

Renato Ramos da Silva

**RECONHECIMENTO DE IMAGENS DIGITAIS UTILIZANDO REDES
NEURAS ARTIFICIAIS**

Monografia de Graduação apresentada ao Departamento de Ciência da Computação da Universidade Federal de Lavras como parte das exigências do curso de Ciência da Computação para obtenção do título de Bacharel

Orientador
Prof. Mário Luiz Rodrigues Oliveira

Lavras
Minas Gerais - Brasil
2005

Renato Ramos da Silva

**RECONHECIMENTO DE IMAGENS DIGITAIS UTILIZANDO REDES
NEURAS ARTIFICIAIS**

Monografia de Graduação apresentada ao Departamento de Ciência da Computação da Universidade Federal de Lavras como parte das exigências do curso de Ciência da Computação para obtenção do título de Bacharel

Aprovada em *20/01/2005*

Prof. Guilherme Bastos Alvarenga

Prof. Mário Luiz Rodrigues Oliveira
(Orientador)

Lavras
Minas Gerais - Brasil

Dedico todo este trabalho a Deus e a santíssima mãe Maria que sempre estão ao meu lado

Agradecimentos

Agradeço aos meus pais pela minha vida e estarem sempre ao meu lado dando todo apoio para que isto pudesse ser realizado. Por todos os professores que contribuíram na minha formação, de um modo maior pela confiança dos professores Joaquim, Guilherme e Mario na minha pessoa. Por todos os colegas de classe por esses anos passados. Aos amigos da Crediesal e Devex, especialmente ao Robson, Julio e Ulisses, por cada dia de convivência. Finalmente, agradeço de modo especial por cada irmão do Projeto Universidades Renovadas que estiveram juntos nessa caminhada, e permaneceremos .

Resumo

RECONHECIMENTO DE IMAGENS DIGITAIS UTILIZANDO REDES NEURAS ARTIFICIAIS

O presente projeto estudou as etapas do processamento digital de imagens, a saber: aquisição da imagem, pré-processamento, segmentação, representação e descrição, reconhecimento e interpretação. Foi implementada uma rede neural artificial utilizando o algoritmo de treinamento backpropagation para realização do reconhecimento de objetos em imagens digitais.

RECOGNITION OF DIGITAL IMAGES USING ARTIFICIAL NEURAL NETWORKS

The present project studied the stages of the digital image processing, to know: image acquisition , preprocessing, segmentation, representation and description, recognition and interpretation. The project implemented a artificial neural networks using the training algorithm backpropagation to carry out of objects recognition in digital images.

Sumário

1	Introdução	1
1.1	Motivação	2
1.2	Objetivo	2
1.3	Escopo do trabalho	2
2	Imagem digital	5
2.1	Introdução	5
2.2	Definição de Imagem Digital	5
2.3	Processamento digital de imagens	6
2.3.1	Aquisição de imagem	7
2.3.2	Pré-processamento	7
2.3.3	Segmentação	8
2.3.4	Representação e descrição	8
2.3.5	Reconhecimento	9
2.3.6	Interpretação	9
3	Método Estatístico	11
3.1	Introdução	11
3.2	Método Estatístico no reconhecimento de padrões	11
3.3	Classificação por máxima verossimilhança	12
3.4	Distância Euclidiana	12
3.5	Aproximação Bayesiana	12
4	Redes Neurais Artificiais	15
4.1	Introdução	15
4.2	Definição de Redes Neurais Artificiais	15
4.3	Características de uma Rede Neural Artificial	17

4.4	Backpropagation	18
5	Metodologia	19
5.1	Introdução	19
5.2	Algoritmo implementado	19
5.3	Linguagem Utilizada	21
5.4	Teste dos Algoritmos	22
6	Resultados e Discussões	23
6.1	Introdução	23
6.2	Grupos de imagens	23
6.3	Classificação por grupo	24
6.4	Classificação nos Grupos	25
6.4.1	Grupo 1	25
6.4.2	Grupo 2	25
6.4.3	Grupo 3	26
6.5	Tratamento de imagens	26
7	Conclusões	29
8	Trabalhos Futuros	31

Lista de Figuras

2.1	Representação numérica de uma imagem ampliada de 10x10 pixels com 256 tons de cinza.	6
2.2	Elementos do processo de análise da imagem [Augusto (2002)] . .	7
4.1	Representação artificial de um neurônio natural. As entradas x_1, x_2, \dots, x_n , representam os sinais vindos de outros n neurônios. Os pesos $w_{k1}, w_{k2}, \dots, w_{kn}$ representam os pesos sinápticos das conexões entre os neurônios da camada anterior e da camada k . O Bias (b_k) é o limiar para a ativação do neurônio [Cordeiro (2002)].	16
5.1	A esquerda contém o arquivo Valores.txt a direita o arquivo Image.txt	20
5.2	Exemplo do arquivo de configuração da rede neural	21
5.3	Arquivo de saída gerado pela rede neural	21
6.1	Imagens do Grupo 1: Circulo e Avião	23
6.2	Imagens do Grupo 2: Esfera e Árvore	24
6.3	Imagens do Grupo 3: Árvore sombreada e Face	24
6.4	Imagens das Figuras após pré-tratamento	27

Lista de Tabelas

6.1	Resultados de Grupos de Imagens	25
6.2	Resultados do Grupo 1	25
6.3	Resultados do Grupo 2	26
6.4	Resultados do Grupo 3	26

Capítulo 1

Introdução

A todo momento estamos fazendo reconhecimento de padrões em nossas vidas. Neste exato instante, ao ler este texto, você reconhece estes traços pretos como letras, reconhece um grupo de letras como uma palavra e reconhece o significado de cada palavra. Enfim, reconhecer o rosto de uma pessoa, distinguir um cachorro de um gato, compreender a fala, ler as mais diversas caligrafias e até mesmo interpretar um exame de eletrocardiografia, tudo isso é reconhecimento de padrões [Henrique (2003)].

Reconhecimento de padrões (RP) é, por natureza, uma ciência inexata, e assim admite muitas abordagens, algumas vezes complementares, outras competitivas, para aproximar, indicar ou encontrar soluções a um dado problema. Em problemas complexos de RP, a abordagem comum utilizada é dividir o problema em dois módulos subseqüentes: a extração de características e o módulo de classificação. Tanto para o primeiro como para o segundo módulo já existem várias técnicas desenvolvidas que vão desde técnicas estatísticas e matemáticas até técnicas de Inteligência Artificial ([Artur (1999)], [Bishop (1995)]).

Um dos ramos da ciência da computação que se ocupa da automação do comportamento inteligente é a inteligência artificial, onde nós, estudantes e pesquisadores, buscamos entender e explorar os mecanismos da mente possibilitando o pensamento e a ação inteligente.

E um desses mecanismo utiliza para realizar uma tarefa particular ou função de interesse, a maneira como o cérebro realiza e é conhecido como rede neural artificial.

1.1 Motivação

Há uma enorme quantidade de aplicações para a classificação de objetos. Dentre as diversas utilidades, pode-se destacar:

- análise de radiografias, diagnóstico de doenças, identificação de células, simulação de funções cerebrais;
- caracterização de rochas, prospecção mineral, sensoriamento remoto;
- detecção de alvos, classificação de sinais de radar;
- efeitos especiais, animação;
- visão computacional, controle de manipuladores, análise de situações;
- simulação de processos químicos, processamento de sinal, diagnóstico de falhas;
- análise de texturas e fases.

1.2 Objetivo

O trabalho visou atender aos seguintes objetivos:

- Estudar todas as etapas do processamento digital de imagens;
- estudar as técnicas utilizadas no reconhecimento de imagens digitais;
- estudar a utilização das redes neurais artificiais no reconhecimento de imagens digitais.

1.3 Escopo do trabalho

A seguir é apresentada uma breve descrição dos capítulos deste trabalho. O capítulo dois apresenta uma descrição sobre imagens digitais e como é realizado o seu processamento. O capítulo três mostra a classificação através dos métodos estatísticos. Em seguida, o capítulo quatro apresenta as redes neurais sob o ponto de vista do reconhecimento de imagens. O capítulo cinco apresenta a metodologia usada na constituição desta pesquisa. O capítulo seis os resultados desta pesquisa.

O capítulo sete apresenta as principais conclusões referentes ao trabalho bem como algumas propostas de trabalhos futuros e algumas contribuições. E para finalizar as referências bibliográficas.

Capítulo 2

Imagem digital

2.1 Introdução

O capítulo apresenta uma definição sobre imagem e sua forma digital. Descreve também todo o processamento de imagens digitais.

2.2 Definição de Imagem Digital

Segundo [Gonzales and Woods (1992)], uma imagem pode ser definida como uma função $f(x, y)$, onde o valor nas coordenadas espaciais x e y corresponde ao brilho (intensidade) da imagem nessa coordenada.

A única forma de se representar uma imagem em um computador é quando ela está digitalizada tanto no domínio espacial como no das amplitudes.

Uma imagem digital é a representação numérica e discreta de um objeto, ou especificamente, é uma função quantificada e amostrada, de duas dimensões, geradas por meios ópticos, disposta em um grade padrão, retangular igualmente espaçada, quantificada em iguais intervalos de amplitude. Assim, uma imagem digital é um vetor retangular bidimensional de amostras de valores quantificados ([Cordeiro (2002)], [Castleman (1996)]).

A menor unidade de uma imagem digital é denominada picture element (pixel). Um pixel é a representação numérica da luminosidade de um ponto da imagem.

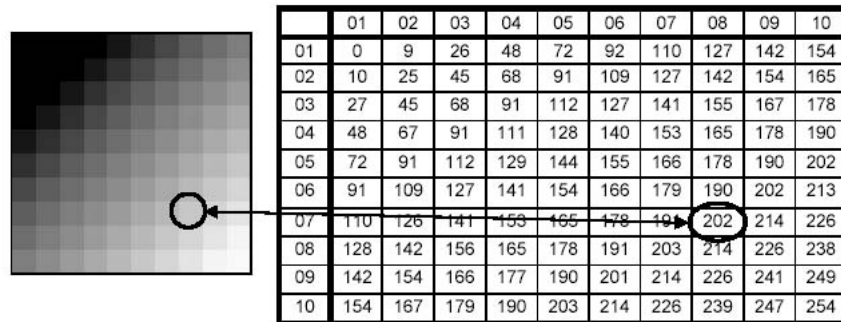


Figura 2.1: Representação numérica de uma imagem ampliada de 10x10 pixels com 256 tons de cinza.

2.3 Processamento digital de imagens

Segundo [Maria (2000)], entende-se por Processamento Digital de Imagens a manipulação de uma imagem por computador de modo que a entrada e a saída do processo sejam imagens. O objetivo de se usar processamento digital de imagens é melhorar o aspecto visual de certas feições estruturais para o analista humano e fornecer outros subsídios para a sua interpretação, inclusive gerando produtos que possam ser posteriormente submetidos a outros processamentos.

As técnicas em análise de imagem podem ser divididas em três áreas básicas: (1) processamento de baixo nível, com funções que podem ser vistas como reações automáticas, ou seja, reações que não requerem comportamento inteligente; (2) processamento de nível intermediário, com processos de extração e caracterização de componentes em uma imagem; e (3) processamento de alto nível, que envolve os processos de reconhecimento e interpretação. A Figura 2.2 mostra os processos de cada uma dessas áreas [Gonzales and Woods (1992)].

Nas seções de 2.3.1 a 2.3.6 serão apresentado os passos para o processamento digital de imagens com enfoque no reconhecimento de imagens. São eles:

- Aquisição de imagem;
- Pré-processamento;
- Segmentação;
- Representação e descrição;

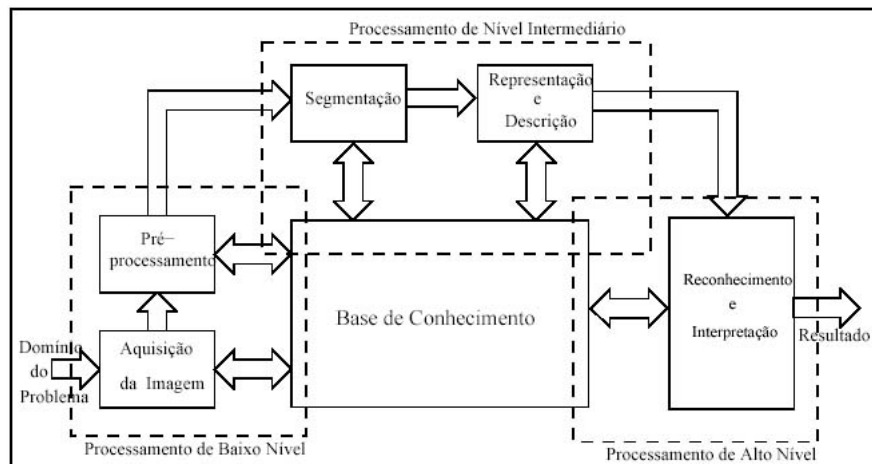


Figura 2.2: Elementos do processo de análise da imagem [Augusto (2002)]

- Reconhecimento;
- Interpretação.

2.3.1 Aquisição de imagem

Segundo [Gonzales and Woods (1992)], o primeiro passo do processo requer apenas um sensor de imagens e a capacidade para digitalizar o sinal produzido pelo sensor.

São necessários dois elementos para a aquisição da imagem: um aparelho físico que é sensível à faixa espectral de energia eletromagnética na qual produza um sinal elétrico de saída proporcional ao nível de energia sentido. E um digitalizador, que converte o sinal elétrico capturado na sua forma digital.

2.3.2 Pré-processamento

Após a aquisição e digitalização da imagem, o próximo passo é o pré-processamento. A função chave do pré-processamento é melhorar a imagem, com o objetivo de aumentar as chances de sucesso dos processos seguintes [Gonzales and Woods (1992)]. Nesta etapa, são utilizadas técnicas para aumento de contraste, remoção de ruídos, realce e normalização, com o objetivo de converter os padrões para uma

forma que possibilite uma simplificação do posterior processo de reconhecimento [Augusto (2002)].

2.3.3 Segmentação

O próximo estágio é o processo chamado de segmentação. De um modo geral, a segmentação subdivide uma imagem de entrada em suas partes constituintes ou objetos. Cada uma destas partes é uniforme e homogênea com respeito a algumas propriedades da imagem, como por exemplo, cor e textura. Algoritmos de segmentação para imagens monocromáticas são geralmente baseados em duas propriedades básicas de valores em escala de cinza: descontinuidade e similaridade. Na primeira categoria, o particionamento da imagem é baseado no subconjunto de pontos de um objeto que o separa do restante da imagem. As técnicas de segmentação nesta categoria buscam evidenciar os limites entre os objetos, através da detecção de pontos isolados e da detecção de linhas e bordas na imagem. Na segunda categoria, a segmentação é baseada nas técnicas de limiarização, crescimento por regiões, união e divisão de regiões ([Augusto (2002)], [Gonzales and Woods (1992)]).

2.3.4 Representação e descrição

Geralmente, a saída do estágio de segmentação são dados brutos de pixel. Neste caso pode ser necessário converter os dados para uma forma conveniente, possibilitando o processamento por computador. Dois tipos de representação podem ser utilizados: representação limite ou representação regional. A representação limite é apropriada quando o foco está em características da forma, como por exemplo em cantos. Representação regional é apropriada quando o foco está em propriedades reflectivas, tais como textura ou cor. No entanto, em algumas aplicações estas representações coexistem. Escolher a representação é apenas parte da solução para a transformação de dados brutos em uma forma conveniente para o processamento computacional subsequente. Um método para descrever os dados tal que as características de interesse sejam realçadas, também deve ser utilizado. Descrição, também chamada de seleção de característica, lida com a extração de características que resultam em algum diferenciar uma classe de objetos de outra [Gonzales and Woods (1992)].

2.3.5 Reconhecimento

Após feita separação de classe de objetos com características semelhantes passamos a uma fase na qual tem interesse identificar o que cada uma dessas classes representa, e assim identificá-las com um respectivo valor.

O reconhecimento de padrões é importante devido às ocorrências na vida humana tomarem forma de padrões. A formação da linguagem, o modo de falar, o desenho das figuras, o entendimento das imagens, tudo envolve padrões. Reconhecimento de Padrões é uma tarefa complexa, onde o homem busca, sempre, avaliar as situações em termos dos padrões das circunstâncias que as constituem, descobrir relações existentes no meio, para melhor entendê-lo e adaptar-se [Artur (1999)].

Existem, hoje, muitas técnicas de RP desenvolvidas no qual podemos dividir em três abordagens:

- Estatística - conjuntos de medidas de características (na forma de n-tupla ou vetores) são extraídos das imagens e métodos estatísticos são utilizados para separar as classes. Dentre os métodos utilizados podemos citar classificadores bayesianos, métodos probabilísticos e regras de decisão.
- Estrutural - padrões são representados em uma forma simbólica (tais como strings e árvores), e os métodos de reconhecimento são baseados em casamento de símbolos ou em modelos que tratam padrões de símbolos como sentenças, a partir de uma linguagem artificial;
- Neural - na abordagem neural, como o próprio nome diz, o reconhecimento é realizado utilizando-se Redes Neurais Artificiais. Alguns autores consideram o reconhecimento via Redes Neurais como sendo um tipo particular de reconhecimento estatístico, já que as características também são na forma de n-tuplas ou vetores e existe uma equivalência entre alguns modelos de Redes Neurais e técnicas estatísticas fundamentais ([Augusto (2002)], [Bishop (1995)]).

2.3.6 Interpretação

Para [Gonzales and Woods (1992)], interpretação envolve a fixação de significado a um grupo de objetos reconhecidos, em outras palavras, estamos interessados em dar significados a imagem.

Interpretação de imagens computadorizadas é um processo extremamente complexo. As dificuldades aparecem tanto pela quantidade de dados a serem processadas e também pela falta de ferramentas de processamento fundamental para receber o dados iniciais para gerar os resultados desejados (detalhamento do conteúdo

da imagem). Assim como não há ferramentas para realizar interpretação de imagens não estruturada, somos forçados a considerar aproximações que oferecem razoável chance de sucesso. Essa limitação leva-nos a considerar duas restrições: limitar a generalidade do problema e a incorporação de conhecimento humano no processo.

Os três principais formalismos utilizados para a representação do conhecimento são baseados em:

- **lógica formal** - onde o conhecimento pode ser expresso através de regras lógicas;
- **redes semânticas** - que através de grafos direcionados nos demonstra o relacionamento entre elementos de uma imagem;
- **sistema de produção ou regra base** - que apesar de ser restrito a um determinado domínio, é utilizado numa larga escala de aplicativo em processamento de imagem digital.

Capítulo 3

Método Estatístico

3.1 Introdução

No capítulo contém um explicação sobre métodos estatísticos e alguns métodos utilizados na classificação de objetos.

3.2 Método Estatístico no reconhecimento de padrões

[Schalkoff (1992)] diz que o reconhecimento de padrões estatístico, como sugere o próprio nome, assume uma base estatística para os algoritmos de classificação. Um conjunto de medidas, que denotam as características, é extraído dos dados de entrada e usado para associar cada vetor de características a uma dentre classes. Presume-se que as características são geradas por um estado da natureza, e, portanto, existe um modelo subordinado a um estado com um conjunto de probabilidades e/ou funções de densidade de probabilidade correspondente, passível de ser usado para representá-lo.

A forma mais geral e natural de formular soluções para o RP é o Reconhecimento de Padrões estatístico, através do qual é reconhecida a natureza estatística tanto da informação que se quer representar quanto dos resultados que devem ser expressos ([Bishop (1995)], [Artur (1999)]).

Existem vários métodos estatísticos que apresentam um bom resultado na área de reconhecimento de imagens. Dentre eles estão Classificação por máxima verossimilhança, distância euclidiana [Maria (2000)], aproximação Bayesiana, curva normal, *parzen window* e vizinhança mais próxima [Henrique (2003)].

Nas seções 3.1, 3.2 e 3.3 apresenta os métodos mais utilizados para a classificação de objetos em imagens digitais.

3.3 Classificação por máxima verossimilhança

Segundo [Maria (2000)], Classificação por máxima verossimilhança é o método de classificação pixel a pixel mais utilizado em dados de sensoriamento remoto.

Diferentes tipos de equações podem ser utilizados para aproximar a densidade de probabilidade das classes de um sistema de reconhecimento de padrões. Para cada classe deverão ser encontrados diferentes parâmetros para a equação de aproximação utilizada. O método de aproximação por máxima verossimilhança é uma maneira de encontrar os parâmetros que melhor aproximarão a densidade de probabilidade, com base em um conjunto de amostras. Por ser paramétrica, a densidade de probabilidade fica representada por $p(x|c, \nu)$, sendo ν o vetor de parâmetros. Cada classe é tratada independentemente uma da outra; ou seja, a densidade de probabilidade de uma classe não influencia na outra; assim, suas amostras são variáveis aleatórias independentes e identicamente distribuídas [Henrique (2003)].

3.4 Distância Euclidiana

Nesse método, as amostras de treinamento são usadas para calcular a média de cada classe. Cada pixel será atribuído a classe cuja distância entre o valor do pixel e a média da classe seja menor. O pixel será incorporado ao agrupamento que apresenta a menor distância Euclidiana. Este procedimento é repetido até que toda a imagem seja classificada [Maria (2000)].

O método é utilizado quando o número de amostras de treinamento por classe é limitado. O método da distância Euclidiana possui a vantagem de ser, computacionalmente, mais simples.

3.5 Aproximação Bayesiana

As soluções pela aproximação bayesiana são quase sempre idênticas às do método de máxima verossimilhança, sendo a diferença mais conceitual do que prática. Enquanto no método da máxima verossimilhança procura-se fixar o valor do vetor de parâmetros, na aproximação bayesiana o vetor de parâmetros é considerado uma variável aleatória [Henrique (2003)].

A análise dos resultados da aproximação bayesiana normalmente é mais complexa, comparada com a máxima verossimilhança. Isso porque, pela aproximação bayesiana, o resultado é uma distribuição de probabilidades do vetor de parâmetros, enquanto na máxima verossimilhança obtém-se um vetor de parâmetros [Henrique (2003)].

Capítulo 4

Redes Neurais Artificiais

4.1 Introdução

O capítulo defini rede neural artificial, o seu funcionamento e algumas aplicações.

4.2 Definição de Redes Neurais Artificiais

Um dos ramos da inteligência artificial que mais se desenvolveu ultimamente é o de redes neurais, também conhecido como conexista [Cordeiro (2002)]. Isso se deve a capacidade de fazer suposições mais delicadas a respeito da distribuição dos dados de entrada do que métodos estatísticos tradicionais e a capacidade de formar fronteiras de decisão altamente não-lineares no espaço de características, levou ao seu crescente uso [Artur (1999)].

Assim como na natureza, um neurônio artificial é uma estrutura relativamente simples, que responde a estímulos de outros neurônios conectados a ele. Essa arquitetura dá as redes neurais características marcantes de intenso paralelismo e robustez [Cordeiro (2002)].

Na maioria das aplicações de RP usando-se redes neurais, são estabelecidas conexões entre os valores de todas as características que definem os padrões com camadas intermediárias de neurônios e a todas as classes objetivos que são consideradas como a camada de saída. Dentre as aplicações que utilizam Redes Neurais, [Augusto (2002)] cita: reconhecimento de objetos [Tu and Li (1999)], mineração de dados [Craven and Shavlik (1997)], reconhecimento de fala [Yuk and Flanagan (1999)], robótica [de A. Barreto et al. (2001)].

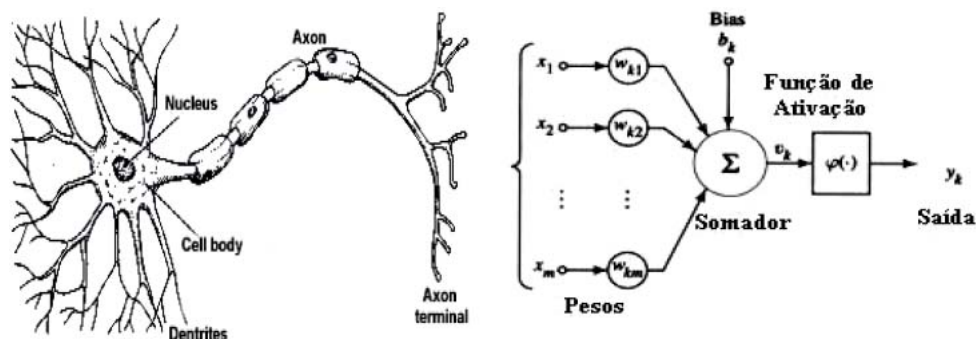


Figura 4.1: Representação artificial de um neurônio natural. As entradas x_1, x_2, \dots, x_n , representam os sinais vindos de outros n neurônios. Os pesos $w_{k1}, w_{k2}, \dots, w_{kn}$ representam os pesos sinápticos das conexões entre os neurônios da camada anterior e da camada k . O Bias (b_k) é o limiar para a ativação do neurônio [Cordeiro (2002)].

O treinamento da rede é realizado corrigindo os pesos nas conexões para estabelecer-se as relações entre as características e classes que promovam a melhor discriminação possível entre os padrões de classes diferentes. Assim, ao ser apresentado à rede um novo padrão, esta indicará a classe que o representa melhor na camada de saída [Artur (1999)].

O conhecimento pode ser representado na rede como padrões de atividade distribuída entre muitas unidades, e uma unidade pode tomar parte em diversos padrões diferentes. Em outras palavras, redes neurais artificiais combinam uma arquitetura altamente paralela com uma forma distribuída de representação de conhecimento. São formas de se extrair informação difusa ou indeterminada, a partir de um conjunto de dados previamente escolhido. São modelos indicados para tratar de sistemas abertos ou complexos, pouco conhecidos e que não podem ser adequadamente descritos por um conjunto de regras ou equações. São indicadas para aplicações que requeiram tolerância a falhas, detecção de padrões, diagnóstico, abstração ou generalização, onde haja dados com ruído, incompletos ou distorcidos ([Carlos (1994)], [Widrow (1990)]).

O conhecimento numa rede neural artificial não está armazenado em locais determinados. Não é possível olhar num endereço de memória e verificar o conteúdo de uma variável. O conhecimento está armazenado na topologia e nos pesos [Carlos (1994)].

Uma rede neural artificial é caracterizada por uma arquitetura, um método de

aprendizado ou treinamento, uma função de ativação e um bias.

4.3 Características de uma Rede Neural Artificial

A seção apresenta uma descrição das características básicas das redes neurais artificiais.

- Função de Ativação

Função usada para restringir a amplitude de saída de um neurônio. Tipicamente, o intervalo normalizado da amplitude de saída de um neurônio é escrito como o intervalo unitário fechado $[0,1]$ ou alternativamente $[-1,1]$ [Haykin (2001)].

- Aprendizagem

Aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural artificial são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre [Haykin (2001)].

Existe dois paradigmas de aprendizagem. O primeiro é conhecido como aprendizagem supervisionada na qual existe um mentor, pessoa ou professor, capaz de fornecer à rede neural uma resposta desejada para um vetor de treinamento inicial. O segundo é chamado de aprendizagem não-supervisionada, onde não há um crítico para supervisionar o processo de aprendizagem, neste caso é fornecido condições para realizar uma medida independente da tarefa da qualidade da representação cuja rede deve aprender.

A seção 4.3 apresenta um dos algoritmos de aprendizagem supervisionada, backpropagation, utilizado neste projeto. Os algoritmos Adaline, Cascade Correlation, Levenberg-Marquardt, Redes de Elman e Gradiente conjugado são exemplos de aprendizagem supervisionadas. Mapas de Kohonen e Redes ART são exemplos de algoritmos de aprendizagem não-supervisionada.

- Arquitetura

A maneira pela qual os neurônios de uma rede neural estão estruturados, sendo este intimamente ligado com o algoritmo de aprendizagem. Em geral, podemos identificar três classes de arquiteturas de rede fundamentalmente diferentes ([Haykin (2001)]):

1. **Redes Alimentadas Adiante com Camada Única** - rede neural artificial acíclica que possui uma camada de nós de entrada e uma camada de nós computacionais de saída.
2. **Redes Alimentadas Diretamente com Múltiplas Camadas** - rede neural artificial acíclica que possui uma camada de nós de entrada, uma ou mais camadas de nós computacionais ocultos e uma camada de nós computacionais de saída.
3. **Redes Recorrentes** - rede neural artificial que possui pelo menos um laço de realimentação e pelo menos uma camada de nós.

- Bias

O neurônio pode ainda apresentar um bias que tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, ou seja, o termo bias age como um peso extra nas conexões das unidades cuja entrada é sempre um.

4.4 Backpropagation

Algoritmo de retropropagação de erro que é baseado na regra de aprendizagem por correção de erro e é constituído pela rede alimentada diretamente com múltiplas camadas.

Ele é dividido basicamente em duas etapas: a propagação e a retropropagação. A propagação consiste na aquisição dos dados pela camada de entrada e sua propagação por toda rede, produzindo um conjunto de saída. Esse conjunto é comparado com valores desejados. Apartir dai começa o processo de retropropagação que consiste no ajustes dos pesos segundo a diferença produzida, para fazer com que a resposta da rede se aproxime do desejado.

Capítulo 5

Metodologia

5.1 Introdução

Este capítulo apresenta a metodologia utilizada no desenvolvimento deste trabalho.

Este projeto implementou uma rede neural treinada com o algoritmo Backpropagation para reconhecimento digital de imagens. Ele será baseado unicamente na classificação de pixel a pixel de determinados objetos escolhido pelo usuário.

A utilização do aplicativo wxPID, desenvolvido por [Arley (2003)] em projeto de conclusão de curso, na qual foi implementado alguns dos algoritmos clássicos em processamento digital de imagens para a realização do pré-processamento, segmentação e representação sobre imagens.

5.2 Algoritmo implementado

Foram feitas duas adaptações do software wxPID para a realização deste trabalho:

- Foi adicionado um evento no qual o usuário ao clicar sobre determinado objeto com o botão esquerdo do *mouse*, aparece uma janela na qual ele escolhe uma determinada representação para aquele objeto.
- Também foi adicionado um evento ao botão direito do *mouse* na qual o usuário salva cada pixel da imagem em um arquivo texto.

Para cada uma das alterações feitas é gerado um arquivo para a rede neural. O evento associado ao botão esquerdo do *mouse* gera o arquivo **Valores.txt** que é

constituído por valores indicativos do pixel(red, blue, green) e uma string associada ao pixel indicando a representação do objeto na imagem . O evento associado ao botão direito do *mouse* gera o arquivo **Image.txt**, fornecendo o número de pixel constituinte da imagem e os valores respectivos de cada pixels(red, blue, green) da imagem.

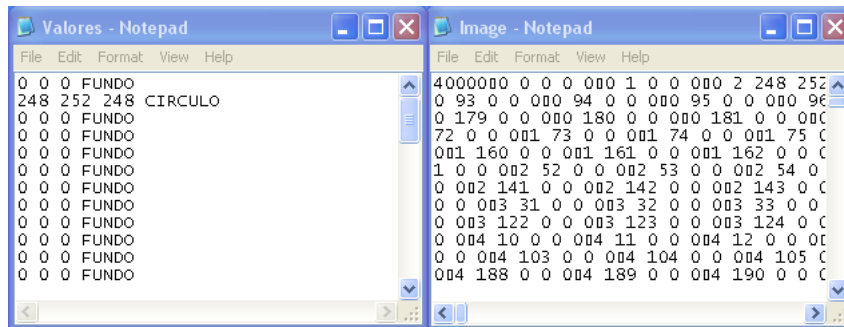


Figura 5.1: A esquerda contém o arquivo Valores.txt a direita o arquivo Image.txt

Para a configuração da arquitetura, definida como alimentada diretamente com três camadas, da rede neural o usuário deverá criar um arquivo texto contendo os valores indicados:

- Os três primeiros indicam o número de nós de cada camada (entrada, escondida e de saída).
- O último valor indica o número elementos passados a rede para treinamento.

Cada camada está armazenada em um vetor de neurônios, estrutura na qual é constituída por um valor e um vetor de pesos. Após configurada as camadas, ajustamos os valores de saída desejados de acordo com o número de classes. Ajustamos os bias e então é passado os valores iniciais para o treinamento da rede neural.

Na etapa de propagação, um determinado nó da camada oculta recebe de cada ligação da camada de entrada um valor, que é o valor inicial multiplicado pelo peso. Esses valores são somados passando a ser o valor do nó. Esse mesmo processo é feito para o nó de saída.

O resultado obtido é comparado com o desejado e a diferença existente é utilizada para o ajuste de pesos. Etapa essa conhecida como retropropagação e terminando aqui a etapa de treinamento.

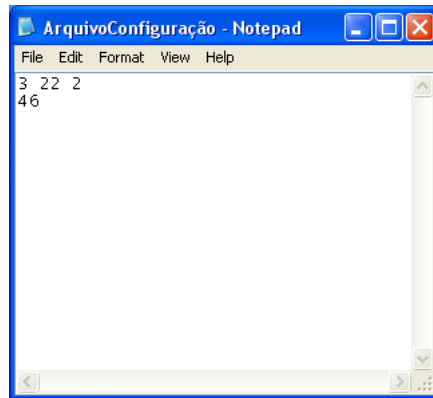


Figura 5.2: Exemplo do arquivo de configuração da rede neural

Com a rede neural configurada fazemos os reconhecimentos de cada objeto da figura, a partir da leitura do arquivo **Image.txt**. O resultado é passado para o arquivo **ResultadoBack.txt**, Figura 5.3, que contém a posição e o valor do pixel, e uma string associada, que indica a classificação feita pela rede neural.

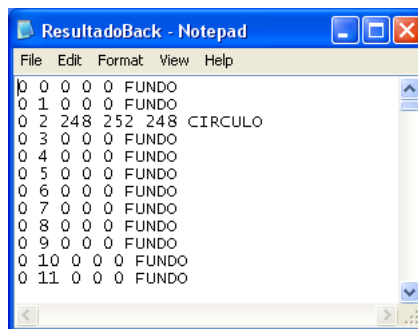


Figura 5.3: Arquivo de saída gerado pela rede neural

5.3 Linguagem Utilizada

A linguagem escolhida para execução do projeto foi C++. Dentre as principais características que motivaram tal escolha estão: código multiplataforma, paradigma de orientação à objeto, poder de processamento e manter um padrão com o projeto

já realizado por [Arley (2003)].

5.4 Teste dos Algoritmos

A avaliação do algoritmo implementado será feita a partir da análise do arquivo **ResultadoBack.txt**, observando se na posição indicada pela imagem o pixel é classificado corretamente. As imagens foram divididas em três grupos, cada grupo contendo duas imagens e a separação em grupos de acordo com complexidade da figura. Assim podemos verificar a ação da rede neural artificial entre os grupos de imagens e sobre as imagens contidas nos grupos.

Capítulo 6

Resultados e Discussões

6.1 Introdução

Este capítulo apresenta os resultados obtidos e as discussões relativos ao projeto.

6.2 Grupos de imagens

A seção tem por objetivo demonstrar as imagens escolhidas para a realização dos testes e um explicação deste agrupamento.

No grupo 1 foram separadas as imagens em preto e branco. Todas elas com um objeto apenas.

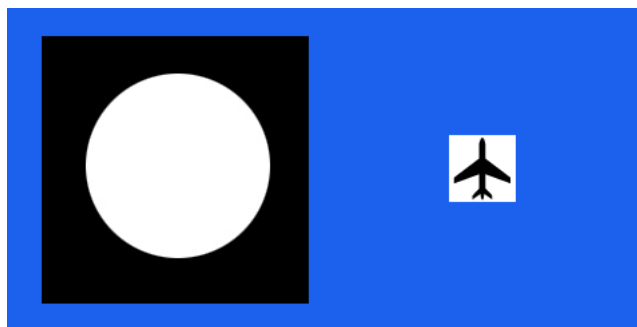


Figura 6.1: Imagens do Grupo 1: Circulo e Avião

No grupo 2 estão as imagens que possuem uma pequena variação de tom de

cinza entre o objeto e a figura e outra que possui uma variação de tons de cinza no objeto muito grande.

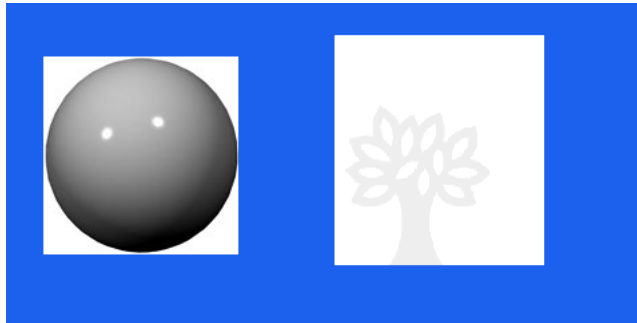


Figura 6.2: Imagens do Grupo 2: Esfera e Árvore

No grupo 3 estão contidas as imagens com mais de um objeto a serem classificados.



Figura 6.3: Imagens do Grupo 3: Árvore sombreada e Face

6.3 Classificação por grupo

Analisando-se cada pixel de resposta da rede neural artificial construiu-se uma tabela. A tabela é composta pelo grupo de imagens, número de pixels reconhecidos corretamente(NPRC), número de pixels reconhecidos incorretamente(NPRI) e a porcentagem de pixels reconhecidos corretamente em cada grupo(PPRCG).A tabela abaixo mostra claramente o comportamento da rede neural frente a cada grupo. E a partir desta tabela que iremos fazer a análise da rede neural.

Como pode ser visto pela tabela as imagens do grupo 1 e 2 obtiveram um bom resultado, e o principal fator no excelente resultado é a imagem não possui objetos

Grupo	NPRC	NPRI	PPRCG
1	44638	288	99,35
2	75542	946	98,76
3	74550	34440	68,40

Tabela 6.1: Resultados de Grupos de Imagens

similares, em contraposição ao ocorrido no grupo 3, na qual possui mais de um objeto similar a ser classificado.

6.4 Classificação nos Grupos

Apresenta-se nesta seção um comparação dos resultados obtidos em cada grupo. Em cada grupo a análise também será feita apartir de uma tabela. A tabela é composta pela imagem classificada, número de pixels reconhecidos corretamente(NPRC), número de pixels reconhecidos incorretamente(NPRI) e a porcentagem de pixels reconhecidos corretamente em cada imagem(PPRCI).

6.4.1 Grupo 1

A diferença nos resultado encontrado no grupo é causado pelo contorno do objeto. Como pode se ver a figura que cujo conteúdo é um avião possui uma superfície de contato maior com o fundo do que a figura do que contém um circulo.

Figura	NPRC	NPRI	PPRCI
Circulo	42204	230	99,45
Avião	2434	58	98,86

Tabela 6.2: Resultados do Grupo 1

6.4.2 Grupo 2

O resultado encontrado na classificação do objeto nas duas imagens demonstra que a variação dos tons de cinza de um objeto(esfera) em uma figura e a proximidade do valor do pixel de um objeto com o fundo da imagem tem pouca influência na classificação.

Figura	NPRC	NPRI	PPRCI
Esfera	21246	362	98,32
Árvore	54296	584	98,93

Tabela 6.3: Resultados do Grupo 2

6.4.3 Grupo 3

A diferença encontrada na avaliação das imagens do grupo 3 pode ser explicada pelo número de pixel que define determinado objeto. No caso da figura com o objeto da árvore sombreada existia apenas um pixel definindo determinado objeto o que não ocorre com os objetos da outra imagem.

Figura	NPRC	NPRI	PPRCI
Árvore sombreada	54382	20858	72,28
Face	20168	13582	59,76

Tabela 6.4: Resultados do Grupo 3

6.5 Tratamento de imagens

Toda a análise feita até o momento é relativo a classificação de imagens sem a nenhuma etapa previa, ou seja, foi feito somente o reconhecimento das imagens.

Esta seção tem por objetivo a análise do tratamento de imagens frente a classificação de objetos. Para isso a Figura nos mostra como fica duas de nossas imagens após o tratamento.

Ao realizar o tratamento os objetos passam a ser reconhecidos apenas pelo seu contorno. Para imagens do grupo 1 e 2 houve uma melhora pois os erros aconteciam na borda dos objetos. No entanto o grupo 3 com apenas um contorno, atribuindo determinado objeto, se tornou difícil a classificação.



Figura 6.4: Imagens das Figuras após pré-tratamento

Capítulo 7

Conclusões

O presente projeto verificou a eficiência da utilização das redes neurais no reconhecimento de imagens pixel a pixel, principalmente quando a imagem não possui objetos similares.

Dentre os fatores na qual influenciaram no reconhecimento dos objeto podemos verificar para os grupos 1 e 2, onde estão as imagens mais simples, a superfície de contato com o fundo como um fator relevante e no grupo 3 o número de objetos a ser classificado.

Nota-se também que o treinamento supervisionado se mostrou adequado para a realização do reconhecimento, esperando que o não-supervisionado mantenha a qualidade de classificação.

Capítulo 8

Trabalhos Futuros

Agora será passada algumas sugestões de trabalhos futuros para continuação do projeto:

- Métodos de classificação utilizando modelos estatísticos ou estrutural;
- Implementar uma rede neural não supervisionada;
- Classificação utilizando bordas;
- Desenvolver método de Interpretação;

Referências Bibliográficas

- [Augusto (2002)] Augusto, F. R. Localização e Reconhecimento de Placas de Sinalização Utilizando um Mecanismo de Atenção Visual e Redes Neurais Artificiais. Universidade Federal de Campina Grande, 2002.
- [Arley (2003)] Arley, A. C. Estudo e Implementação de Algoritmos Clássicos para Processamento Digital de Imagens, Lavras, 2003
- [Artur (1999)] Artur, J. S. Reconhecimento de Padrões Usando Indexação Recursiva. Universidade Federais de Santa Catarina, 1999.
- [Bezdek e Pal (1992)] BEZDEK, J. C. and PAL, S. K. (Editores) Fuzzy Models Pattern Recognition: Methods That Search for Structures in Data, IEEE, 1992.
- [Bishop (1995)] Bishop, C. M. Neural Networks for Pattern Recognition. Clarendon Press. Oxford, 1995.
- [Carlos (1994)] Carlos, J. F. P. Aplicação de Redes Neuronais no Processamento Digital de Imagens. Universidade Federal de Minas Gerais, 1994.
- [Castleman (1996)] Castleman, Kenneth R. Digital Image Processing. Upper Saddle River: Prentice Hall, Inc. 1996
- [Cordeiro (2002)] Cordeiro, F. M. Reconhecimento e Classificação de Padrões de Imagens de Núcleos Linfócitos do Sangue Periférico Humano com a Utilização de Redes Neurais Artificiais. Universidade Federal de Santa Catarina, 2002.
- [Craven and Shavlik (1997)] Craven, M. W. and Shavlik, J. W. Using neural networks for data mining. Future Generation Computer Systems, 1997.

- [de A. Barreto et al. (2001)] de A. Barreto, G., Araújo, A. F. R., Dücker, C., and Ritter, H. Implementation of a distributed robotic control system based on a temporal selforganizing neural network. In Proceedings of the IEEE International Conference on System, Man and Cybern (SMC'01), pages 335-340. Tucson, Arizona, 2001.
- [Duda e Hart (1973)] DUDA, O., HART, P. E. Pattern classification and scene analysis. John Wiley Sons, Inc., 1973.
- [Faussett (1994)] Fausett, L. V. Fundamentals of Neural networks: architectures, algorithms, and applications. Florida Institute of Technology, 1994.
- [Gonzales and Woods (1992)] GONZALES, R. C. WOODS, R. E. Digital Image Processing. University of Tennessee Perceptics Corporation, 1992.
- [Haykin (2001)] Haykin, S. Redes Neurais: Princípios e prática. McMaster University, Hamilton, Ontário, Canadá, 2001.
- [Henrique (2003)] Henrique, J. B. R. Desenvolvimento de uma técnica de reconhecimento de padrões baseada em distância. Universidade Federal de Santa Catarina, 2003.
- [Luger (2004)] Luger, G. F. Inteligência Artificial: estruturas e estratégias para a resolução de problemas complexos. University of New Mexico at Albuquerque, 2004.
- [Maria (2000)] MARIA, L. G. F. Processamento Digital de Imagens. INPE, Junho de 2000.
- [Pao (1989)] PAO, Y. Adaptive Pattern recognition and neural networks. Addison-Wesley, 1989.
- [Schalkoff (1992)] SCHALKOFF, Robert J. Pattern Recognition: Statistical, Structural and Neural Approaches. John Wiley Sons, Inc., 1992.
- [Tu and Li (1999)] Tu, Z. and Li, R. (1999). A multilayer hopfield neural network for 3-d object recognition. In Proceedings of International Mobile Mapping Workshop, pages 7A.3.1-6. Bangkok, Thailand, 1999.
- [Widrow (1990)] Widrow, B. and Lehr, M. A. 30 Years of adaptive neural network: Perceptron, Madaline and Backpropagation. Proceedings of IEEE, v. 78, n. 9, Sep. 1990.

[Yuk and Flanagan (1999)] Yuk, D. and Flanagan, J. Telephone speech recognition using neural network. In Proceedings of the IEEE International Conference on Acoustics and Signal Processing, volume 1, pages 157-160, 1999.