. 1990.