

UNIVERSIDADE REGIONAL DE BLUMENAU
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
CURSO DE CIÊNCIAS DA COMPUTAÇÃO
(Bacharelado)

**PROTÓTIPO DE UM SISTEMA DE RECONHECIMENTO DE
CARACTERES BASEADO EM REDES NEURAIS**

TRABALHO DE CONCLUSÃO DE CURSO SUBMETIDO À UNIVERSIDADE
REGIONAL DE BLUMENAU PARA A OBTENÇÃO DOS CRÉDITOS NA
DISCIPLINA COM NOME EQUIVALENTE NO CURSO DE CIÊNCIAS DA
COMPUTAÇÃO — BACHARELADO

MÁRCIO RODRIGO CAMPESTRINI

BLUMENAU, JUNHO/2000

2000/1-46

PROTÓTIPO DE UM SISTEMA DE RECONHECIMENTO DE CARACTERES BASEADO EM REDES NEURAIS

MÁRCIO RODRIGO CAMPESTRINI

ESTE TRABALHO DE CONCLUSÃO DE CURSO, FOI JULGADO ADEQUADO
PARA OBTENÇÃO DOS CRÉDITOS NA DISCIPLINA DE TRABALHO DE
CONCLUSÃO DE CURSO OBRIGATÓRIA PARA OBTENÇÃO DO TÍTULO DE:

BACHAREL EM CIÊNCIAS DA COMPUTAÇÃO

Prof. Roberto Heinzle — Orientador na FURB

Prof. José Roque Voltolini da Silva — Coordenador do TCC

BANCA EXAMINADORA

Prof. Roberto Heinzle

Prof. Dalton Solano dos Reis

Prof. Marcel Hugo

AGRADECIMENTOS

Agradeço a todas as pessoas que durante estes anos me apoiaram, incentivando-me a não esmorecer diante dos obstáculos encontrados.

A meus pais e familiares, pelo apoio e incentivo constante a conclusão deste curso.

Ao professor Roberto Heinzle, pelo incentivo, orientação e atenção no desenvolvimento deste trabalho.

A meus amigos, pela incentivo nos momentos de fraqueza e ajuda em determinados momentos do trabalho.

A Deus, sem o qual eu não estaria aqui.

AGRADECIMENTOS	iii
LISTA DE FIGURAS	vi
LISTA DE QUADROS	vii
RESUMO	viii
ABSTRACT	ix
1 INTRODUÇÃO	1
1.1 JUSTIFICATIVA	1
1.2 OBJETIVOS	2
1.3 ORGANIZAÇÃO DO TEXTO	2
2 RECONHECIMENTO DE PADRÕES VISUAIS	3
2.1 INTRODUÇÃO	3
2.2 RECONHECIMENTO DE CARACTERES	3
2.2.1 INTRODUÇÃO	3
2.2.2 PROCESSAMENTO DA IMAGEM	3
2.2.2.1 IMAGENS DIGITAIS.....	4
2.2.2.2 IMAGENS BI-LEVEL	4
2.2.2.2.1 CONECTIVIDADE	4
2.2.2.2.2 OPERAÇÕES COM IMAGENS BI-LEVEL	6
2.2.2.2.2.1 EROÇÃO E DILATAÇÃO	6
2.2.2.2.2.2 ESQUELETIZAÇÃO.....	7
2.2.3 CLASSIFICAÇÃO DA IMAGEM.....	8
2.2.3.1 QUANTIDADE DE ORIFÍCIOS E PONTOS FINAIS DA IMAGEM	9
2.2.3.2 DENSIDADE DA IMAGEM	10
2.2.4 SEGMENTAÇÃO DE CARACTERES.....	10
2.2.4.1 SEGMENTAÇÃO POR REGIÃO	11
2.2.4.2 SEGMENTO POR TEXTURA	11
2.2.4.3 SEGMENTAÇÃO POR CONTORNO	12
2.2.4.4 SEGMENTO POR ENLACE DE BORDAS E DETECÇÃO DE CONTORNOS.....	12
3 REDES NEURAIS	14
3.1 INTRODUÇÃO	14
3.2 O NEURÔNIO BIOLÓGICO	14
3.3 O ELEMENTO DE PROCESSAMENTO	15
3.4 ANALOGIA COM O CÉREBRO	16
3.5 REDES NEURAIS ARTIFICIAIS	16
3.5.1 CAMADAS	17
3.5.2 CONEXÕES.....	17
3.6 TOPOLOGIAS DE REDES NEURAIS	18
3.7 TIPOS DE REDES NEURAIS ARTIFICIAIS	18
3.8 MÉTODOS DE CONTROLE DE APRENDIZADO	19
3.8.1 APRENDIZADO SUPERVISIONADO	19
3.8.2 APRENDIZADO NÃO-SUPERVISIONADO	19

3.9	MODELOS DE REDES NEURAS ARTIFICIAIS	19
3.9.1	MODELO FEEDFORWARD BACKPROPAGATION	21
3.9.2	MODELO PERCEPTRON.....	22
4	DESENVOLVIMENTO DO TRABALHO	23
4.1	ESPECIFICAÇÃO	23
4.1.1	DEFINIÇÃO DA REDE NEURAL	26
4.1.2	TREINAMENTO DA REDE NEURAL	26
4.2	IMPLEMENTAÇÃO	26
4.2.1	VISUAL BASIC.....	26
4.2.1.1	CARACTERÍSTICAS DO VISUAL BASIC	26
4.2.2	APRESENTAÇÃO DO PROTÓTIPO	27
4.2.2.1	ARQUIVO.....	27
4.2.2.2	OPERAÇÕES.....	28
4.2.2.2.1	TREINAR REDE.....	29
4.2.2.2.2	RECONHECER PALAVRA	29
4.2.3	RESULTADOS OBTIDOS	31
5	CONCLUSÃO	33
5.1	LIMITAÇÕES	33
5.2	EXTENSÕES	34
6	REFERÊNCIAS BIBLIOGRÁFICAS.....	35

LISTA DE FIGURAS

FIGURA 1 - REGIÕES (A) 4-CONEXA E (B) 8-CONEXA	5
FIGURA 2 - IMAGENS ANTES DA EROSÃO	7
FIGURA 3 - IMAGENS APÓS A EROSÃO	7
FIGURA 4 - EXEMPLO DE ESQUELETIZAÇÃO	8
FIGURA 5 - EXEMPLOS DA LETRA E	9
FIGURA 6 - NEURÔNIO BIOLÓGICO	14
FIGURA 7 - NEURÔNIO ARTIFICIAL	15
FIGURA 8 - EXEMPLOS DE TOPOLOGIAS DE REDES NEURAIS	18
FIGURA 9 - FUNÇÃO SOMADORA DO ELEMENTO DE PROCESSAMENTO	21
FIGURA 10 - PROCESSO DE FUNCIONAMENTO DO PROTÓTIPO	23
FIGURA 11 - EXEMPLO DA LETRA "A" E SEU RESPECTIVO MAPA DE BITS	24
FIGURA 12 - DIAGRAMA DO RECONHECIMENTO DE UMA PALAVRA	25
FIGURA 13 - O ITEM ABRIR IMAGEM DO MENU ARQUIVO	28
FIGURA 14 - MENU OPERAÇÕES	28
FIGURA 15 - TELA DE TREINAMENTO DA PALAVRA	29
FIGURA 16 - TELA DE RECONHECIMENTO DA PALAVRA	30
FIGURA 17 - PALAVRA RECONHECIDA	30
FIGURA 18 - PALAVRA NÃO RECONHECIDA	31

LISTA DE QUADROS

QUADRO 1 – RELAÇÃO DE VIZINHANÇA ENTRE OS <i>PIXELS</i>	5
QUADRO 2 – ANALOGIA ENTRE O NEURÔNIO BIOLÓGICO E NEURÔNIO ARTIFICIAL.....	16
QUADRO 3 – APLICAÇÕES – REDES NEURAIS.....	20
QUADRO 4 – NÚMERO DE ITERAÇÕES PARA CADA LETRA	31

RESUMO

Neste trabalho é apresentado um estudo sobre o reconhecimento de caracteres através da utilização de redes neurais. Serão abordados aspectos referentes ao processamento de imagem, ao reconhecimento de caracteres e às redes neurais. Ao final, é apresentado um protótipo de software para reconhecimento de caracteres utilizando uma rede neural e os resultados obtidos com essa implementação.

ABSTRACT

This work presents a letter recognition study using neural network concepts. Image processing aspects and neural network utilities are going to be shown. Finally, a software prototype to recognize letters will be presented, including the results and conclusions of this implementation.

1 INTRODUÇÃO

As redes neurais deixaram de ser um sonho para se tornarem realidade na computação. Elas possuem a característica de serem apropriadas ao reconhecimento de padrões, graças à sua capacidade de sofrer aprendizado, modificando seu comportamento frente a um conjunto de estímulos de entrada. Portanto, a rede pode aprender a dar uma resposta específica para um determinado conjunto de estímulos fornecidos. Logo, as redes neurais são muito adequadas para o reconhecimento e classificação de padrões, pois podem se adaptar para responder a um padrão específico [LOE1996].

Devido às suas características inerentes, as redes neurais podem realizar tarefas que não são executadas de forma satisfatória em sistemas tradicionais, mas que para o ser humano são tarefas triviais. Entre as aplicações características para a utilização de redes neurais destacam-se [OSO1991]:

- a) reconhecimento e síntese da fala;
- b) reconhecimento de padrões;
- c) processamento de sinais e eliminação de ruídos.

Os sistemas de reconhecimento ótico de caracteres – OCR (*Optical Character Recognition*) são sistemas desenvolvidos para, de certa forma, reproduzir a capacidade humana de ler textos. Além de servirem como uma forma alternativa de interação homem-máquina, os sistemas de reconhecimento de caracteres tem sido utilizados para a compactação de imagens textuais. Este caso se aplica muito na utilização de sistemas de *fac-símile* para transmissão de textos. Através do reconhecimento dos caracteres em uma imagem a ser transmitida por este tipo de equipamento, pode-se obter a compactação das informações, resultando em economia (menor número de informações transmitidas) [OSO1991].

1.1 JUSTIFICATIVA

Cada vez mais faz-se uso do armazenamento magnético de informações, e também de textos existentes em papel. Apesar de ser uma tarefa corriqueira para o ser humano, o reconhecimento de caracteres, ação necessária para o armazenamento correto desses textos, é uma tarefa bastante complexa aos olhos da informática. Para poder resolver esse problema,

são desenvolvidos sistemas de reconhecimento de caracteres, os quais irão efetuar esse processo. Esse processo é semelhante ao executado facilmente pelo homem, e é baseado numa área da computação conhecida por Inteligência Artificial, que através das Redes Neurais, procura simular o funcionamento do cérebro humano.

1.2 OBJETIVOS

O principal objetivo deste trabalho é desenvolver um protótipo de um sistema de reconhecimento de caracteres aplicando-se redes neurais.

Como objetivos secundários temos:

- a) desenvolvimento de uma rede neural;
- b) estudo dos conceitos de reconhecimento de caracteres.

1.3 ORGANIZAÇÃO DO TEXTO

Abaixo será brevemente descrita a estrutura do trabalho.

Na introdução (capítulo um), é apresentada uma visão geral do trabalho, importância e objetivos.

No segundo capítulo é apresentada a fundamentação teórica do reconhecimento de padrões, abordando mais detalhadamente o reconhecimento de caracteres.

O terceiro capítulo apresenta os conceitos básicos de redes neurais: técnicas envolvidas e principais ferramentas.

O capítulo quatro mostra o desenvolvimento do trabalho, comentando as técnicas e ferramentas envolvidas.

No quinto capítulo são apresentadas as principais conclusões e sugestões para futuros trabalhos.

2 RECONHECIMENTO DE PADRÕES VISUAIS

2.1 INTRODUÇÃO

O objetivo principal da utilização do reconhecimento de padrões visuais é extrair, detectar e identificar elementos existentes em uma cena (conjunto de objetos, pertencentes ao mundo real, que podem ser percebidos a olho nu). Desta maneira, procura-se imitar o ser humano e suas habilidades, tentando criar dispositivos e algoritmos capazes de realizar as mesmas funções humanas como localizar objetos, classificar padrões e detectar relações entre estes. A principal vantagem é a de capacitar máquinas a realizar tarefas até então só possíveis de serem realizadas por seres humanos.

As aplicações de reconhecimento de padrões incluem funções como: compactação de dados, aumento da capacidade de aquisição de dados pelo computador, sensoriamento remoto, controle de qualidade e classificação de objetos, além de uma série de outras aplicações científicas e comerciais.

2.2 RECONHECIMENTO DE CARACTERES

2.2.1 INTRODUÇÃO

Segundo [OSO1991], “os sistemas de reconhecimento de caracteres (OCR) são sistemas desenvolvidos para, de certa forma, reproduzir a capacidade humana de ler textos”. Através desta definição, pode-se entender a importância desse tipo de aplicação para o ser humano.

O funcionamento básico de um sistema OCR traduz-se da seguinte forma: o computador recebe uma imagem, processa essa imagem de maneira que ele possa entendê-la, compara-a a uma base ou a aplica a um processo de reconhecimento, donde retornará uma representação de caracter reconhecido.

2.2.2 PROCESSAMENTO DA IMAGEM

De acordo com [HAR1983], o processamento de imagens é uma área de estudos que tem evoluído constantemente dentro da computação. Esse crescimento é devido

principalmente ao desenvolvimento de equipamentos cada vez mais baratos para a captura e tratamento das imagens digitais.

2.2.2.1 IMAGENS DIGITAIS

O processo de digitalização, segundo [OSO1991], consiste em realizar a aquisição de uma imagem e enviá-la ao computador em um formato adequado para que este possa manipulá-la. As informações visuais são convertidas em sinais elétricos, que são convertidos em valores binários a serem armazenados na memória do computador. Neste processo, os sinais são amostrados espacialmente e quantificados em amplitude, de forma a obter a imagem digital.

No processamento de imagens, a imagem digital é definida como sendo uma matriz formada por $m \times n$ elementos. Cada elemento dessa imagem é denominado *pixel* (*picture element*), que possui, associado, uma informação referente a cor e luminosidade [OSO1991].

2.2.2.2 IMAGENS BILEVEL

Segundo [SCH1997], as imagens *bi-level* contém somente dois níveis de cinza: o preto e o branco. As imagens precisam de apenas 1 bit por *pixel* para serem representadas. Geralmente imagens *bi-level* podem ser produzidas através do processo de limiarização (*thresholding*), que consiste em comprimir as variações de cinza em sua escala até que a imagem fique somente com dois níveis, resultando em um conjunto menor de informações do que o presente na imagem original.

Em geral, o primeiro passo para efetuar o processamento de imagens no formato *bi-level*, e realizar a distinção entre o que é objeto e o que é fundo. Após realizado esse reconhecimento, ocorre a tentativa seguinte de identificar o objeto através de suas características, tais como: moldura, tamanho e orientação, que são propriedades mensuráveis, para em seguida comparar com as medidas encontradas em classes já conhecidas do objeto.

2.2.2.2.1 CONECTIVIDADE

Segundo [SCH1997], o relacionamento entre os *pixels* individuais de uma pequena área, determina muitas características da imagem. O relacionamento mais utilizado é a

conectividade. Considera-se na imagem um *pixel* P , uma linha i e uma coluna j . Visualizando uma pequena região deste *pixel*, pode ser rotulada a vizinhança dos *pixels* [PAR1993]:

- Os vizinhos horizontais a P são os números 3 e 7;
- Os vizinhos verticais são 1 e 5;
- Os *pixels* restantes são vizinhos diagonais.

O quadro abaixo mostra a relação de vizinhança entre *pixels*:

Quadro 1 – relação de vizinhança entre os *pixels*

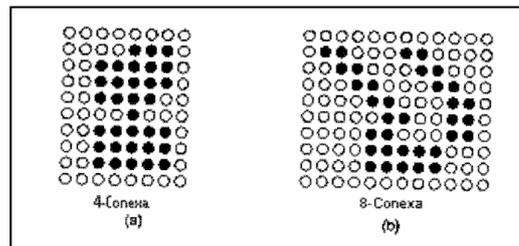
8	1	2
($i-1, j-1$)	($i-1, j-1$)	($i-1, j-1$)
7	P	3
($i-1, j-1$)	($i-1, j-1$)	($i-1, j-1$)
6	5	4
($i-1, j-1$)	($i-1, j-1$)	($i-1, j-1$)

Fonte: [PAR1993]

Deve-se também, em muitos casos, conhecer a conectividade das regiões dos *pixels*. De acordo com [PAR1993], as regiões conexas são definidas como (figura 1):

- 4-Conexa: cada *pixel* da figura pode ser alcançado através de movimentos na horizontal e na vertical.
- 8-Conexa: para alcançar um determinado *pixel* da imagem, são utilizados movimentos horizontais, verticais e diagonais.

Figura 1 - regiões (a) 4-conexa e (b) 8-conexa



Fonte: [PAR1993]

A localização de regiões conexas é de fácil identificação, pois freqüentemente elas correspondem aos objetos ou um grupo de objetos pertencentes à uma imagem. Antes de um objeto ser descrito ou reconhecido, é necessário conhecer a sua posição, e todos os pixels pertencentes ao objeto devem ser identificados.

Para se visualizar os objetos de uma região conexa, assume-se que a região pode ser identificada como o preto sendo representado pelo número zero (0), que representa o objeto, e o branco representado pelo número 255, assumindo-se que este valor represente o *background* da imagem. Os *pixels* da região conexa estão conectados uns aos outros, é necessário encontrar um pixel pertencente à região. Então, quando todos os *pixels* conexos forem encontrados, a procura termina.

2.2.2.2.2 OPERAÇÕES COM IMAGENS BI-LEVEL

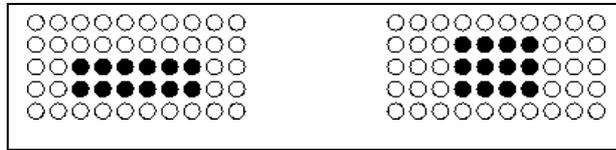
A alteração ou a representação de uma imagem *bi-level* é simples, mas, em virtude do menor número de informações disponíveis em cada pixel para compor a imagem (somente as cores preta e branca), poderá existir uma diminuição da qualidade da imagem, seja em realce, nitidez ou mesmo, em certas ocasiões, a impossibilidade de reconhecimento da mesma.

De acordo com [PAR1993], para manipular imagens *bi-level*, algumas das técnicas utilizadas são a erosão e dilatação, além da técnica de esqueletização, que são descritas a seguir.

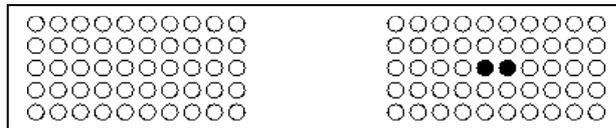
2.2.2.2.2.1 EROSÃO E DILATAÇÃO

A erosão consiste em “arrancar” a camada externa de *pixels* de uma imagem. Ela é uma operação destrutiva realizada sobre a imagem, ou seja, não há como determinar quais os *pixels* a serem removidos e, assim, não há como restaurar. Uma região pode sofrer um processo de erosão repetidamente até desaparecer.

Considerando as duas regiões existentes na figura 2, ambas tem a mesma área (12 *pixels*), mas, após aplicar a erosão de um *pixel*, vê-se que as duas imagens não são idênticas, conforme demonstrado na figura 3.

Figura 2 - imagens antes da erosão

Fonte: [PAR1993]

Figura 3 - imagens após a erosão

Fonte: [PAR1993]

A técnica de dilatação funciona de maneira contrária a erosão: ela adiciona uma nova camada de *pixels* ao redor da imagem.

Essas operações tem como principais objetivos a correção e a comparação de imagens, já que, através da adição (dilatação) e remoção (erosão) de *pixels*, os defeitos existentes em uma imagem serão corrigidos.

De acordo com [PAR1993], existem dois processos que combinam erosão e dilatação:

- a) *opening*: realiza uma erosão seguida de uma dilatação, obtendo uma imagem mais suave que o original.
- b) *closing*: é primeiramente realizada uma dilatação e depois uma erosão. Após esse processo, as lacunas existentes na imagem original estarão fechadas.

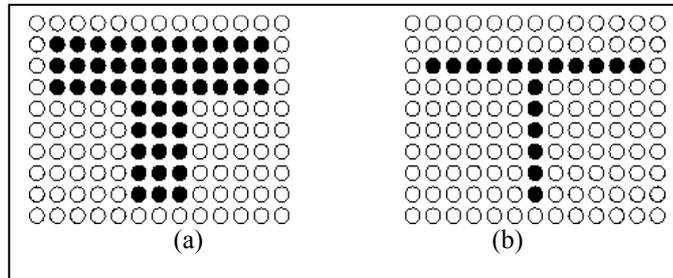
2.2.2.2.2 ESQUELETIZAÇÃO

As imagens que são compostas de muitas linhas, como as imagens de texto, tem a posição e a orientação das linhas e seus relacionamentos umas com as outras extremamente significativos. *Pixels* extras são bagagem indesejável que complicam a análise, tornando necessária a sua remoção.

O processo efetuado para isso é conhecido por esqueletização (*thinning* ou *skeletonizing*), e tem por objetivo remover da imagem todos os *pixels* extras que ela possui,

obtendo assim uma imagem mais simples com a mesma representatividade da imagem original (figura 4).

. **Figura 4** - exemplo de esqueletização (a) imagem original (b) esqueleto da imagem



Fonte: [PAR1993]

A esqueletização é executada em dois momentos: inicialmente a imagem é analisada e os *pixels* a serem eliminados são marcados. Após esse momento, a imagem é novamente varrida e os *pixels* então são efetivamente removidos. Estes dois processos se repetem até não existirem mais *pixels* redundantes na imagem, obtendo assim o esqueleto da imagem.

A esqueletização possui algumas características que devem ser levadas em consideração [PAR1993]:

- a) o objeto deve consistir-se de regiões delgadas, um *pixel* de largura;
- b) os *pixels* que compõem o esqueleto devem estar próximos ao centro, cruzando uma seção da imagem;
- c) os *pixels* do esqueleto devem estar conectados uns aos outros para formar o mesmo número de regiões da imagem original.

Contudo, os métodos de esqueletização de imagens não são perfeitos, pois a maioria dos algoritmos corroem a fronteira do objeto, sendo completamente sensíveis a pequenas irregularidades presentes no contorno.

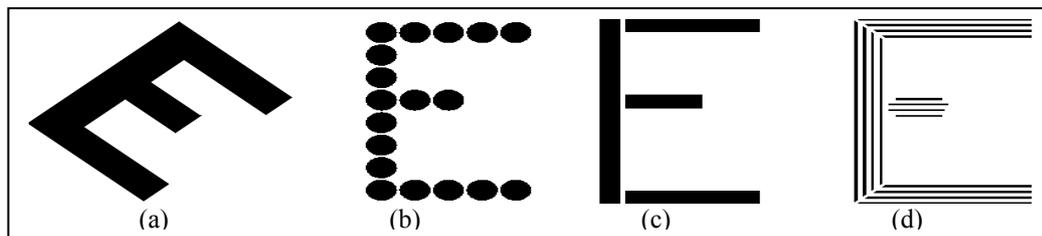
2.2.3 CLASSIFICAÇÃO DA IMAGEM

Um dos maiores problemas encontrados no processo de reconhecimento de caracteres é encontrar uma característica que possa identificar o caracter dentro de um conjunto.

Encontrada essa característica, o processo de reconhecimento do caracter se torna mais simples de ser resolvido.

Na figura 5, são mostrados quatro exemplos diferentes de letras “E” que são reconhecidas visualmente, mas onde os *softwares* de reconhecimento fracassam, em virtude da enorme gama de fontes e estilos que podem ser aplicados.

Figura 5 - exemplos da letra E (a) girada em 45° (b) formada por pontos individuais (c) formado por linha não conectadas (d) fonte estilizada



Fonte: [PAR1993]

Alguns dos métodos utilizados para classificar uma imagem, segundo [SCH1997], são os descritos a seguir.

2.2.3.1 QUANTIDADE DE ORIFÍCIOS E PONTOS FINAIS DA IMAGEM

Algumas características básicas dividem os caracteres em classes, permitindo a análise de cada classe por diferentes métodos. O número de orifícios divide o grupo de caracteres em três partes [SCH1997]:

- a) os caracteres B e 8 possuem dois orifícios;
- b) os caracteres A, D, O, P, Q, a, b, d, e, g, o, p, q, 6, 9, 0, e, às vezes, o número 4 possuem um orifício;
- c) os caracteres restantes não possuem orifício.

O número de pontos finais pode também ser usado para agrupar imagens de caracteres. Um ponto final é frequentemente, mas não sempre, indicado por um *pixel* 1-conexo na versão delgada da imagem. A letra A, por exemplo, tem dois pontos finais, localizados na base. A letra X tem quatro pontos, um em cada canto. Combinando essa característica e o número de orifícios consegue-se uma classificação melhor do que utilizados isoladamente. Por outro lado, existem caracteres que possuem ponto final e orifício, sendo eles: O, o e 0 (zero).

2.2.3.2 DENSIDADE DA IMAGEM

Segundo [OSO1991] o cálculo da densidade da imagem é uma técnica empregada e que obtém bons resultados. Ela consiste em dividir o caractere em regiões (comumente 9 ou 16, através de uma matriz 3x3 ou 4x4), e verificar a quantidade de pontos pretos existentes nas regiões. Assim pode-se detectar a presença ou ausência de pontos em determinadas regiões, o que simplifica a classificação do caractere a ser reconhecido.

2.2.4 SEGMENTAÇÃO DE CARACTERES

De acordo com [SCH1997], a segmentação tem por objetivo obter, a partir de uma imagem digitalizada, um conjunto de primitivas – os segmentos significativos – que contém a informação semântica relativa a imagem. Geralmente, as primitivas utilizadas são naturais: contornos ou regiões.

A segmentação é um passo extremamente necessário e crítico em um sistema OCR. Qualquer erro cometido durante esta etapa não poderá ser corrigido no passo seguinte, o reconhecimento. Em outras palavras, erros de segmentações afetam diretamente o reconhecimento [CAS1983].

Em [CAS1983] encontram-se alguns métodos de segmentação:

- a) Método de segmentação TSSRS (*Two-Stage Segmentation and Registration Scheme*), que tem como base dois estágios de segmentação:
 - para ser possível a segmentação, primeiramente verifica-se as informações referentes aos objetos e seus limites;
 - após, algumas rotinas especiais são aplicadas à região de fronteira da imagem, sendo que a segmentação da região é efetuada analisando-se o objeto e buscando os limites que ele possui.
- b) segmentação vertical: este método de segmentação baseia-se em uma grade padrão, que é determinado pelo campo e pela linha base, sendo também chamada de segmentação por campo;
- c) segmentação de caracteres que se tocam: neste método, primeiramente, a região é dividida em várias zonas horizontais. Dentro de cada zona, um conjunto de lacunas

verticais contíguas é procurado. O ponto mais a esquerda e mais a direita de cada lacuna são registrados, respectivamente, como “L” e “R”;

- d) segmentação horizontal: neste método, são utilizadas as referências das coordenadas da imagem para definir a segmentação da região confinada.

Segundo [FAC1993], existem quatro formas diferentes de se efetuar a segmentação de imagens:

- a) segmentação por região;
- b) segmentação por textura;
- c) segmentação por contorno; e
- d) segmentação por enlace de bordas e detecção.

2.2.4.1 SEGMENTAÇÃO POR REGIÃO

Na segmentação por região, a região de uma imagem é um conjunto de pontos ligados, podendo-se chegar a outro ponto da região partindo de um ponto inicial. Em geral, as regiões que se deseja detectar são homogêneas ou seja, possuem alguma propriedade constante (em geral, a continuidade do nível de cinza) [SCH1997].

Existem dois métodos para segmentação por região:

- a) limiarização (*thresholding*): separa as regiões de uma imagem quando a mesma apresenta duas classes, o fundo e o objeto, estando baseada no histograma para realizar esta operação. Quando os níveis de cinza são bem distintos, formam no histograma picos separados por vales, facilitando o processo de limiarização;
- b) divisão e fusão (*split and merge*): os *pixels* são agrupados, de maneira a formar uma região, onde será verificada uma condição de homogeneidade. Aplicam-se sucessivas divisões e fusões até obter-se regiões que satisfaçam o critério definido.

2.2.4.2 SEGMENTO POR TEXTURA

Deve permitir o destaque dos motivos da imagem, ou seja, das características existentes em uma imagem, considerando propriedades de regularidade e repetição. O motivo deve ser um agrupamento aleatório de subconjuntos da imagem. Os três métodos usados na segmentação por textura são [SCH1997]:

- a) técnica estatística: permite caracterizar uma textura do tipo suave, irregular, granulada, etc.;
- b) técnica estrutural: preocupa-se com a disposição das primitivas da imagem, como exemplo, textura baseada em linhas paralelas dispostas regularmente;
- c) técnica espectral: levam às propriedades do espectro de Fourier e permitem detectar a periodicidade global, identificando altas energias.

2.2.4.3 SEGMENTAÇÃO POR CONTORNO

Um contorno é uma mudança brusca no nível de cinza entre duas regiões relativamente homogêneas [SCH1997]. Existem dois métodos aplicáveis à segmentação por contorno:

- a) detecção de pontos: executa-se a segmentação através de pontos particulares relevantes à imagem, que devem conter muita informação e não devem ser deteriorados pelo ruído. Para ser eficiente, deve ser eliminado o ruído presente na imagem;
- b) detecção de contornos: ocorre através de um operador de derivada local. O gradiente $G(x, y)$ e o laplaciano $L(x, y)$, respectivamente operadores de primeira e segunda ordem, podem ser usados. Para maior nitidez, é utilizado o processo de limiarização da imagem.

2.2.4.4 SEGMENTO POR ENLACE DE BORDAS E DETECÇÃO DE CONTORNOS

De acordo com [SCH1997], as técnicas descritas anteriormente não são muito precisas devido a ruídos, cortes no contorno e outras perturbações existentes nas imagens. Abaixo são descritas técnicas que visam aumentar a precisão da segmentação:

- a) perseguição de contornos: necessita de uma ferramenta que permita a análise das características dos pixels que possam pertencer ao contorno e escolher o mais relevante. O princípio deste algoritmo reside na busca de pontos similares do contorno, onde cada ponto deve respeitar um critério de pertinência ao contorno escolhido da vizinhança do ponto anterior;
- b) transformada em *Hough*: é aplicada quando tem-se informações precisas sobre a forma da curva. Os dados de base da transformada de *Hough* são geralmente pontos

de uma imagem obtidos através das transformações de gradiente e do *thresholding*. A idéia é aplicar na imagem uma transformação de forma que todos os pontos de uma mesma curva sejam mapeados num único ponto do espaço, usando os parâmetros da curva procurada.

- c) técnicas do grafo: método global baseado na representação de segmentos de reta na forma de uma estrutura de grafo e na pesquisa de menor custo correspondente a segmentos significativos. Esse procedimento é mais complicado, porém tem a vantagem de fornecer resultados satisfatórios na presença de ruídos.

3 REDES NEURAIS

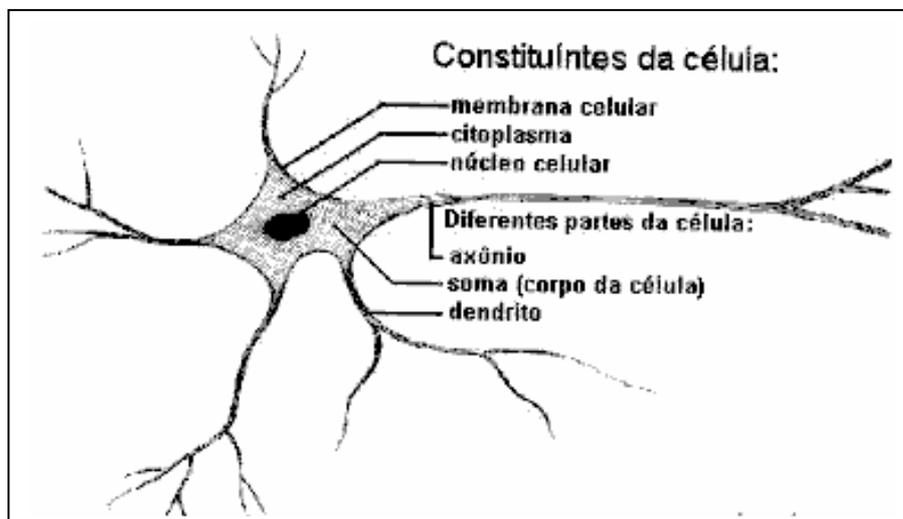
3.1 INTRODUÇÃO

As redes neurais são “máquinas” concebidas para trabalharem segundo o mesmo processo de um sistema nervoso. Seus elementos de processamento são neurônios artificiais, interconectados, que efetuam a soma das entradas e geram uma saída através da aplicação dessa soma a uma função de transferência.

3.2 O NEURÔNIO BIOLÓGICO

O neurônio (figura 6), basicamente, capta o estímulo nos seus dendritos e o transmite pelo axônio para que o estímulo atinja um outro neurônio ou célula. Essa captação do estímulo é realizada por receptores sensoriais especializados que, segundo [TAF1995], podem ser mecânicos, térmicos, nociceptores, químicos, entre outros.

Figura 6 - neurônio biológico



Os estímulos entram nos neurônios através das sinapses (conexões existentes entre os neurônios), que conectam os dendritos de um neurônio aos axônios de outro neurônio ou com o sistema nervoso. Esses estímulos, na forma de sinais, são passados para o somador (corpo celular), que os adiciona e os aplica a um sensor de limiar, responsável por definir o nível de energia mínima de entrada acima do qual o neurônio irá disparar.

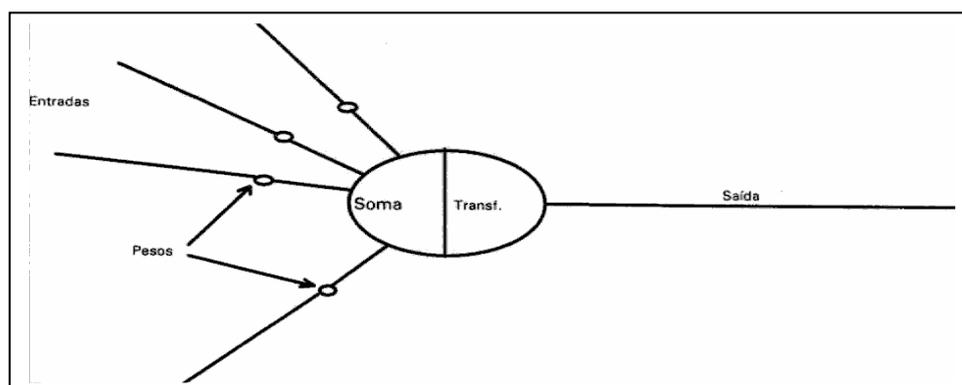
Se a soma dos sinais é maior que o nível de limiar, o neurônio envia energia através do axônio, de onde a energia é transmitida para outras sinapses (ou às vezes, diretamente para axônios ou corpos celulares, dependendo da conexão do axônio original), ou também realimenta a sinapse original; mas se a soma for menor que o limiar definido, nada acontece.

As sinapses tem grande importância na memorização da informação, pois a maioria destas é que retém a informação.

3.3 O ELEMENTO DE PROCESSAMENTO

Um processo semelhante ao que ocorre no cérebro humano ocorre nos neurônios de uma rede neural (também conhecidos como elementos de processamento (figura 7) ou nodos): os sinais entram nos nodos através de conexões ponderadas, onde a informação está armazenada. Cada sinal que entra pelo sistema, seja ele positivo ou negativo, é multiplicado por um número, ou peso, que indica sua influência na saída do nodo. É feita a soma ponderadas dos sinais, que produz um certo nível de atividade. Como num neurônio natural, uma resposta só é gerada num nodo se esse nível de atividade ultrapassar um limite. Se não atingir esse nível, nenhuma saída é gerada, o que indica que não há informação a ser fornecida pelo nodo.

Figura 7 - neurônio artificial



A operação dos elementos processados envolve os seguintes passos:

- a) classificação dos sinais de entrada;
- b) computação das somas ponderadas individuais;
- c) combinação das somas ponderadas individuais;

d) aplicação da função de transferências na soma.

3.4 ANALOGIA COM O CÉREBRO

No quadro 2 é exibida uma analogia do neurônio artificial com o neurônio biológico.

Quadro 2 – analogia entre o neurônio biológico e neurônio artificial

Neurônio biológico	Elementos de processamento
Dendritos	Entradas
Sinapses	Conexões ponderadas
Somador	Função somadora
Limiarização (<i>threshold</i>)	Função de transferência
Axônio	Saída global

3.5 REDES NEURAIS ARTIFICIAIS

Uma rede neural é um modelo matemático que consiste em vários nodos, semelhantes aos neurônios. Estes nodos são organizados em camadas, onde cada um deles têm conexões ponderadas com outros elementos, sendo que estas conexões podem tomar diferentes configurações, dependendo da aplicação desejada [LOE1996].

Geralmente, uma rede neural é constituída de uma camada de entrada, responsável por receber os estímulos, e de uma camada de saída que irá produzir a resposta. Algumas redes podem ter uma ou mais camadas intermediárias, denominadas camadas ocultas. A capacidade computacional de uma rede neural está nas conexões entre os nodos.

Similarmente às redes biológicas, as redes neurais podem ser organizadas de diversas maneiras. Entretanto, enquanto as redes biológicas tem como uma das características o processamento paralelo, as redes neurais artificiais implementadas em máquinas com processamento seqüencial são essencialmente seriais [TUR1993].

Basicamente, uma rede neural artificial é composta de uma coleção de neurônios (nodos), agrupados em camadas, tendo a seguinte estrutura.

3.5.1 CAMADAS

Uma camada de uma rede neural é uma coleção de neurônios com a mesma função.

Basicamente, as camadas de uma rede neural podem ser classificadas em:

- a) entrada: o processamento é realizado sobre valores numéricos, onde, em alguns casos, cada valor de entrada corresponde a um único atributo [SIL1999]. O pré-processamento é necessário quando da análise de problemas envolvendo atributos qualitativos ou imagens, para efetuar sua conversão para valores numéricos equivalentes [TUR1993];
- b) saída: onde encontra-se a solução para o problema apresentado. Tal qual a camada de entrada, esta só pode assumir valores numéricos. Nos casos em que o pré-processamento é efetuado na camada de entrada, não é exigido também na camada de saída, porém é comum a ocorrência em ambas as camadas.

3.5.2 CONEXÕES

Em relação às conexões, podem ser feitas as seguintes considerações:

- a) peso das conexões: os pesos são elementos fundamentais à rede. Eles exprimem a força relativa, em valor numérico, dos dados alimentados ou transmitidos às várias conexões. Segundo [FAU1994], são eles que exprimem a importância relativa de cada entrada de um neurônio artificial;
- b) tipos de conexões: tem-se as redes *feedback*, onde o resultado tanto pode ser encaminhado para a camada posterior quanto para a camada anterior; as redes *feedforward*, onde o resultado é enviado somente para a próxima camada; e as redes *feedlateral*, onde o resultado do processamento é enviado para os elementos da mesma camada, devido à presença de conexões laterais [TAF1995];
- c) função soma: efetua os cálculos necessários para estabelecer a média dos pesos que compreendem as conexões dos neurônios da camada de entrada [TAF1995];
- d) função de transferência: analisando o resultado de entrada, o neurônio produz ou não um resultado de saída. A partir disso então, a função de transferência modifica o valor de saída do neurônio para que o mesmo respeite uma determinada faixa de

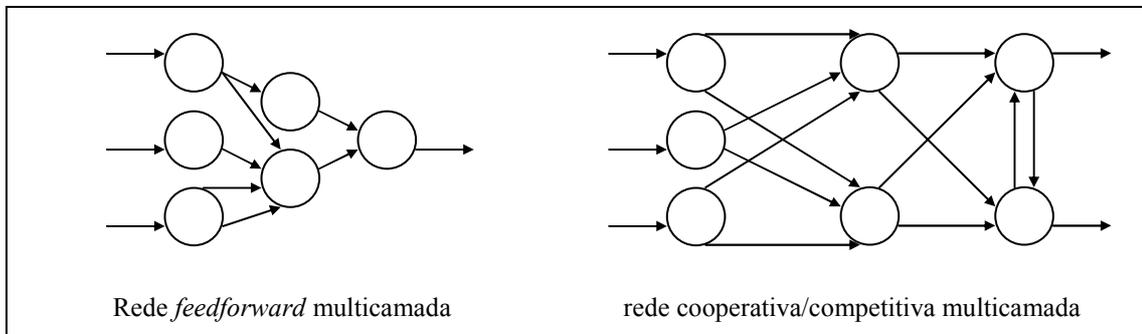
domínio [TAF1995].

3.6 TOPOLOGIAS DE REDES NEURAIIS

A topologia de uma rede é a maneira com que uma rede neural está organizada.

As redes neurais artificiais incluem vários neurônios, interconectados de diferentes maneiras, o que possibilita classificá-las entre as topologias conhecidas (figura 8) [KAR1996].

Figura 8 - exemplos de topologias de redes neurais



Fonte: [KAR1996]

3.7 TIPOS DE REDES NEURAIIS ARTIFICIAIS

Com relação a associatividade, existem dois tipos de associação: auto-associativa, que correlaciona um padrão a ele mesmo, e hetero-associativa, que relaciona um padrão de entrada a um padrão de saída [LOE1996]:

- a) rede auto-associativa: segundo [LOE1996] as redes auto-associativas tem como principal característica a possibilidade de gerar um padrão completo, sem as distorções de um padrão ruidoso (com problemas). Cada tipo de rede auto-associativa tem uma maneira diferente de padrão de aprendizado, e pode manter somente um certo número de padrões em sua memória.
- b) rede hetero-associativa: de acordo com [TAF1995], uma rede de memória hetero-associativa mapeia um conjunto de padrões (entradas) relacionados a outros (saídas). É típica para casos em que uma determinada entrada já esteja associada a uma determinada saída.

3.8 MÉTODOS DE CONTROLE DE APRENDIZADO

A monitoração é o processo de controle de aprendizado da rede. Este processo de aprendizado acontece, basicamente, de duas formas: o aprendizado supervisionado e o aprendizado não-supervisionado [TAF1995].

3.8.1 APRENDIZADO SUPERVISIONADO

Durante a sessão de treinamento, um estímulo de entrada é aplicado e resulta numa saída. Esta saída é então comparada com um padrão inicial informado a rede. Caso a resposta atual da rede seja diferente da resposta fornecida pelo padrão, a rede gera um sinal de erro, que é utilizado para a correção dos pesos atuais da rede [KAR1996].

Concluído esse processo, o estímulo inicial é fornecido novamente à rede, resultando numa nova saída. Se essa saída coincidir com o padrão informado, a rede estará treinada para esse estímulo; caso contrário, todo o processo de correção será repetido, efetuando-se novo cálculo dos pesos e nova apresentação do estímulo inicial [KAR1996].

3.8.2 APRENDIZADO NÃO-SUPERVISIONADO

De acordo com [SIL1999], as redes não-supervisionadas tem a capacidade de determinar uma correlação entre os possíveis padrões de entrada e são particularmente úteis nos problemas em que as entradas variam com o tempo de forma conhecida. Segundo [SIL1999], este mecanismo pode ser considerado como sendo a capacidade da rede de abstrair correlações entre os estímulos de modo a obter as respostas desejadas.

3.9 MODELOS DE REDES NEURAIIS ARTIFICIAIS

No quadro 3 são apresentadas os principais modelos de redes neurais e algumas características e aplicações.

Quadro 3 – Aplicações – Redes Neurais

Modelo	Aplicações Básicas	Ano	Vantagens	Desvantagens
Adaline/Madaline	Filtragem de sinal adaptativo, equalização adaptativa	1960	Rápida e fácil de implementar	Somente é possível classificar espaços linearmente separáveis
Adaptive Resonance Theory (ART)	Reconhecimento de padrões	1983	Capacidade de aprender novos padrões	Natureza dos exemplares categóricos podem mudar com o aprendizado
Backpropagation Perceptron	Reconhecimento de padrões, filtragem de sinal, controle robótico, compressão de dados, etc.	1974 - 1986	Operação rápida, bem compreendida e com muitas aplicações de sucesso	Tempo de treinamento longo
BAM – Memória Associativa Bidirecional	Hetero-associativa	1987	Simples, prova clara da estabilidade dinâmica	Capacidade de armazenamento, precisão e recuperação pobres
Boltzman Machine, Cauchy Machine	Reconhecimento de padrões, otimização	1984	Capaz de formar representação ótima das características dos padrões	Tempo de aprendizado longo
Barin-State-is-a-Box (BSB)	Revocação auto-associativa	1977	Bom desempenho	Incompleta em termos de aplicações em potencial
Hopfield	Evocação auto-associativa, otimização	1982	Simples e estabilidade dinâmica	Incapaz de aprender novos estados
Neocognitron	Reconhecimento de caracteres manuscritos/imagens	1975	Capaz de reconhecer padrões independente da escala, rotação e translação.	Estrutura complexa
Quantização de vetor de aprendizagem	Revocação auto-associativa, compressão de dados	1981	Capaz de auto-organizar vetores	Treinamento lento
Recurrent	Controle robótico, reconhecimento da fala...	1987	Excelente para classificação, mapeamento de informações variando no tempo	Rede complexa
Redes de ligações funcionais	Classificação, mapeamento	1988	Rápida para treinar	Não é claro o modo de identificar funções adotadas
Time-Delay	Reconhecimento de fala	1987	Rápida operação	Janela fixada para a atividade temporal representada

3.9.1 MODELO *FEEDFORWARD BACKPROPAGATION*

O modelo *feedforward backpropagation* surgiu em meados da década de 80 e constitui, segundo pesquisadores [WIL1995 *apud* SIL1999], a mais difundida e largamente usada entre todas as arquiteturas e modelos de redes neurais conhecidas. Como principais novidades deste modelo em relação aos anteriores estão as múltiplas camadas, com a possibilidade de valores de entrada e saída contínuos, tais como 1,0001 até 2,0009. Ainda assim, existem limitações quanto ao número de camadas, pois, a partir de certa quantidade, as camadas anteriores são repetidas, de forma a não demonstrar avanço no processo de convergência [LOE1996].

As aplicações modeladas nesta topologia necessitam de padrões de entrada e saída, para os quais a rede irá convergir e se estabilizar (treinamento ou aprendizado). Este aprendizado é supervisionado, pois para cada padrão de entrada está associado um padrão de saída desejado [LUN1999].

Os neurônios das camadas ocultas concedem à rede capacidade de abstração e generalização, ou seja, é capaz de classificar um padrão complexo mesmo quando este não pertenceu ao conjunto de treinamento, portanto, a rede é imune a pequenas falhas.

De uma maneira sistemática, a forma com que os valores são passados de uma camada a outra é a seguinte: a soma ponderada de cada elemento de processamento com os pesos associados às conexões existentes entre os neurônios, feita pelo módulo somador, é levada ao módulo de ativação, que gera um sinal de saída [SIL1999].

A função somadora do neurônio é definida por (figura 9):

Figura 9 - função somadora do elemento de processamento

$$Y_{ij} = \sum_{i=1}^n X_i * W_{ij}$$

Onde:

a) Y_j = saída da rede neural do elemento j ;

- b) Σ = somatório dos valores de entrada multiplicados por seus pesos;
- c) X_i = valor de entrada do índice i ;
- d) W_{ij} = peso associado ao elemento j de índice i .

O treinamento deste modelo consiste em ajustar os pesos de conexões das camadas para que o conjunto de entradas atinja o conjunto de saídas desejado. Para tanto, o algoritmo executa vários passos, a saber:

- a) introduzir os pares de entrada;
- b) calcular as saídas (função soma e transferência);
- c) verificar se as saídas obtidas estão próximas das saídas desejadas, considerando a taxa de erro;
- d) caso o erro seja superior ao parametrizado para o modelo, corrigir os pesos;
- e) repetir os passos anteriores até que a rede assimile os padrões a ela submetidos.

A revocação da rede é feita com os pesos armazenados em arquivos. Nesta etapa, um novo padrão de entrada é oferecido a rede. Ela irá, então, computar e gerar a nova saída.

Sua rápida operacionalização, apresentando capacidade de generalização, robustez e abstração são as suas principais vantagens. O fato de requerer um longo tempo de treinamento, em virtude da necessidade de estabilização e convergência da rede é a sua desvantagem, segundo Loesch [LOE1996].

3.9.2 MODELO PERCEPTRON

Segundo [SIL1999], o Perceptron foi a primeira rede neural a emergir. Possui o conceito substancial de neurônio artificial o qual ainda hoje é usado, onde cada neurônio computa uma soma ponderada de suas entradas, e passa esta soma em uma função não-linear com limiarização. A criação do modelo Perceptron é creditada a Frank Rosenblatt, que o criou para que servisse como ferramenta de pesquisa de modelagem de possíveis mecanismos cerebrais e para o reconhecimento de padrões.

4 DESENVOLVIMENTO DO TRABALHO

4.1 ESPECIFICAÇÃO

Neste trabalho está sendo proposta a automatização do processo de reconhecimento de palavras em imagens previamente digitalizadas. Esta operação irá envolver desde o preparo da imagem com a palavra até a apresentação da palavra reconhecida (figura 10).

O protótipo desenvolvido irá efetuar o reconhecimento de imagens criadas com fonte *Courier New*, maiúsculas, tamanho 12, utilizando os efeitos negrito e itálico, além da fonte normal. Não foram reconhecidas letras minúsculas, caracteres acentuados, caracteres especiais e números.

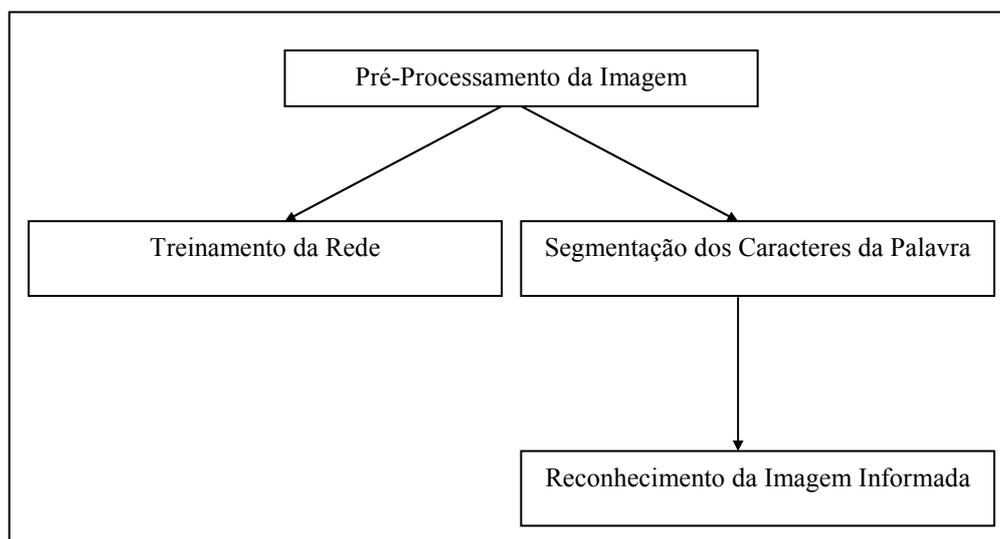


Figura 10 – Processo de funcionamento do protótipo

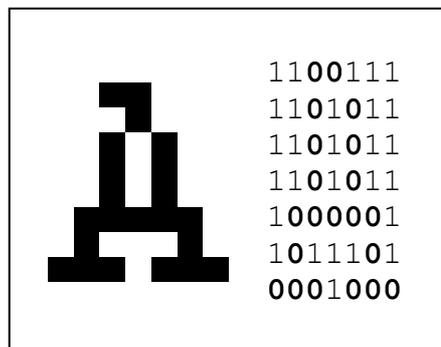
Existem dois processos distintos a serem efetuados pelo protótipo: o treinamento da rede e o reconhecimento da palavra.

O treinamento da rede é efetuado utilizando-se imagens no formato *bitmap* (mapa de bits) com fonte *Courier New*, tamanho 12, utilizando-se dos efeitos normal, itálico e negrito (para melhorar o grau de generalização da rede), representando as 26 letras do alfabeto. As imagens foram desenhadas no Adobe Photoshop 5.0.

Para eliminar possíveis ruídos existentes nas figuras antes de submetê-las ao processo de treinamento da rede, foi aplicado um algoritmo para efetuar a binarização da imagem. Esse algoritmo utilizou, como fator de limiarização, a cor cinza, que representa a metade da escala normal utilizando-se 256 cores. Os valores que estiverem acima desse valor, ou seja, as cores que estiverem entre o cinza e o branco, serão convertidas para branco, e as cores que estiverem na faixa existente entre o cinza e o preto, serão convertidas para o preto.

Após efetuado a binarização da imagem, é gerado um arquivo texto contendo o mapa de bits da imagem a ser aplicada (figura 11).

Figura 11 – exemplo da letra “A” e seu respectivo mapa de bits



Essa imagem é então aplicada a uma rede neural modelada de acordo com o especificado mais adiante, na seção sobre a definição da rede neural. Quando a rede convergir e estabilizar frente ao padrão desejado, será criado um arquivo texto armazenando os pesos de cada conexão para a padrão selecionado.

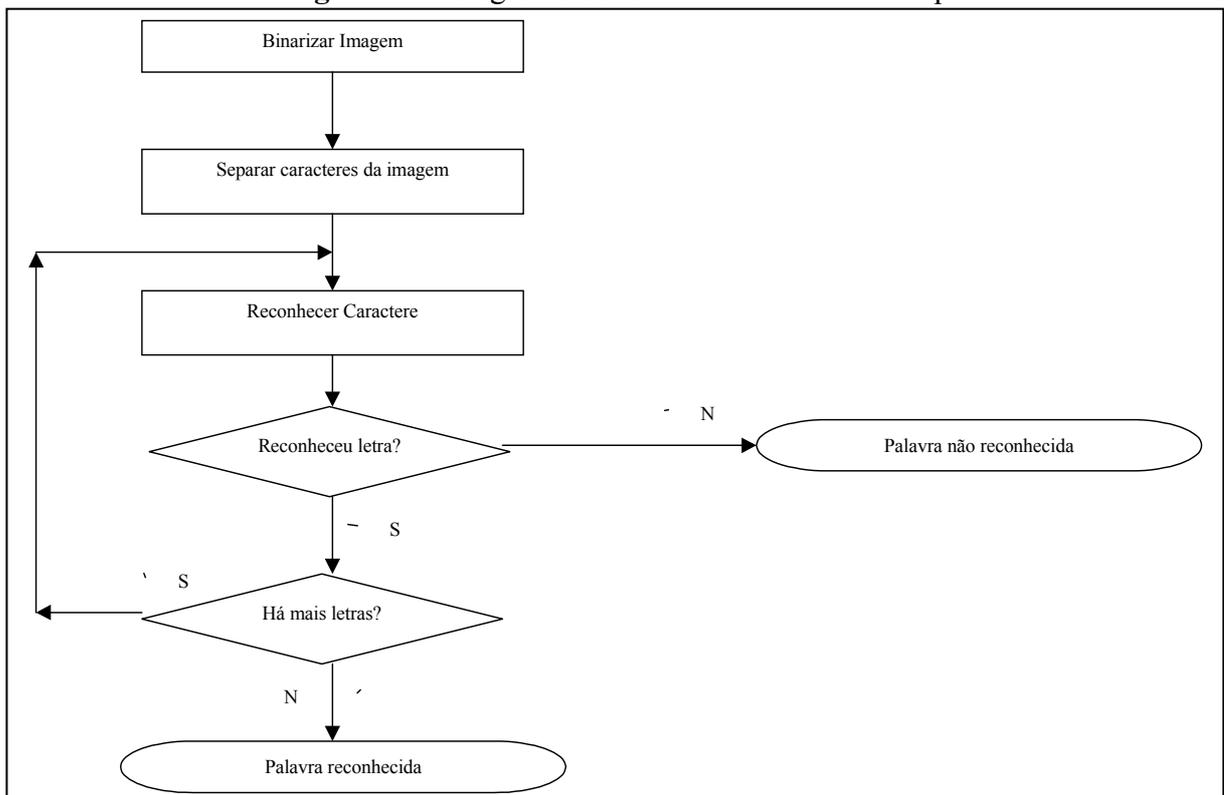
O outro processo executado pelo sistema é o processo de reconhecimento de uma imagem contendo uma palavra (figura 12). Para esse processo, também foram utilizadas imagens com as características já descritas quando da descrição do processo de treinamento da rede neural.

Inicialmente é realizado, também, a binarização da imagem para eliminar possíveis ruídos existentes na figura. Quando do término da binarização, é realizada a segmentação da palavra selecionada, isolando os caracteres existentes nela. Tendo em vista que o protótipo

apresentado somente trabalha com fonte de tamanho fixo, o algoritmo de segmentação considera espaços fixos (7 *pixels*) para a separação das letras [OSO1991]. Para cada letra identificada na palavra é gerado um arquivo em formato texto contendo o mapa de bits da mesma, que será utilizado durante a etapa seguinte, de reconhecimento.

No reconhecimento dos caracteres da palavra, o mapa de bits gerado para cada uma das letras é aplicado à rede neural, iniciando-se o processo sempre pela letra “A”. Caso a letra não seja reconhecida, será testada contra o padrão da letra “B”, e assim sucessivamente, até que se identifique qual a letra está sendo analisada. Caso não seja possível identificar a letra entre as 26 letras do alfabeto, a palavra é imediatamente definida como não reconhecida. Para o processo de reconhecimento foi escolhida uma rede neural, devido a sua grande capacidade de generalização, utilizando como modelo o *Feedforward Backpropagation*, por se tratar de uma rede própria para o reconhecimento geral de padrões [LOE1996].

Figura 12 – diagrama do reconhecimento de uma palavra



4.1.1 DEFINIÇÃO DA REDE NEURAL

Para o desenvolvimento deste trabalho optou-se por utilizar o modelo de rede neural *Backpropagation*. Esta rede tem sua estrutura composta por 56 neurônios de entrada, que representam uma matriz de 8 x 7 de cada letra a ser reconhecida, uma camada oculta composta por 28 neurônios, e uma camada de saída contendo 7 neurônios. Esses sete neurônios da saída representam o valor binário do caractere reconhecido de acordo com a tabela ASCII.

A rede foi treinada de maneira que, para cada caractere informado na camada de entrada, existisse apenas uma combinação possível de neurônios ativos, que correspondesse às informações iniciais da rede.

4.1.2 TREINAMENTO DA REDE NEURAL

Para o treinamento da rede foram utilizados figuras representando as vinte e seis letras do alfabeto, maiúsculas. As imagens utilizadas foram criadas no formato *bitmap* com fonte *Courier New*, tamanho 12, utilizando-se dos efeitos normal, itálico e negrito. Com isso, obteve-se uma matriz de pesos que representasse cada letra do alfabeto, uma vez que o treinamento foi efetuado para cada letra individualmente.

4.2 IMPLEMENTAÇÃO

4.2.1 VISUAL BASIC

O ambiente de desenvolvimento escolhido para o desenvolvimento do protótipo foi o Visual Basic (versão 5.0, da Microsoft), já que o autor deste está familiarizado com o mesmo, e tendo em vista que o ambiente também se adequa aos objetivos deste protótipo.

4.2.1.1 CARACTERÍSTICAS DO VISUAL BASIC

McKelvy [MCK1997], em seu livro, descreve algumas das características do Visual Basic:

- a) descendente do Basic;

- b) programação orientada a eventos;
- c) linguagem compilada e não interpretada (embora essa opção possa ainda ser escolhida);
- d) desenvolvimento rápido de aplicações;
- e) utilização de SQL para acesso a bases de dados;
- f) conectividade através de ODBC.

Também podem ser citados alguns recursos disponíveis na linguagem, tais como:

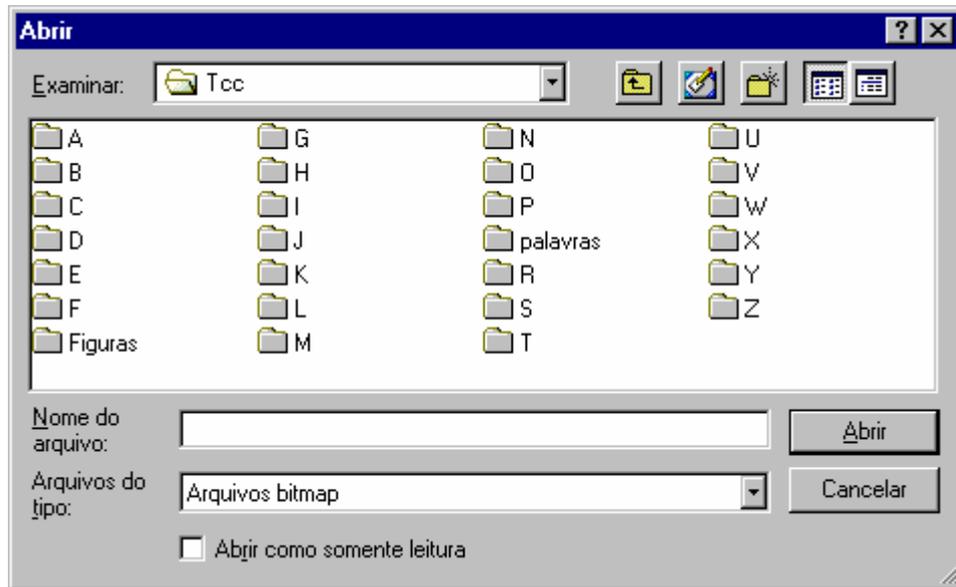
- a) suporte a controles OCX;
- b) modelos de aplicativos e formulários;
- c) criação de bibliotecas de funções (DLLs – *Dynamic Link Libraries*);
- d) assistentes para criação de formulários, classes.

4.2.2 APRESENTAÇÃO DO PROTÓTIPO

Ao iniciar o protótipo, o usuário irá deparar com uma janela e um menu (figura 12), apresentados segundo o padrão *windows*.

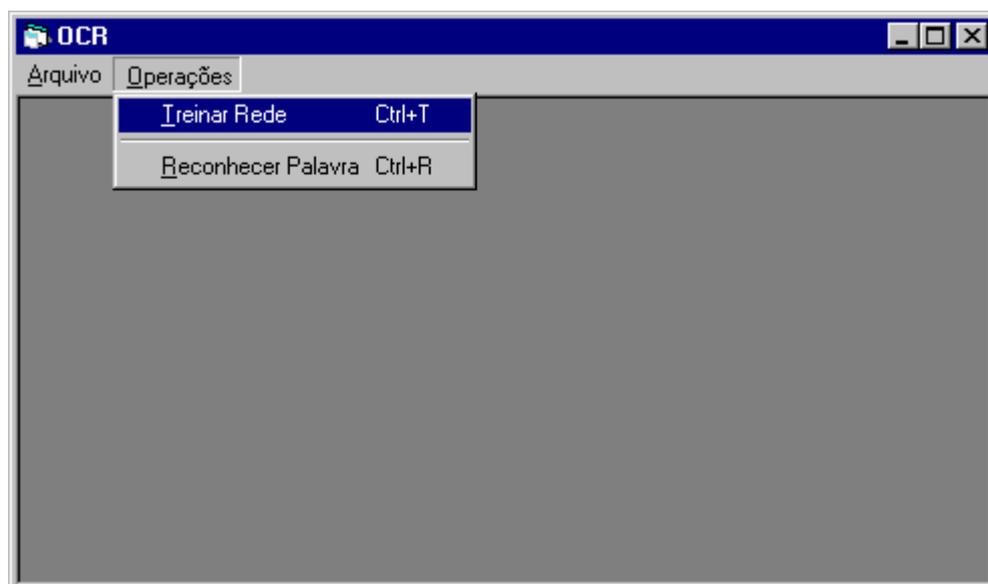
4.2.2.1 ARQUIVO

Na opção **Arquivo** do menu do protótipo, temos a opção “**Abrir Imagem**” que irá exibir uma caixa de diálogo (figura 13) para selecionar a imagem desejada para o reconhecimento. Deve salientar também que essa opção só estará ativa caso o usuário já tenha acessado a tela de reconhecimento.

Figura 13 – O item Abrir Imagem do menu Arquivo

4.2.2.2 OPERAÇÕES

Neste menu encontram-se as opções para Treinar a Rede e Reconhecer a Palavra (figura 14).

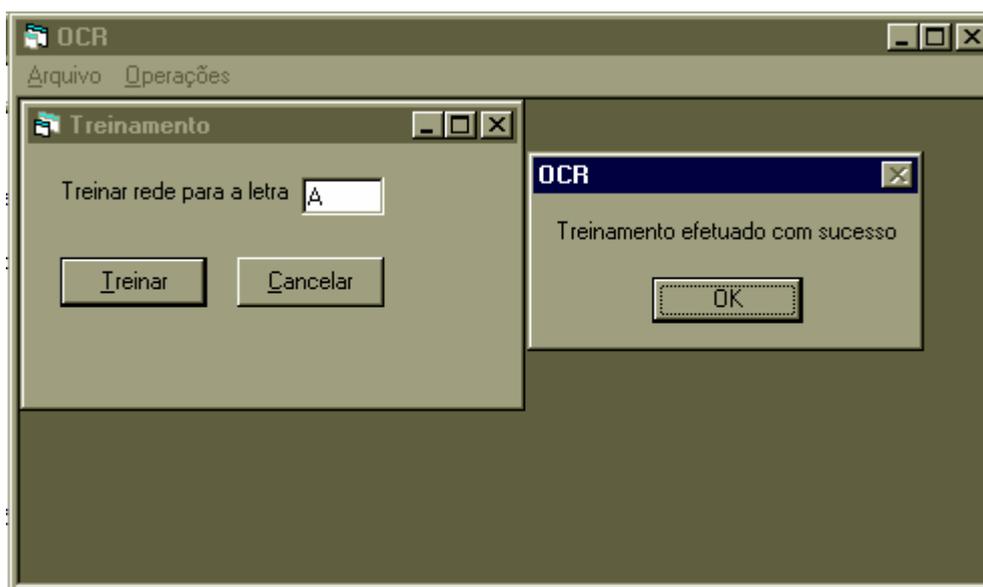
Figura 14 – Menu Operações

4.2.2.2.1 TREINAR REDE

Caso não exista uma matriz de pesos para iniciar o processo, o protótipo irá criar uma matriz com os pesos iguais a 1 (um), a partir de onde será iniciado o treinamento e ajuste dos pesos da rede.

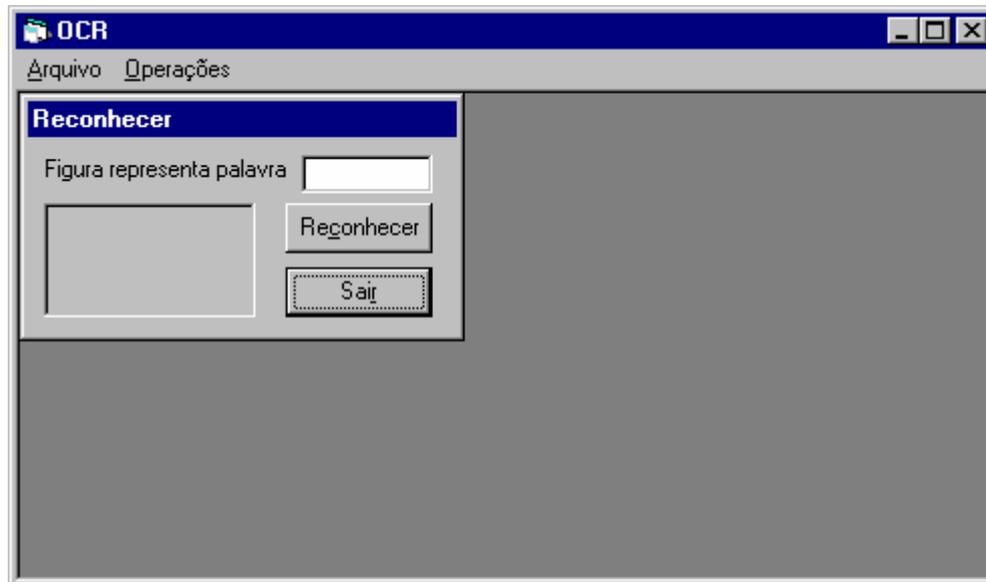
Para efetuar o treinamento da rede, deve ser informado qual a letra que se deseja treinar. Feito isso, basta pressionar o botão “Treinar” para iniciar o treinamento da rede. Ao concluir o treinamento da letra selecionada, o sistema irá informar que concluiu o processo de treinamento com sucesso. Não é necessário salvar a matriz de pesos, uma vez que o protótipo irá efetuar o salvamento da mesma após o treinamento com sucesso de cada letra (figura 15).

Figura 15 – tela de treinamento da palavra

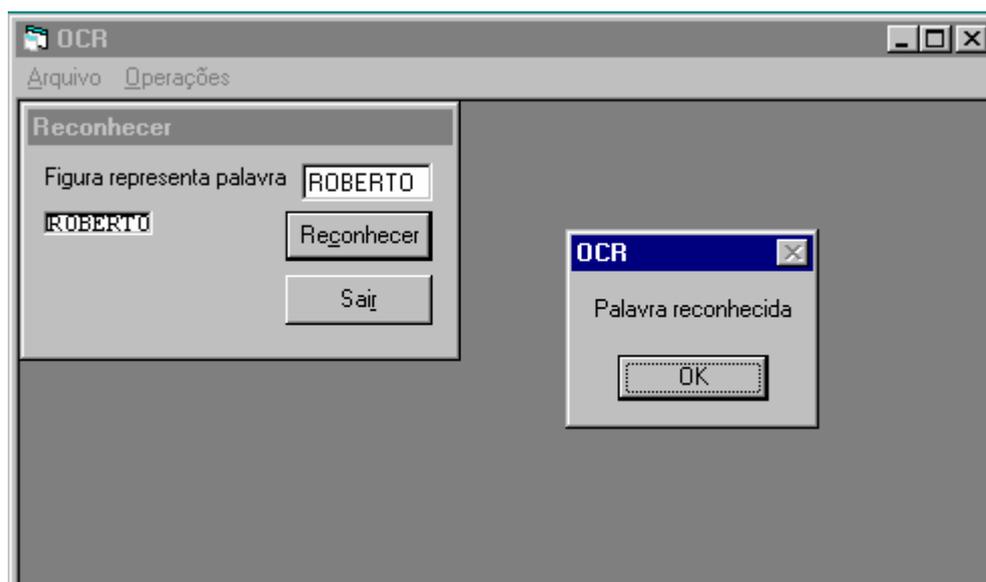


4.2.2.2.2 RECONHECER PALAVRA

Ao selecionar a opção “Reconhecer”, a tela para realizar o reconhecimento das palavras será exibida (figura 16).

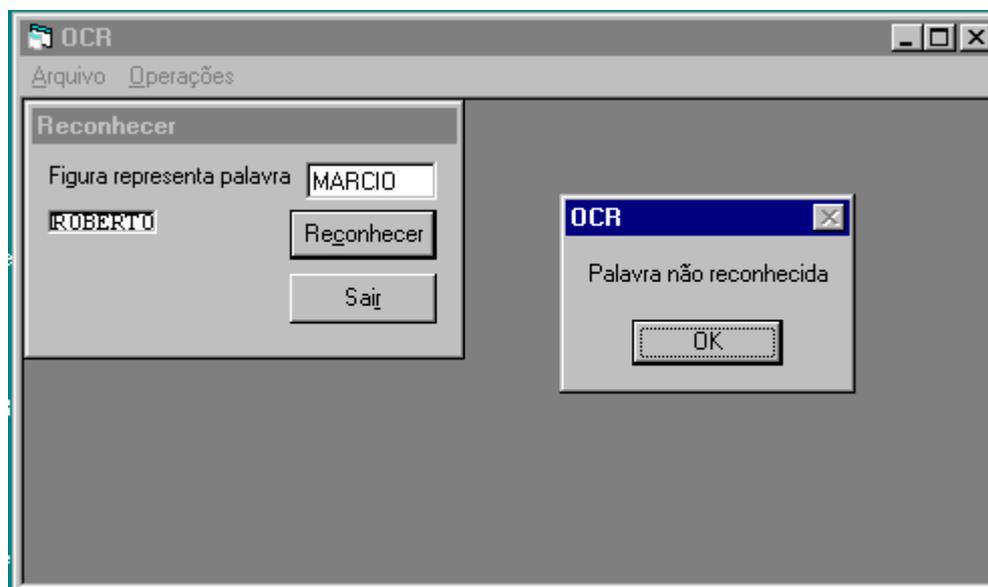
Figura 16 – tela de reconhecimento da palavra

Para efetuar o reconhecimento, deve primeiro ser selecionado uma figura com a palavra a ser reconhecida, através do menu “Arquivo → Abrir”. Após isso, basta pressionar o botão “Reconhecer” para submeter a figura selecionada ao tratamento e reconhecimento. Ao concluir o reconhecimento, o protótipo irá exibir uma mensagem com a palavra reconhecida (figura 17).

Figura 17 – Palavra reconhecida

Caso seja utilizada alguma figura não correspondente ao padrão de fonte determinado no protótipo, ou a palavra informada não corresponde a imagem selecionada, o protótipo exibirá uma mensagem informando que não conseguiu reconhecer a figura selecionada (figura 18).

Figura 18 – palavra não reconhecida



4.2.3 RESULTADOS OBTIDOS

Em virtude do ambiente altamente controlado que foi utilizado durante o desenvolvimento do protótipo, a rede neural não apresenta taxas de erro durante a fase de reconhecimento. Assim, para uma avaliação do desempenho da rede, segue na tabela abaixo o número de iterações necessárias para o aprendizado de cada letra (Quadro 4).

Quadro 4 – número de iterações para cada letra

A	B	C	D	E	F	G	H	I	J	K	L	M
2	3	3	6	5	6	5	12	8	10	7	13	9

N	O	P	R	S	T	U	V	W	X	Y	Z
10	7	21	10	13	14	9	9	7	18	10	12

5 CONCLUSÃO

Neste trabalho foi ilustrada a utilização das redes neurais para o reconhecimento de caracteres. Durante o desenvolvimento do trabalho, concluiu-se que a fase de segmentação é uma das etapas mais críticas, pois, qualquer problema nessa etapa irá afetar diretamente o desempenho do protótipo.

Durante o decorrer do trabalho, foram observadas vantagens e desvantagens do uso de redes neurais. A principal vantagem observada foi a característica de auto-aprendizado, já que as redes neurais não necessitam de conhecimentos de especialistas para tomar decisões, baseando-se unicamente nos exemplos que lhe são fornecidos.

Entre as desvantagens observadas das redes neurais, destacam-se:

- a) caixa preta: segundo [SIL1999] é impossível saber porque a rede chegou a uma determinada conclusão; os pesos utilizados não aceitam uma interpretação lógica: sabe-se apenas que funcionam;
- b) trabalho artesanal: não há regras gerais para se determinar o volume de dados para treinamento, quantas camadas utilizar, a melhor estratégia de treinamento, enfim, como desenvolver a rede: esses parâmetros só podem ser estabelecidos através de experiências com redes e através de tentativa-e-erro.

Considera-se que o objetivo do trabalho, desenvolver um protótipo de software para reconhecimento de caracteres utilizando redes neurais, foi alcançado. No entanto, por se tratar de um protótipo, existem várias melhorias que podem ser implementadas, como descrito abaixo.

5.1 LIMITAÇÕES

O protótipo desenvolvido apresenta as seguintes limitações:

- a) só reconhece letras em fonte *Courier New*, tamanho 12;
- b) não foram reconhecidos caracteres com acentuação;
- c) não reconhece letras minúsculas;
- d) números e caracteres especiais não foram reconhecidos;
- e) não foi efetuado o cálculo do limiar para geração do histograma. Sendo assim, não

foi gerado um histograma.

Apesar das limitações que foram apresentadas, o protótipo desenvolvido serviu para validar os algoritmos e idéias apresentadas pelos autores referenciados.

5.2 EXTENSÕES

Como extensões possíveis para este trabalho, pode-se citar:

- a) reconhecimento dos caracteres minúsculos, caracteres especiais e números, que não são reconhecidos hoje;
- b) reconhecer diversos tipos de letras, não só a especificada hoje;
- c) aprimorar os algoritmos para o pré-processamento da imagem, de forma que o limiar seja calculado, melhorando a qualidade do reconhecimento.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- [CAS1983] CASEY, Richard G.; JIH, Centung Robert. *A processor-based OCR system*. **IBM Journal of Research and Development**. Armonk, v.27, n.4, p386-399, Jul. 83.
- [FAC1993] FACON, Jacques. **Processamento e análise de imagens**. Córdoba : CEFET, 1993.
- [FAU1994] FAUSSET, Laurene. *Fundamentals of neural networks*. New Jersey : Prentice Hall, 1994.
- [HAR1983] HARALICK, Robert M. *Pictorial Data Analysis*. Berlin : Springer, 1983
- [KAR1996] KARTALOPOULOS, Stamatios V. *Understanding neural networks and fuzzy logic: basic concepts and applications*. New Jersey : IEEE Press. 1996.
- [KLA1996] KLABUNDE, Charles Cristiano. Trabalho de Conclusão de Curso. Curso de Ciências da Computação. **Aplicação do modelo de rede neural RBF-Fuzzy ARTMAP para reconhecimento de fala**. Blumenau : FURB, 1999.
- [LOE1996] LOESCH, Cláudio; SARI, Solange Teresinha. **Redes neurais artificiais: fundamentos e modelos**. Blumenau : FURB, 1996.
- [LUN1999] LUNA, Paulo de Tarso Mendes. Notas de Aula.
- [MCK1997] MCKELVY, Mike; et al. **Usando Visual Basic 5**. Rio de Janeiro : Campus, 1997.
- [PAR1993] PARKER, J. R. **Practical computer vision using C**. USA : Wiley Professional Computing, 1993.
- [OSO1991] OSÓRIO, Fernando Santos. Dissertação submetida à Universidade Federal do Rio Grande do Sul para obtenção do Grau de Mestre em Computação. **Um**

estudo sobre o reconhecimento visual de caracteres através de redes neurais. Porto Alegre : UFRGS, 1991.

- [SAN1999] SANTOS, Reinaldo Fischer dos. Trabalho de Conclusão de Curso. Curso de Ciências da Computação. **Aplicação de uma nova proposta de pré-processamento utilizando o modelo de rede neural RBF-Fuzzy ARTMAP para reconhecimento de fala.** Blumenau : FURB, 1996.
- [SCH1997] SCHRAMM, Cátia Regina. Trabalho de Conclusão de Curso. Curso de Ciências da Computação. **Desenvolvimento de um protótipo de um software de reconhecimento de caracteres.** Blumenau : FURB, 1997.
- [SIL1999] SILVA, Alex Sandro da. Trabalho de Conclusão de Curso. Curso de Ciências da Computação. **Protótipo de Software para Classificação de Impressão Digital.** Blumenau : FURB, 1999.
- [TAF1995] TAFNER, Malcon; XEREZ, Marcos de; RODRIGUES FILHO, Ilson W.. **Redes Neurais Artificiais.** Blumenau : FURB, 1995.
- [TUR1993] TURBAN, Efraim. *Decision support and expert systems.* New York : MacMillan, 1993.
- [WIL1995] WILHELM, Pedro Paulo Hugo; LOPES, Maurício Capobianco; *et al.* Sistema Inteligente de Apoio à Decisão. **Revista de Negócios.** Blumenau, v.1 n.1, Dezembro de 1995.